



# МАТЕМАТИКАЛЫҚ ЛОГИКА ЖӘНЕ КОМПЬЮТЕРЛІК ҒЫЛЫМДАР

Халықаралық ғылыми конференцияның еңбектері  
2022 жыл, 7-8 қазан, Астана, Қазақстан

# MATHEMATICAL LOGIC AND COMPUTER SCIENCE

Proceedings of the International scientific conference  
October 7-8, 2022, Astana, Kazakhstan

# МАТЕМАТИЧЕСКАЯ ЛОГИКА И КОМПЬЮТЕРНЫЕ НАУКИ

Труды Международной научной конференции  
7-8 октября 2022 г., Астана, Казахстан

АСТАНА  
2022  
ASTANA

**Математикалық логика және компьютерлік ғылымдар халықаралық ғылыми  
конференциясы**  
**Еуразия ұлттық университеті, Астана, ҚР**  
**7-8 қазан, 2022 жыл**

Конференция ғылымның қазіргі жай-күйін, математикалық логика мен компьютерлік ғылымдардың өзара ықпалын және заманауи тенденциялары мен мәселелерін таныстыру және талқылау мақсатында университетімізде өтті.

Осы форумда қарастырылған мәселелер мен нәтижелер әр жылдарда өткізілген «Модельдік теория бойынша қазақ-француз коллоквиумы» және «Модельдік теория бойынша француз-қазақ коллоквиумы» атты алты конференция, сондай-ақ 2019 жылы «Он алтыншы Азияның логика конференциясы» жұмыстарының жалғасы ретінде талқыланды.

Азияның логика конференциясы (АЛК) математикалық логикадағы ірі халықаралық оқиға болды. Онда математикалық логика және оның қосымшалары, компьютерлік ғылымдардағы логика және философиялық логика саласындағы соңғы ғылыми жұмыстар ұсынылды. Математикалық логика және компьютерлік ғылымдар бойынша өткізілген бұл конференция математикалық логика жетістіктерінің компьютерлік ғылымдардағы жасанды интеллект, сараптамалық жүйелер және т.б. бөлімдеріне зор әсерін көрсетеді, сонымен қатар, еуразия кеңістігінде, атап айтқанда Қазақстанда математикалық логика мен компьютерлік ғылымдарды ілгерілету және ақпарат пен ой алмасу үшін ғалымдарды біріктіруге бағытталған. Л.Н. Гумилев атындағы Еуразия ұлттық университеті ғылыми-зерттеу университеті атағын жақында ғана алған, математикалық логика және алгебра бойынша зерттеулер механика-математика факультетінің «Алгебра және геометрия» кафедрасында жүргізілуде, онда ҚР Ұлттық ғылым академиясының академигі Өмірбаев Уалбай, Қазақ КСР Ғылым академиясының академигі Тайманов Асан Дабсұлының шәкірттері, сонымен қатар Сібір алгебра және логика мектебінің өкілдері жұмыс жасайды.

Л.Н. Гумилев атындағы Еуразия ұлттық университетінің Ақпараттық технологиялар факультеті конференцияда жарияланған мамандықтар бойынша білікті ғалымдарды дайындайды, дегенмен, ақпараттық технологиялардың DATA Science, жасанды интеллект, компьютерлік лингвистика, ақпараттық қауіпсіздік, ақпараттық жүйелер және бағдарламалық инженерия сияқты заманауи бағыттары бойынша мамандарды даярлауды күшейту қажет.

Конференцияда семантикалық және логикалық бағдарламалау бойынша үш баяндамадан тұратын пул ұсынылды. Конференция осы ғылымның математикалық логика жетістіктеріне негізделген теориялық негіздеріне бағытталған баяндамалармен ашылды: «Көпмүшенің есептелуінің логикалық-математикалық сипаттамасы» тақырыбында дайындаған біріккен баяндамасымен академик С.С. Гончаров және ф.-м.ғ. докторы Свириденконың (РҒА СБ Математика институты), осы идеяларды ақпараттық өнім түрінде жүзеге асыру туралы «Онтологиялық интеллектуалдық мәліметтерді игеру» баяндамасымен ф.-м.ғ. докторы Е.Е. Витяев (РҒА СБ Математика институты) және «Онтологиялар және микросервис объектіге бағытталған бағдарламалау» баяндамасымен ф.-м.ғ. докторы А.В. Манцивода (Иркутск мемлекеттік университеті) жалғастырды. Бұл баяндамалар үлкен қызығушылық тудырды және конференцияға қатысушыларды өздерінің іргелі сипатымен, кең ауқымымен және өмірдің нақты салаларында, мысалы, компьютерлік ғылымдар, экономика, медицина және т.б. енгізуімен таң қалдырды.

Математикалық логика, алгебра және сандар теориясы бағытында мынадай баяндамалар ұсынылды: «Полурешетки Роджерса» тақырыбында С.А. Бадаев (Қазақ-Британ техникалық университеті, Алматы), сақина теориясының өзекті мәселесінің шешімі туралы Қазақстан Республикасы Ұлттық ғылым академиясының академигі

У.Өмірбаев (Уэйн университеті, АҚШ). «On Small Models in Jonsson Theory» тақырыбында А.Р. Ешкеев (Қарағанды мемлекеттік университеті), «Жай сандарды бағалау туралы» тақырыбында М.Б. Ғаббасов («FACTOR» жүйелік зерттеу компаниясы, Астана, Қазақстан). Сондай-ақ Қазақстан ғалымдарының бірқатар баяндамалары, мысалы, И.В. Латкин, М. Бекенов, Н. Мархабатов және т.б. талқыланды.

Инновациялық технологиялар негізінде Ұлыбритания тұрғындарын әлеуметтік, медициналық қолдаудың өзекті мәселелерінің шешімі Линкольн университетінің Компьютерлік ғылымдар мектебінің профессоры Салах Аль-Маджидтің «Health and Social Care Assistive Technologies Innovation» баяндамасында берілді.

Ақпараттық қауіпсіздік мәселелері бойынша профессор В.А. Десницкий (РФ ҒА Санкт-Петербург федералды зерттеу орталығы, РФ) «Анализ атак истощения энергоресурсов в системах интернета вещей» баяндамасын талқыға салды.

Новосібір және Еуразия ұлттық университеті ғалымдарының математикалық логикаға және ақпараттық технологияларға негізделген ақпараттық жүйелер туралы және РҒА Сібір бөлімшесінің Информатика мәселелері институты ғалымдарының компьютерлік лингвистика, Қазақстан Республикасы Білім және ғылым министрлігінің Есептеу және ақпараттық технологиялар институтынан М.А. Самбетбаева «Тезаурусы и онтологии в научно-образовательных информационных системах» және С.К. Сагнаева (Л.Н. Гумилев атындағы ЕҰУ) «Модели и методы семантического анализа и представления смысла текста в компьютерной лингвистике» баяндамаларын ғалымдар талқысына ұсынды.

Есептеу құрылғылары мен әртүрлі ақпараттық жүйелер мен технологияларды бағдарламалық қамтамасыз ету арқылы энергияны тұтыну саласындағы ақпараттық технологиялардың заманауи мәселелерін компьютерлік мектебінің профессоры А.Л. Ластовецкий (Дублин, Ирландия) «Energy of computing: Optimization of software applications for energy and performance» баяндамасында ұсынды. Бұл баяндамада әртүрлі елдердің энергетикасы мен экономикасына қатысты мәселелер көрсетілді және Қазақстан Республикасының аумағында атом электр станцияларын салудың ерекше өзекті мәселесі көтерілді. Конференцияға әртүрлі елдерден ондаған ғалымдар қатысып, жүзден астам мақалалар мен тезистер жіберілді. Конференция жұмысы жоғары деңгейде ұйымдастырылып, әр елден келген белгілі ғалымдардың еңбектері таныстырылды, сондай-ақ өзекті мәселелерді шешуге бағытталған отандық жас ғалымдардың еңбектері белгіленген ғылыми зерттеулерге қойылатын заманауи талаптарға сай жүргізілуде.

«Математикалық логика және компьютерлік ғылымдар» халықаралық конференциясы Қазақстан Республикасы Білім және ғылым министрлігінің АР08855497 «Алгебралық жүйелердің модельдік-теориялық және алгоритмдік қасиеттері» грантының қаржылық қолдауымен өтті.

***Ұйымдастыру комитетінің төрағасы, физика-математика ғылымдарының докторы, профессор Тусупов Д.А.***

**Международная научная конференция по математической логике  
и компьютерным наукам  
Евразийский национальный университет,  
Астана РК,  
7-8 октября 2022 года**

Данная конференция проводилась в нашем университете для ознакомления и обсуждения современного состояния наук, взаимного влияния математической логики и компьютерных наук и современных трендов и проблем.

Проблемы и результаты, которые обсуждались на этом форуме, являются продолжением работ шести конференций «Казахско-французский коллоквиум по теории моделей» и «Франко-казахский коллоквиум по теории моделей» в разные годы, а также работ «Шестнадцатой азиатской логической конференции» 2019 году.

Азиатская конференция по логике (АКЛ) являлся крупным международным событием в математической логике. В нем были представлены последние научные разработки в области математической логики и ее приложений, логики в компьютерных науках и философской логики.

Данная конференция по математической логике и компьютерным наукам отражает огромное влияние достижений математической логики на разделы компьютерных наук такие как искусственный интеллект, экспертные системы и другие, направлена на продвижение математической логики и компьютерных наук в евразийском пространстве, в частности в Казахстане, и объединение ученых для обмена информацией и идеями.

В ЕНУ имени Л.Н. Гумилева, которому недавно присвоено звание исследовательского университета, исследования по математической логике и алгебре проводятся на кафедре «Алгебры и геометрии» механико-математического факультета, где работают ученики академика НАН РК Умирбаева Уалбая, академика АН Каз.ССР Тайманова Асана Дабсовича, а также другие представители Сибирской школы алгебры и логики.

Факультет информационных технологий ЕНУ им. Л. Н. Гумилева осуществляет подготовку компетентных ученых по специальностям, заявленных на конференции, однако необходимо усиление в подготовке специалистов по современным направлениям информационных технологий таких как DATA Science, искусственный интеллект, компьютерная лингвистика, информационная безопасность, информационные системы и программная инженерия. На конференции был представлен пул из трех докладов по семантическому и логическому программированию. Конференция была открыта с теоретических основ этой науки, которые опираются на достижения математической логики – совместный доклад «Логико-математическое описание полиномиальной вычислимости» академика РАН С.С. Гончарова и д.ф.-м.н. Свириденко (Институт математики СО РАН) и была продолжена докладами о реализации этих идей в виде информационных продуктов - доклад «Онтологический интеллектуальный анализ данных» д.ф.-м.н. Витяева Е.Е. (Институт математики СО РАН) и доклад "Онтологии и микросервисное объектно-ориентированное программирование" д.ф.-м.н. Манциводы А.В. (Иркутский государственный университет)). Данные доклады вызвали большой интерес и впечатлили участников конференции своей фундаментальностью, широкой областью применения и внедрением в реальные области жизнедеятельности такие, как компьютерные науки, экономика, медицина и т.д.

По направлению математическая логика, алгебра и теория чисел были представлены обзор Бадаева С.А. «Полурешетки Роджерса» (Казахско-Британский технический университет, Алматы), решение актуальной проблемы теории колец академика НАН РК У. Умирбаева (Университет Уэйна, США). Доклад «On Small Models

in Jonsson Theory» был сделан Ешкеевым А.Р. (Карагандинский государственный университет). Габбасов М.Б. (Компания системных исследований «ФАКТОР», Астана, Казахстан) представил доклад «Об одной оценке простых чисел. А также был предложен ряд докладов ученых из Казахстана таких как И.В. Латкин, М. Бекенов, Н. Мархабатов и другие.

Решение актуальных проблем социальной, медицинской поддержки жителей Великобритании на основе инновационных технологий было дано в докладе «Health and Social Care Assistive Technologies Innovation» профессора Salah Al-Majeed школы компьютерных наук университета Линкольн Великобритании.

По проблемам информационной безопасности был предложен доклад «Анализ атак истощения энергоресурсов в системах интернета вещей» профессора Десницкого В.А. (Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, РФ).

Обзоры цикла работ по информационным системам на основе математической логики и информационных технологий ученых Новосибирска и Евразийского национального университета и компьютерной лингвистике ученых института проблем информатики СО РАН, ЕНУ имени Гумилева и Института вычислительных и информационных технологий МОН РК были предложены Самбетбаевой М.А. в докладе «Тезаурусы и онтологии в научно-образовательных информационных системах» и Сагнаевой С.К. в докладе «Модели и методы семантического анализа и представления смысла текста в компьютерной лингвистике».

Современные проблемы информационных технологий в области потребления энергии вычислительными устройствами и программным обеспечением различных информационных систем и технологий были представлены в докладе «Energy of computing: Optimization of software applications for energy and performance» профессора Ластовецкого А.Л. из школы компьютерных наук и информатики УКД, Дублин, Ирландия. Данный доклад отразил проблемы, касающиеся энергетики и экономики различных стран, и ставит особенно актуальный вопрос строительства атомных станций на территории РК.

В работе конференции приняли участие десятки ученых из различных стран и были присланы более ста статей и тезисов.

В целом, работа конференции была организована на высоком уровне, были представлены работы известных ученых из разных стран, а также работы молодых отечественных ученых, которые направлены на решения актуальных проблем, выполняются согласно современным требованиям к научным исследованиям.

Международная конференция «Математическая логика и компьютерные науки» была финансово поддержана грантом МОН РК AP08855497 «Теоретико-модельные и алгоритмические свойства алгебраических систем».

***Председатель организационного комитета, доктор физико-математических наук, профессор Тусупов Д.А.***

**International Research Conference on Mathematical Logic  
and Computer Science,  
Eurasian National University,  
Astana, Kazakhstan,  
October 7-8, 2022**

The conference was held at our university to introduce and discuss the current state of science, the mutual influence of mathematical logic and computer science, modern trends and problems.

The problems and results discussed at this forum are a continuation of the series of six conferences as “Kazakh-French Colloquium on Model Theory” and “French-Kazakh Colloquium on Model Theory” held in different years, as well as “16th Asian Logic Conference” series in 2019.

The Asian Logic Conference was a major international event in mathematical logic. It presented the latest research developments in the field of mathematical logic and its applications, logic in computer science and philosophical logic.

The conference on mathematical logic and computer science reflects the huge impact of the achievements of mathematical logic on sections of computer science such as artificial intelligence, expert systems and others. This event is aimed to promote mathematical logic and computer science in the Eurasian space, in particular in Kazakhstan, and to bring researchers together for the information and ideas exchange.

At L.N. Gumilyov Eurasian National University, that recently gained the title of a research university, research in mathematical logic and algebra is carried out at the department of “Algebra and Geometry” of the Faculty of Mechanics and Mathematics by the followers of Ualbay Umirbaev and Asan Taimanov, the academicians of the National Academy of Sciences of RK and the Academy of Sciences of the Kazakh SSR, as well as other representatives of the Siberian School of Algebra and Logic.

The Faculty of Information Technology of the ENU trains competent researchers in majors declared at the conference, but it is necessary to improve training experts in modern areas of information technology such as DATA Science, artificial intelligence, computer linguistics, information security, information systems and software engineering.

A pool of three reports on semantic and logical programming was presented at the conference. The conference was opened with the theoretical foundations of this science, which are based on the achievements of mathematical logic: a joint report on “Logical and mathematical description of polynomial computability” by Academician S.S. Goncharov and Dr. Sviridenko (Institute of Mathematics SB RAS); reports on implementing the ideas in the form of information products on “Ontological Data Mining” by Dr. E.E. Vityaev (Institute of Mathematics SB RAS); and the report on “Ontologies and Microservice Object-Oriented Programming” by Dr. A.V. Mantsivody (Irkutsk State University). These reports aroused great interest and impressed the conference participants with their fundamental nature, wide scope and implementation in real life areas such as computer science, economics, medicine, etc.

On mathematical logic, algebra and number theory the following reports were presented: review by S.A. Badaev on “Rogers Semilattices” (Kazakh-British Technical University, Almaty); the solution of the topical issue of the theory of rings by Academician of the National Academy of Sciences of the Republic of Kazakhstan U. Umirbaev (Wayne University, USA); the report “On Small Models in Jonsson Theory” by A.R. Eshkeev (Karaganda State University); the report “On one estimate of prime numbers” by M.B. Gabbasov (FACTOR System Research Company, Astana, Kazakhstan). Additionally, a number of reports of researchers from Kazakhstan such as I.V. Latkin, M. Bekenov, N. Markhabatov and others were discussed.

The solution to the urgent problems of social, medical support for residents of the UK based on innovative technologies was presented in the report “Health and Social Care Assistive

Technologies Innovation” by Professor Salah Al-Majeed of the School of Computer Science at Lincoln University in the UK.

On the problems of information security, the report “Analysis of energy depletion attacks in the Internet of Things” by Professor V.A. Desnitsky was discussed (St. Petersburg Federal Research Center of the Russian Academy of Sciences, RF).

Reviews of a work cycle on information systems based on mathematical logic and information technologies by scientists from Novosibirsk and the Eurasian National University and computational linguistics of researchers from the Institute of Computer Science Problems of the Siberian Branch of the RAS, L.N. Humilyov ENU and the Institute of Computing and Information Technologies of the Ministry of Education and Science of the Republic of Kazakhstan were presented by M.A. Sambetbayeva in the report “Thesauri and ontologies in scientific and educational information systems” and S.K. Sagnaeva in the report “Models and methods of semantic analysis and representation of the meaning of text in computational linguistics”.

Modern problems of information technologies in the field of energy consumption by computing devices and software of various information systems and technologies were presented by Professor A.L. Lastovetsky from the School of Computer Science and Informatics of UKD, Dublin, Ireland in the report “Energy of computing: Optimization of software applications for energy and performance”. This report reflected the problems related to the energy and economy of various countries, and raises a particularly topical issue of building nuclear power plants on the territory of the Republic of Kazakhstan.

Dozens of researchers from various countries took part in the conference and more than a hundred articles and abstracts were received.

The conference was organized at a high level and in order to solve urgent problems, the works of famous researchers from different countries, as well as the works of young domestic research fellows were presented and carried out in accordance with the modern requirements for scientific research.

The international conference “Mathematical Logic and Computer Science” was financially supported within the grant of the Ministry of Education and Science of the Republic of Kazakhstan AR08855497 “Model-theoretic and algorithmic properties of algebraic systems”.

***Chair of the Organizing Committee, doctor of Physical and Mathematical Sciences,  
Professor J.A. Tussupov***

УДК 51  
ББК 22.1  
М 29

*Бағдарламалық комитет - Программный комитет - Program committee of the conference*

**Председатель:**

**Сыдыков Е.Б.** – академик НАН РК, Председатель Правления – Ректор ЕНУ им. Л.Н. Гумилева, Астана, Казахстан;

**Сопредседатели:**

**Гончаров С.С.** – академик РАН, Институт математики им. С.Л. Соболева, Новосибирск, Россия;

**Умирбаев У.У.** – академик НАН РК, Университет Уэйна, Детройт, США.

**Члены комитета:**

**Курмангалиева Ж.Д.** – Член Правления – Проректор по науке, коммерциализации и интернационализации, ЕНУ им. Л.Н. Гумилева, Астана, Казахстан;

**Адамов А.А.** – проф., ЕНУ им. Л.Н. Гумилева, Астана, Казахстан;

**Амиргалиев Е.Н.** – доктор технических наук, профессор, заведующий лабораторией искусственного интеллекта и робототехники Института информационных и вычислительных технологий КН МОН РК (г.Алматы);

**Бадаев С.А.** – проф., Казахстанско-Британский технический университет, Алматы, Республика Казахстан;

**Байжанов Б.С.** – чл.-корр., НАН РК, Институт математики и математического моделирования, Алматы, Республика Казахстан;

**Баракнин В.Б.** – доктор технических наук, профессор, ведущий научный сотрудник в Федеральном исследовательском центре информационных и вычислительных технологий (ФИЦ ИВТ). Новосибирск, Россия;

**Бектемисов М.А.** – Заместитель председателя Правления-Проректор по академическим вопросам КазНПУ, доктор физико-математических наук, профессор.

**Вербовский В.В.** – проф., Казахский национальный исследовательский технический университет имени К. И. Сатпаева, Алматы, Республика Казахстан;

**Джумадильдаев А.С.** – академик НАН РК, проф., Казахстанско-Британский технический университет, Алматы, Республика Казахстан;

**Ешкеев А.Р.** – проф., КарГУ им. Е. А. Букетова, Караганда, Казахстан;

**Загорюлько Ю.А.** – зав. лаб. искусственного интеллекта Институт систем информатики им. А.П. Ершова СО РАН, Новосибирск, Россия;

**Искаков К.Т.** – проф., ЕНУ им. Л.Н. Гумилева, Астана, Казахстан;

**Калимолдаев М.Н.** – академик НАН РК, РГП «Институт информационных и вычислительных технологий» Комитета науки МОН РК;

**Карчевский А.Л.** – проф., НГУ, Институт математики им. С.Л. Соболева, Новосибирск, Россия;

**Касымов Н.Х.** – проф. Национального Университета Узбекистана, Ташкент, Узбекистан;

**Козыбаев Д.Х.** – декан ММФ ЕНУ им. Л.Н. Гумилева, Астана, Казахстан;

**Кулпешов Б.Ш.** – чл.-корр. НАН РК, проф., Казахстанско-Британский технический университет, Алматы, Республика Казахстан;

**Ластовецкий А.** – проф., директор лаборатории гетерогенных вычислений, Школа компьютерных наук и информатики УКД, Дублин, Ирландия;

**Лилиан Авено** – XLIM, Научно-исследовательский институт, Университет Пуатье, Франция;

**Морозов А.С.** – заведующий лабораторией логических систем Института математики им. С.Л. Соболева СО РАН, д.ф.-м.н., Новосибирск, Россия;

**Мустафа М.** – проф. Назарбаев Университет, Астана, Казахстан;



**Нуракунов А.М.** – проф., Кыргызский национальный университет им. Дж. Баласагына и Институт математики Национальной академии наук, Бишкек, Кыргызская Республика;  
**Нурсултанов Е.** – проф. МГУ им. М.В. Ломоносова, Астана, Казахстан;  
**Пуза Б.** – проф. Институт Камиллы Джордан, Лионский университет Клода Бернара 1, Лион, Франция;  
**Рахметуллина С.Ж.** – Председатель Правления – Ректор НАО ВКТУ им. Д.Серикбаева, Усть-Каменогорск, Казахстан;  
**Сеилов Ш.Ж.** – академик Международной Академии Связи, доктор экон.наук, проф., декан ФИТ ЕНУ им. Л.Гумилева, Астана, Казахстан;  
**Судоплатов С.В.** – зам. директора Института математики им. С.Л. Соболева СО РАН, д.ф.-м.н., Новосибирск, Россия;  
**Старовойтов В.В.** – проф. Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Белоруссия,  
**Темирбеков Н.М.** – вице-президент и академик НИА РК, чл.-корр. НАН РК, Алматы, Казахстан;  
**Хисамиев Н.Г.** – проф. ЕНУ им. Л.Н. Гумилева, Астана, Казахстан.

***Ұйымдастырушы комитет - Организационный комитет - Organizing committee of the conference***

**Тусупов Д.А.** – проф. ЕНУ им. Л.Н. Гумилева, Астана, Казахстан;  
**Муханова А.А.** – доктор философии (PhD), ЕНУ им. Л.Н. Гумилева, Астана, Казахстан;  
**Абдикеримова Г.Б.** – доктор философии (PhD), ЕНУ им. Л.Н. Гумилева, Астана, Казахстан;  
**Серикбаева С.К.** – доктор философии (PhD), ЕНУ им. Л.Н. Гумилева, Астана, Казахстан;  
**Махажанова У.Т.** – доктор философии (PhD), ЕНУ им. Л.Н. Гумилева, Астана, Казахстан;  
**Сисенов Н.А.** – ст.препод. ЕНУ им. Л.Н. Гумилева, Астана, Казахстан;  
**Мусабеков А.К.** – ст.препод. ЕНУ им. Л.Н. Гумилева, Астана, Казахстан;  
**Есенова М. Б.** – докторант ЕНУ им. Л.Н. Гумилева, Астана, Казахстан.

*Конференция және халықаралық ғылыми конференцияның еңбектерін басып шығару Қазақстан Республикасы Ғылым және жоғары білім министрлігінің №AP08855497 «Алгебралық құрылымдардың теоретикалық,-модельдік және алгоритмикалық қасиеттері» гранты есебінен қаржыландырылды.*

*Конференция и издание трудов международной научной конференции профинансированы грантом № AP08855497 «Теоретико-модельные и алгоритмические свойства алгебраических структур» Министерство науки и высшего образования Республики Казахстан.*

*The conference and the publication of the proceedings of the international scientific conference were funded by grant No. AP08855497 “Theoretic-model and algorithmic properties of algebraic structures” of the Ministry of Science and Higher Education of the Republic of Kazakhstan.*

**«Математикалық логика және компьютерлік ғылымдар»** атты Халықаралық ғылыми конференция еңбектері – Астана, 2022. – 341 бб.

**ISBN 978-601-337-776-6**

В сборнике опубликованы материалы конференции по математической логике и компьютерным наукам, современным трендам и проблемам таким, как семантическое и логическое программирование, математическая логика, алгебра и теория чисел, проблемы информационной безопасности, компьютерная лингвистика, современные проблемы информационных технологий в области потребления энергии вычислительными устройствами и программным обеспечением различных информационных систем и технологий.

Данная конференция по математической логике и компьютерным наукам отражает огромное влияние достижений математической логики на разделы компьютерных наук такие как искусственный интеллект, экспертные системы и другие, направлена на продвижение математической логики и компьютерных наук в евразийском пространстве, в частности в Казахстане, и объединение ученых для обмена информацией и идеями.

**УДК 51  
ББК 22.1  
М 29**

**ISBN 978-601-337-776-6**

**© НАО ЕНУ имени Л.Н. Гумилева, 2022**

## МАЗМУНЫ - СОДЕРЖАНИЕ - CONTENT

<b>1. СЕКЦИЯ. АЛГЕБРА И МАТЕМАТИЧЕСКАЯ ЛОГИКА</b>	
Судоплатов С.В.. Об алгебраических замыканиях в моделях элементарных теорий.....	14
Yeshkeyev A.R., Tungushbayeva I.O., Omarova M.T. Forcing companions of the hereditary jonsson ap-theory.....	17
Gabbasov M.B., Kuanov T.Z. On one estimate of prime numbers.....	21
Кемелбек М.Н., Алимжанова Л.М. IT-жобалар мен стартаптардың инвестициялық тартымдылығын бағалау ерекшеліктерін зерттеу.....	24
Yeshkeyev A.R., Ulbrikht O.I., Mussina N.M. Similarities of hybrids of jonsson quasivarieties.....	27
Yeshkeyev A.R., Ulbrikht O.I., Issayeva A.K. On small models in jonsson theory.....	29
Вербовский В В., Ершигешова А.Д. Упорядоченная стабильность и обогащения чистого линейного порядка одноместной функцией.....	32
Латкин И.В., Селиверстов А.В. Вычисления над упорядоченными кольцами посредством обобщённых регистровых машин.....	36
Kozybaev D. Example of a non-locally finite novikov coalgebra.....	41
Бекенов М., Касатова А., Нуракунов А. Формульно-определимые классы алгебраических систем.....	43
Markhabatov N.D. Model theory of finite and pseudofinite graphs.....	46
Хисамиев Н.Г., Тусупов Д.А., Тыныбекова С.Д. Критерий вычислимости $\langle p_\omega, N \rangle$ –разложимой абелевой группы без кручения.....	47
Basheyeva A.O. About generated finite pointed abelian group.....	50
Lutsak S.M., Voronina O.A. On some properties of the quasivariety lattice of lukasiewicz algebras.....	51
<b>2. СЕКЦИЯ. КОМПЬЮТЕРНЫЕ НАУКИ, ПРЕДСТАВЛЕНИЕ ДАННЫХ И ПРОГРАММНАЯ ИНЖЕНЕРИЯ</b>	
Idrissova I.A., Mukhanova A.A. Methodological foundations of integrated information security of the organization's information and communication infrastructure.....	55
Тохметов А.Т., Абдикеримова Г.Б., Ли В.Э., Токашев Н.К. Технология блокчейн и безопасность конечных устройств интернета вещей.....	60
Омирбекова И.С., Баймулдина Н.С. Повышение эффективности работы информационной системы мониторинга инцидентов безопасности в банковской деятельности.....	64
Муратхан Р., Туткуше А.Е. Ақпараттық қауіпсіздіктің қатерлер инцидентін модельдеу .....	68
Қинойт Д.Д. Хеш-функция ұғымы және бағдарламалық көрінісі.....	71
Самбетбаева М.А. Баймырзаева Г.Ж. Тезаурусы и онтологии в научно-образовательных информационных системах.....	75
Қуатова Ж.Қ., Ерманов Ә.М., Азанов Н.П. Анализ технологических, экономических факторов, которые повлияют на будущее SIEM.....	82
Азаматова Д.Т., Мусиралиева Ш.Ж., Голлманн Д. Фэйк аккаунт анықтау мәселесінде машиналық оқытуды қолдану.....	86
Кубигенова А.Т., Актаева А.У., Исмаилова А.А., Есмагамбетова Г.К. Проблемы обеспечения информационной безопасности хранения и обработки больших данных.....	90
Kystaubayeva A.S., Nebessayeva Z O., Teaching 3d modeling in art schools.....	95

<b>3.СЕКЦИЯ. КОМПЬЮТЕРНЫЕ НАУКИ, МКМ И АНАЛИЗ ДАННЫХ</b>	
Ластовецкий А. Энергия вычислений: оптимизация энергии и производительности программ.....	98
Ахметбаев Д.С., Джандигулов А.Р., Быстрова С.В., Бежежанова Ш.С. Совершенствование алгоритмов топологической модели установившегося режима электроэнергетических систем.....	103
Өмірзақова Ф.Н. 3DS MAX бағдарламасының көмегімен үш өлшемді сахналарды модельдеу.....	109
Алимжанова Л.М., Нуриева Д.Я. Применение gfm анализа данных для сегментирования клиентов и удержания их лояльности.....	113
Қапбар Г.С., Карымсакова А.Е., Гапбарова П.С. ЖОО студенттеріне python тілінде бағдарламалауды оқыту.....	118
Рзаев Б.Т., Бельдеубаева Ж.Т., Увалиева И.М. Детектирование аномальных ситуаций информационных систем.....	123
Мауина Г.М., Айтимова У.Ж. Шешім қабылдаудың көп критериялы есебінің математикалық моделі.....	128
Қабдешов Т., Хажайв К. Адам денесіндегі температураны цилиндрлік бөлудің математикалық моделі.....	133
Орман И.М. Система функций типа хаара.....	139
Шевцов В.А., Каиржанова А.Д., Исмаилова А.А. Разработка скрипта для выявления переменных локусов бактериальных штаммов <i>francisella tularensis</i> .....	150
Муратова Ж.М., Жетписов К., Иксебаева Ж.С. Көпөлшемді модельдерді классификациялау, сипаттау және салыстыру ортасы.....	155
Голенко Е. С., Исмаилова А. А. Модель машинного обучения для предсказания функций белков.....	159
Дрозд В.Г., Целых У.Е., Сорокин А.Д., Ярославцев С.В. Компьютерное моделирование ведения складского учета в строительном бизнесе.....	164
Алдабергенова К.М., Кантуреева М.А. Ауыл шаруашылық бағытын автоматтандыру бойынша ақпараттық жүйелерді құру мәселелері.....	169
Десницкий В.А. Анализ атак истощения энергоресурсов в системах интернета вещей.....	173
Серік М., Жумагулова С.К., Копыльцов А.В. Вопросы применения параллельных вычислений в базе данных обучающимися вузов.....	177
Сағатова А.С., Алимжанова Л.М., Сарбасова А.К. Компаниядағы тиімді ішкі коммуникация үшін it инфрақұрылымын оңтайлы таңдау.....	179
Некесова А.А. Деректерді зияткерлік талдау(data mining).....	183
Темиргазиева Ш., Базаркулова И.Е., Омаров Б.С. Кескіндерді тануда қолданылатын терең оқытудың әдістеріне шолу.....	186
<b>4 СЕКЦИЯ ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ</b>	
Гончаров С., Нечесов А. Объектно-ориентированный логический язык программирования для искусственного интеллекта и робототехники.....	191
Сериков М.К., Алимжанова Л.М., Сербин В.В. Маркетинговые возможности продвижения бренда с помощью интернет-маркетинга.....	196
Бөрібаев Б.Б., Кенесова Н.А. Ақпараттық технологиялар терминдерін мемлекеттік тілге аудару және жүзеге асыру.....	201

<b>Байбатшаева А.Е., Иманкулова З.А. Карабаева Ж.Ж.</b> Киім топтамасының эскиздерін жобалаудағы компьютерлік графиканың қолданылуы.....	205
<b>Alikul S.M.</b> Using sentiment analysis to assess the investment attractiveness of an organization.....	209
<b>Ибраева Қ.Қ., Досымова К.Б., Қожабай А.Б.</b> Білім беру жүйесіндегі сандық технологиялар.....	213
<b>Касекеева А.Б., Марат М.М., Кашкимбаева Н.М.</b> Нахождение похожих объектов с помощью mapreduce.....	217
<b>Шонагулов О.А.</b> Структура баз знаний интеллектуальной системы в области ветеринарии.....	223
<b>Танченко Л.А., Одноуров Д.В.</b> Применение ogm-технологии при разработке образовательного онлайн портала с элементами социальной сети.....	228
<b>Байшоланова К.С., Бақытжанұлы А.</b> Жоғары білім беру жүйесіне заттар интернет (iot) технологиясын енгізу мүмкіндіктері.....	233
<b>Yessenova M., Abdikerimova G., Bolsynbek M.</b> Use of methods to determine damage to agricultural crops (akmola region, shortandy district).....	236
<b>Максутова Н.Ш., Хамитова Ж.Ж.</b> Использование метода ближайшего соседа для классификации изображений.....	241
<b>Максутова Н.Ш., Хамитова Ж.Ж.</b> Сверточная нейронная сеть применительно к классификации изображений.....	246
<b>Хасенов А.Ш., Мусабеков А.К.</b> Адамдардың тобырлану динамикасын мультиагентті модельдеу.....	252
<b>Есимова Н.Б., Есендаулетова Ж.Т.</b> Разговорный интерфейс в среде bot framework.....	256
<b>Серикбаева С.К., Болсынбек М.Қ., Мусабеков А.К.</b> Ғылыми-білім беру қызметін қолдайтын әкімшілік топтамаларға қолжетімділігінің басқару жүйесі.....	262
<b>Амирбай А.А., Муханова А.А.</b> SIEM – жүйесі негізінде компьютерлік жүйенің ақпараттық ресурстарын тиімді қорғау әдістері.....	267
<b>Кадиркулов К., Исмаилова А.</b> Автоматизация интерпретации результатов исследований медицинских лаборатории.....	272
<b>Shukurova A., Kumalakov B., Zhukabayeva T.</b> An intelligent agent of finite state machine in example of games.....	276
<b>Сабырбаева К.М., Жүзбаев С.С.</b> Беттердің бейне кескіндерін талдау негізінде қатысушының онлайн іс-шарадағы зейінін анықтау.....	282
<b>Садирмекова Ж.Б., Самбетбаева М.А., Боранкулова Г.С.</b> Разработка модели портала научных знаний.....	287
<b>Сарсенбаева А.К., Төлепберген А.А., Сүйіндік Н.М., Алимова Ж.С.</b> Байес логикалық қорыту жүйесіне негізделген сараптамалық жүйені жасаудың мысалы.....	292
<b>Orazayeva A., Tussupov J., Pavlov S., Babyuk N.</b> Methods and system for analysis of biomedical images in dynamic changes.....	296
<b>Қасенхан А.М., Ускенбаева Р.К., Нуралықызы С., Джунусова С.М.</b> Үлгіні тануда қолданылатын әдістерді қарастыру.....	301
<b>Шормакова А.Н., Тукеев У.А.</b> Ағылшын-қазақ машиналық аударма пост-редакциялау технологиясы.....	305
<b>Омарова Г.С., Тұрсынбай Е.</b> Программная реализация комбинации адаптивных методов повышения контраста для улучшения рентгеновских изображений.....	308
<b>Мурзахметов А.Н., Бапанов А.А., Бейсов Н.Қ., Тасжуреков Ж.Қ.</b> Naming game моделі негізінде пікір алмасу.....	312
<b>Salah Al-Majeed.</b> Health and social care innovation assistive technologies.....	316

<b>Зүнімова Г.Д.</b> ЖОО-ның басқару шешімдерін қабылдау үшін білім сапасын қабылдау.....	323
<b>Abdikerimova G., Tussupov J.</b> Methods for detecting and highlighting homogeneous areas on texture images.....	327
<b>Кайбасова Д.Ж.</b> Білім беру құжаттарының ұқсастығын анықтауда кластерлік талдаудың қолданылуы.....	333
<b>Shekerbek A., Abdikerimova G.</b> Methods for detecting pathology of the chest on radiographic images.....	336

ОБ АЛГЕБРАИЧЕСКИХ ЗАМКНАНИЯХ В МОДЕЛЯХ ЭЛЕМЕНТАРНЫХ ТЕОРИЙ

С.В. Судоплатов

Институт математики им. С.Л. Соболева СО РАН,  
Новосибирский государственный технический университет,  
Новосибирский государственный университет  
e-mail: sudoplat@math.nsc.ru

В книге С. Шелаха [1] рассматриваются два вида замыканий в алгебраических системах, алгебраическое и определимое, а также связанные с ними следующие понятия:

**Определение** [1, 2]. Пусть  $M$  – некоторая структура,  $T = \text{Th}(M)$ ,  $\bar{a}$  и  $\bar{b}$  – кортежи элементов из  $M$ .

1. Говорим, что кортеж  $\bar{b}$  *определяется* или *определен* формулой  $\varphi(x, \bar{a})$ , если  $\varphi(x, \bar{a})$  имеет единственное решение  $\bar{b}$  в  $M$ . Говорим, что кортеж  $\bar{b}$  *определяется* или *определен* типом  $p$ , если  $\bar{b}$  – единственный кортеж, реализуемый типом  $p$ . Кортеж  $\bar{b}$  называется *определимым* над множеством  $A$ , если тип  $\text{tp}(\bar{b}/A)$  определяет этот кортеж.

2. Для множества  $A$  теории  $T$  объединение множеств решений формул  $\varphi(x, \bar{a})$ ,  $\bar{a} \in A$ , таких что  $\models \exists^{\leq n} x \varphi(x, \bar{a})$  для некоторого  $n \in \omega$  (соответственно  $\models \exists^{\leq 1} x \varphi(x, \bar{a})$ ) называется *алгебраическим (определимым) замыканием* множества  $A$ . Алгебраическое замыкание множества  $A$  обозначается через  $\text{acl}(A)$ , а его определимое замыкание – через  $\text{dcl}(A)$ .

В этом случае мы говорим о том, что формулы  $\varphi(x, \bar{a})$  *свидетельствуют* об этом алгебраическом/определимом замыкании, и эти формулы называются *алгебраическими/определяющими*.

Любой элемент  $b \in \text{acl}(A)$  (соответственно  $b \in \text{dcl}(A)$ ) называется *алгебраическим (определимым)* над  $A$ . Если множество  $A$  зафиксировано или пусто, то  $b$  называется просто *алгебраическим* или *определимым*.

3. Если  $\text{dcl}(A) = \text{acl}(A)$ , то через  $\text{cl}(A)$  обозначается их общее значение.

4. Если  $A = \text{acl}(A)$  (соответственно  $A = \text{dcl}(A)$ ), то  $A$  называется *алгебраически (определимо) замкнутым*.

5. Тип  $p$  называется *алгебраическим (определяющим)*, если  $p$  реализуется лишь конечным числом кортежей (единственным кортежем), т.е.  $p$  содержит алгебраическую (определяющую) формулу  $\varphi$ . Эта формула  $\varphi$  может быть выбрана с минимальным числом решений, и в этом случае формула  $\varphi$  изолирует тип  $p$ . Число этих решений называется *степенью*  $\text{deg}(p)$  типа  $p$ .

6. Полные алгебраические типы  $p(x) \in S(A)$  это в точности те типы, которые имеют вид  $\text{tp}(a/A)$ , где  $a$  – алгебраический элемент над  $A$ . *Степенью* элемента  $a$  над  $A$ , обозначается  $\text{deg}(a/A)$ , называется степень типа  $\text{tp}(a/A)$ .

В настоящей работе мы рассматриваем модификации алгебраического замыкания, относящиеся к множествам решений формул, ограниченным по мощности некоторым натуральным числом.

**Определение.** 1. Для  $n \in \omega \setminus \{0\}$  и множества  $A$  элемент  $b$  называется  *$n$ -алгебраическим* над  $A$ , если  $a \in \text{acl}(A)$  и это свидетельствуется формулой  $\varphi(x, \bar{a})$ , для  $\bar{a} \in A$ , имеющей не более  $n$  решений.

2. Множество всех  $n$ -алгебраических элементов над  $A$  обозначается через  $\text{acl}_n(A)$ .

3. Если  $A = \text{acl}_n(A)$ , то множество  $A$  называется  *$n$ -алгебраически замкнутым*.

4. Тип  $p$  называется  *$n$ -алгебраическим*, если  $p$  имеет в любой модели не более чем  $n$  реализаций, т.е.  $\text{deg}(p) \leq n$ .

5. Полные  $n$ -алгебраические типы  $p(x) \in S(A)$  это в точности те типы, которые имеют вид  $\text{tp}(a/A)$ , где  $a$  –  $n$ -алгебраический элемент над  $A$ , т.е. элемент с условием  $\text{deg}(a/A) \leq n$ . Здесь  $\text{deg}(a/A) = k \leq n$  определяет  $n$ -степень типа  $\text{tp}(a/A)$  и элемента  $a$  над  $A$ .

6. Если  $\text{acl}(A) = \text{acl}_n(A)$ , то минимальное такое значение  $n$  называется *степенью алгебраизации* над множеством  $A$  и обозначается через  $\text{deg}_{\text{acl}}(A)$ . Если же такое значение  $n$  не существует, то полагаем  $\text{deg}_{\text{acl}}(A) = \infty$ . Супремум значений  $\text{deg}_{\text{acl}}(A)$  по всем множествам  $A$  данной теории  $T$  обозначается через  $\text{deg}_{\text{acl}}(T)$  и называется *степенью алгебраизации* теории  $T$ .

Напомним следующее алгебраическое понятие [3], которое позволяет связать множества реализаций типов с группой автоморфизмов данной насыщенной структуры.

**Определение.** Для множества  $A$  и элемента  $a$   $A$ -орбитой  $\text{Orb}_A(a)$  элемента  $a$  называется множество всех элементов  $b$  данной структуры, связанных с  $a$  некоторым  $A$ -автоморфизмом.

Следующее предложение дает алгебраическую характеристику для  $n$ -алгебраических типов.

**Предложение 1.** Тип  $p$  является  $n$ -алгебраическим над  $A$  тогда и только тогда, когда любая/некоторая  $(|A|+|T|)$ -насыщенная модель  $M$ , содержащая  $A$ , имеет конечное число  $A$ -орбит  $O$ , состоящих из реализаций типа  $p$ , все эти орбиты конечны, и, более того, объединение  $\bigcup O$  имеет не более  $n$ -элементов. Если  $p$  – полный тип, то такая  $A$ -орбита единственна в  $M$ .

Аналогично лемме 6.2 из [1] доказывается следующее:

**Предложение 2.** 1.  $A \subseteq \text{acl}_m(A) \subseteq \text{acl}_n(A) \subseteq \text{acl}(A)$  для любых  $m < n$ .

2. Если  $A \subseteq B$  и  $n \geq 1$ , то  $\text{acl}_n(A) \subseteq \text{acl}_n(B)$ .

3. Если множество  $A$  определимо (алгебраически) замкнуто, то  $A = \text{dcl}(A)$  ( $A = \text{acl}(A)$ ).

4. Если множество  $A$   $n$ -алгебраически замкнуто для  $n \geq 2$ , то  $A = \text{acl}(A)$  тогда и только тогда, когда любая конечная орбита над  $A$  имеет не более чем  $n$  элементов.

5. Кортеж  $\bar{b}$  определен (является алгебраическим) над  $A$  тогда и только тогда, когда  $\bar{b} \in \text{dcl}(A)$  ( $\bar{b} \in \text{acl}(A)$ ).

**Замечание.** По определению  $\text{acl}_1(A) = \text{dcl}(A)$  для любого множества  $A$ . При условии существования  $\text{cl}(A)$  для любого множества  $A$  теории  $T$ , получаем минимальное значение для степени алгебраизации теории  $T$ :  $\text{deg}_{\text{acl}}(T) = 1$ .

Следующая теорема описывает все возможные значения для степени алгебраизации совместной теории.

**Теорема.**

1. Для любой совместной теории  $T$ ,  $\text{deg}_{\text{acl}}(T) \in ((\omega + 1) \setminus \{0\}) \cup \{\infty\}$ .

2. Для любого  $\lambda \in ((\omega + 1) \setminus \{0\}) \cup \{\infty\}$  существует теория  $T_\lambda$  с условием  $\text{deg}_{\text{acl}}(T_\lambda) = \lambda$ .

**Примеры.** 1. Рассмотрим дерево  $D_n$  [4], у которого каждая вершина имеет фиксированную степень  $n \in \omega$ . Обозначим через  $T_n$  теорию  $\text{Th}(D_n)$ .

Если  $n = 0$ , то дерево  $D_n$  одноэлементно,  $\text{acl}(A) = \text{dcl}(A) = D_n$  для любого  $A \subseteq D_n$ , следовательно,  $\text{deg}_{\text{acl}}(T_0) = 1$ .

Если  $n = 1$ , то дерево  $D_n$  двухэлементно,  $\text{acl}(A) = \text{acl}_2(A) = D_n$  для любого  $A \subseteq D_n$ , следовательно,  $\text{deg}_{\text{acl}}(T_1) = 2$ .

Если  $n \geq 2$ , то дерево  $D_n$  счетно. При этом,  $\text{acl}(\emptyset) = \emptyset$ ,  $\text{acl}(A) = \text{acl}_2(A) = D_2$  для любого непустого  $A \subseteq D_2$ , следовательно,  $\text{deg}_{\text{acl}}(T_2) = 2$ . Если же  $n \geq 3$ , то при сохранении  $\text{acl}(\emptyset) = \emptyset$  и  $\text{acl}(A) = D_n$  для любого непустого  $A \subseteq D_n$ ,  $\text{acl}_m(A)$  конечно для любого конечного  $A$  и оператор  $\text{acl}_m$  не обладает свойством транзитивности. Тем самым, при  $n \geq 3$ ,  $\text{deg}_{\text{acl}}(T_n) = \infty$ .

2. Рассмотрим граф  $\Gamma_n = \langle \{a, b_1, \dots, b_n\}, \{(a, b_1), \dots, (a, b_n)\} \rangle$ ,  $n \in \omega \setminus \{0\}$ . Поскольку имеется  $n$ -элементная орбита над  $\emptyset$  и над  $\{a\}$  и эта орбита имеет максимальную



мощность среди всех орбит над подмножествами носителя  $\Gamma_n$ , то имеет  $\text{deg}_{\text{acl}}(\text{Th}(\Gamma_n)) = n$ . Беря дизъюнктивное объединение графов  $\Gamma_n$  по всем натуральным  $n$ , получаем граф  $\Gamma_\infty$ , для которого  $\text{deg}_{\text{acl}}(\text{Th}(\Gamma_\infty)) = \infty$ .

3. Подходящим обогащением  $M$  графа  $\Gamma$ , который получается бесконечным тиражированием графов  $\Gamma_\infty$ , реализуется значение  $\text{deg}_{\text{acl}}(\text{Th}(M)) = \omega$ .

4. Следуя [5], замечаем, что для любой линейно упорядоченной структуры  $M$  и любого подмножества  $A \subseteq M$ ,  $\text{acl}(A) = \text{dcl}(A)$ , откуда получаем  $\text{deg}_{\text{acl}}(\text{Th}(M)) = 1$ . Если же структура  $M$  циклически упорядочена, то  $\text{acl}(A) = \text{dcl}(A)$  для любого непустого подмножества  $A \subseteq M$ , а  $\text{acl}(\emptyset)$  может иметь произвольную конечную мощность. Тем самым, в зависимости от циклически упорядоченной структуры  $M$  значение  $\text{deg}_{\text{acl}}(\text{Th}(M))$  может быть произвольным ненулевым натуральным числом.

5. Если  $T$  – теория алгебраически замкнутого поля ненулевой характеристики, то корни многочленов могут образовывать как угодно большие орбиты над множествами коэффициентов этих многочленов. Таким образом,  $\text{deg}_{\text{acl}}(T) = \infty$ .

Возникают следующие проблемы:

**Проблема 1.** Описать  $n$ -алгебраические замыкания для моделей естественных теорий.

**Проблема 2.** Описать степени алгебраизации для естественных классов теорий.

*Работа выполнена в рамках государственного задания Института математики им. С.Л. Соболева СО РАН, проект № FWNF-2022-0012.*

#### Литература

1. Shelah S. Classification theory and the number of non-isomorphic models. Amsterdam: North-Holland, 1990. — 705 p.
2. Tent K., Ziegler M. A Course in Model Theory. Cambridge: Cambridge University Press, 2012. — 248 p.
3. Hodges W. Model theory. Cambridge: Cambridge University Press, 1993. — 772 p.
4. Судоплатов С.В., Овчинникова Е.В. Дискретная математика: учебник и практикум для вузов. М.: Юрайт, 2022. — 280 с.
5. Kulpeshov B.Sh., Macpherson H.D. Minimality conditions on circularly ordered structures // Mathematical Logic Quarterly, 51:4 (2005), 377-399.

## FORCING COMPANIONS OF THE HEREDITARY JONSSON AP-THEORY

Yeshkeyev Aibat Rafhatuly<sup>1</sup>, Tungushbayeva Indira Orazbekovna<sup>1</sup>, Omarova Makhabat Toleuovna<sup>1,2</sup>

<sup>1</sup>Karaganda Buketov University, Karaganda

<sup>2</sup>Karaganda University of Kazpotrebsoyuz, Karaganda

E-mail: [aibat.kz@gmail.com](mailto:aibat.kz@gmail.com), [intng@mail.ru](mailto:intng@mail.ru), [omarovamt\\_963@mail.ru](mailto:omarovamt_963@mail.ru)

This research focuses on the forcing companions of the Jonsson AP-theories in the language enriched by new constant and predicate symbols. The model-theoretic results obtained here generalize the cases concerning some well-known differential algebras, particularly, differential and differentially closed fields of characteristic  $p$ .

The concept of AP-theory was introduced in [1] by the first author of the paper, Yeshkeyev A.R., and plays a significant role in the study of both algebraic and model-theoretic properties of algebras with inductive theories in terms of the amalgam and joint embedding properties. It is also known that AP-theories form a special subclass within Jonsson theories. We call such theories Jonsson AP-theories. The examples are the theories of differential fields of zero and positive characteristic, such as  $DF_0$ ,  $DCF_0$ ,  $DPF_p$ ,  $DCF_p$ , while  $DF_p$  does not have the amalgam property. The complexity of the case of positive characteristic in differential fields gave rise to the study of the connection between a Jonsson AP-theory and non-Jonsson theory. In this manner, for the first time in the history of Model Theory, an approach to combining such theories is demonstrated in the paper. This will allow us to further develop the methods of studying Jonsson theories and expand the apparatus for studying incomplete theories.

We start with the main definitions and facts concerning the subject of the study. Recall the definitions of a model companion and a forcing companion.

Definition 1 [2, p. 129]. Let  $T$  be a theory of the language  $L$ . A forcing companion of the theory is a theory  $T^f$  that satisfies the following condition:

$$T^f = \{\phi \mid T \vdash \neg\neg\phi\}.$$

The following results were proved by J. Barwise and A. Robinson:

Theorem 1 [2, p. 133]. Let  $T_1$  and  $T_2$  be the theories of the language  $L$ . Then  $T_1$  and  $T_2$  are mutually model consistent if and only if  $T_1^f = T_2^f$ .

Theorem 2 [2, p. 134]. Let  $T$  be mutually model consistent with some inductive theory  $T'$ . Then  $T' \subseteq T^f$ . Therefore, if  $T$  is an inductive theory then  $T \subseteq T^f$ .

We are working within the framework of the following definition of Jonsson theory published in the Russian edition of [3].

Definition 2 [3, p. 80]. A theory  $T$  is called Jonsson if:

1.  $T$  has at least one infinite model;
2.  $T$  is an inductive theory;
3.  $T$  has the amalgam property (AP) and joint embedding property (JEP).

The properties AP and JEP are originally considered to be algebraic, however, take a great part in model-theoretic research. It is known from [4] that AP and JEP are independent of each other. The examples of algebras that support this fact are some classes of unars. On the other hand, there exists a specific link between these two properties while considering some theories including Jonsson ones.

The following definition was proposed by Yeshkeyev A.R. and refines the case of our study.

Definition 3 [1, p. 130]. A theory  $T$  is called an AP-theory if it has the amalgam property and from this, it follows that  $T$  also has the joint embedding property, i.e.  $AP \rightarrow JEP$ .

The examples are mentioned above.

The following concepts and facts play a crucial role in the construction of a model-theoretic apparatus associated with the study of Jonsson theories.

Definition 4 [5, p. 155]. Let  $T$  be a Jonsson theory. A model  $C_T$  of power  $2^{|T|}$  is said to be a semantic model of the theory  $T$  if  $C_T$  is a  $|T|^+$ -homogeneous  $|T|^+$ -universal model of the theory  $T$ .

Theorem 3 [5, p. 155].  $T$  is Jonsson iff it has a semantic model  $C_T$ .

The following definition was introduced by T.G. Mustafin.

Definition 5 [5, p. 155]. A Jonsson theory  $T$  is called perfect if its semantic model  $C_T$  is saturated.

Definition 6 [5, p. 161]. The elementary theory of a semantic model of a Jonsson theory  $T$  is called the center of this theory. The center is denoted by  $T^*$ , i.e.  $Th(C) = T^*$

Theorem 4 [5, p. 158]. Let  $T$  be a Jonsson theory. Then the following conditions are equivalent:

- 1)  $T$  is perfect;
- 2)  $T^* = Th(C)$  is the model companion of the theory  $T$ .

The following theorem is of particular importance for this study:

Theorem 5 [5, p. 162]. Let  $T$  be a perfect Jonsson theory. Then the following statements are equivalent:

- 1)  $T^*$  is the model companion of  $T$ ;
- 2)  $Mod T^* = E_T$ ;
- 3)  $T^* = T^f$ , where  $T^f$  is a forcing companion of the theory  $T$ .

Theorem 6 [6, p. 1243]. Let  $T$  be a Jonsson theory. Then for any model  $A \in E_T$  theory  $T^0(A)$  is Jonsson, where  $T^0(A) = Th_{\forall\exists}(A)$ .

We can see from Theorem 6 that in case of the perfectness of  $T$  its center  $T^*$  is also a Jonsson theory.

The following definition will help us to specify the class of Jonsson theories which we deal with in this paper.

Definition 7 [7, p. 120]. A Jonsson theory is said to be hereditary if, in any of its permissible enrichment, it preserves the Jonssonness.

We move to the setting of our problem. We consider the theories  $\Delta_1, \Delta_2, \Delta_3$  that satisfy the following conditions:

- 1)  $\Delta_1$  is an inductive theory that is not a Jonsson theory, but has a model companion which is the theory  $\Delta_3$ ,
- 2)  $\Delta_2$  is a hereditary Jonsson AP-theory that has a model companion, which is also  $\Delta_3$ .

Based on these conditions, we can see the following. All the mentioned theories are mutually model consistent, because  $\Delta_3$  is mutually model consistent with both  $\Delta_1$  and  $\Delta_2$ , for which  $\Delta_3$  is the model companion. This means that  $\Delta_1$  and  $\Delta_2$  are mutually model consistent with each other. At the same time, according to Theorem 1, the forcing companions of mutually model consistent theories must coincide, which means that  $\Delta_1^f = \Delta_2^f$ . is a perfect Jonsson theory, while  $\Delta_2^* = Th(C_2) = \Delta_3$ ,  $C$  is a semantic model of  $\Delta_2$ , which follows from Theorem 4. In addition, Theorem 5 gives us reason to assert that  $\Delta_3$  is also a forcing companion of  $\Delta_2$ , i.e.  $\Delta_3 = \Delta_2^f$ . So we get  $\Delta_1^f = \Delta_2^f = \Delta_3$ .

We consider the following extensions of the theories  $\Delta_1, \Delta_2, \Delta_3$  in the enrichment of  $L$  by adding new constant and predicate symbols  $c$  and  $P$ . Let  $\overline{\Delta_1}$  be a theory that extends  $\Delta_1$  by enriching the language  $L$  with the predicate symbol  $P$  as follows:

$$\overline{\Delta}_1 = \Delta_1 \cup \Delta_1^f \cup \{P, \subseteq\},$$

where  $\{P, \subseteq\}$  is an infinite list of  $\exists$ -sentences and interpretation of  $P$  is an existentially closed submodel in a model of  $\Delta_1$ .

Let  $\overline{\Delta}_2$  be a theory that extends  $\Delta_2$  when a new constant symbol  $c$  is added to the language  $L$  and defined as follows:

$$\overline{\Delta}_2 = \Delta_2 \cup \Delta_2^f \cup Th_{\forall\exists}(C_2, c),$$

where  $C_2$  is a semantic model of Jonsson theory  $\Delta_2$ . Since  $\Delta_2$  is a hereditary Jonsson theory,  $\overline{\Delta}_2$  is also a Jonsson theory.

Here we pose two questions:

1) How will the addition of new symbols  $P$  and  $c$  to the language  $L$  and the subsequent expansion of  $\Delta_1$  and  $\Delta_2$  affect the forcing companion of the received theories?

2) When combining the theories  $\overline{\Delta}_1$  and  $\overline{\Delta}_2$ , can we obtain a consistent theory and what will be its forcing companion?

The answers are the following theorems.

Theorem 7.  $\overline{\Delta}_1^f = \Delta_1^f$ .

Thus, we can conclude that the forcing companion of the inductive theory  $\Delta_1$  does not change when enriching the language of this theory with the predicate symbol  $P$ .

Theorem 8.  $\overline{\Delta}_2^f = \Delta_2^f$ .

This means that the addition of the constant  $c$  to language  $L$  does not affect the forcing companion when expanding theory  $\Delta_2$  to  $\overline{\Delta}_2$ .

The following Theorem, as known as Robinson's Consistency Theorem, plays an important role for proving one of the results of this work, namely, Theorem 10.

Theorem 9 [3, p. 77]. Let  $T$  be a complete theory of the language  $L$ , languages  $L_1$  and  $L_2$  are extensions of  $L$  such that  $L_1 \cap L_2 = L$ , and theories  $T_1$  and  $T_2$  are consistent extensions of  $T$  in  $L_1$  and  $L_2$  respectively. Then  $T_3 = T_1 \cup T_2$  is a consistent theory.

Now we can formulate the following result. Let us introduce the following notation:

$$\overline{\Delta}_3 = \overline{\Delta}_1 \cup \overline{\Delta}_2 \cup P(c),$$

where the sentence  $P(c)$  means that the constant symbol  $c$  added to the language belongs to  $M = P(C_3)$ , i.e. this axiom refines the interpretation of  $P$  in semantic model  $C_3$  of the theory  $\Delta_3 = (\Delta_1)^f = (\Delta_2)^f$  in accordance with the position of  $c$  in  $C_3$ .

Theorem 10. i)  $\overline{\Delta}_3$  is consistent.

$$\text{ii) } (\overline{\Delta}_3)^f = \Delta_1^f = \Delta_2^f$$

The formulated results are described for the general situation in model theory and can be interpreted with using examples of differential algebra, namely, when considering the theory of differential fields of characteristic 0, the theory  $DCF_0$  of differentially closed fields of characteristic 0, the theory  $DF_p$  of differential fields of characteristic  $p$ , the theory  $DPF_p$  of differentially perfect fields of characteristic  $p$ , the theory  $DCF_p$  of differentially closed fields of characteristic  $p$ . In [1], it was proved that  $DF_0$  and  $DPF_p$  are perfect Jonsson theories,  $DCF_0$  and  $DCF_p$  are their centers correspondingly, while  $DF_p$  is not Jonsson. In addition,  $DF_0$  and  $DPF_p$  are strongly convex theories in the classical Robinson sense, which allows us to state that  $DF_0$  and  $DPF_p$  are Jonsson  $AP$ -theories.

Due to the above facts, we can apply the results described in the previous paragraph to the case of differentially closed fields of zero and positive characteristic. Here we assume that  $DF_0$  and  $DPF_p$  are hereditary Jonsson theories.

Based on the reasoning above, we can formulate the following results:

Theorem 11.  $\overline{DF_p}^f = DF_p^f$ .

Theorem 12.  $\overline{DPF_p}^f = DPF_p^f$ .

Theorem 13. i)  $\overline{DF_p} \cup \overline{DPF_p} \cup P(c)$  is consistent.

ii)  $(\overline{DF_p} \cup \overline{DPF_p} \cup P(c))^f = DF_p^f = DPF_p^f$ .

All concepts whose definitions are not given here can be found in [5] and [1].

*This research was funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP09260237).*

## References

1. Yeshkeyev, A.R. Connection between the amalgam and joint embedding properties / A.R. Yeshkeyev, I.O. Tungushbayeva, M.T. Kassymetova // Bulletin of the Karaganda University-Mathematics. — 2022. — 105. — No. 1. — P. 127–135.
2. Barwise, J. & Robinson, A. (1970). Completing Theories by Forcing. *Annals of Mathematical Logic*, 2(2), 119-142.
3. Барвайс, Дж. (1982). Теория моделей: справочная книга по математической логике. Часть 1. Москва: Издательство "Наука".
4. Forrest, W.K. (1977). Model theory for universal classes with the Amalgamation Property: A Study in the foundations of model theory and Algebra. *Annals of Mathematical Logic*, 11(3), 263–366.
5. Ешкеев, А.Р. & Касыметова, М.Т. (2016). Йонсоновские теории и их классы моделей. Караганда: Издательство КарГУ.
6. Yeshkeyev, A.R. & Ulbrikht, O.I. (2019). JSp-косемантичность  $R$ -модулей. *Siberian Electronic Mathematical Reports*, 16, 1233–1244.
7. Yeshkeyev, A.R. & Omarova, M.T. (2021). An essential base of the central types of the convex theory. *Bulletin of the Karaganda University-Mathematics*, 101(1), 119–126.

## ON ONE ESTIMATE OF PRIME NUMBERS

M.B. Gabbassov, T.Z. Kuanov  
System research company "Factor"  
[mars0@mail.ru](mailto:mars0@mail.ru)

French mathematician Pierre Dusart [1] proved the inequalities:

$$\frac{p_k}{k} < \ln(p_k) < \frac{p_k}{k} + 2, \quad (1)$$

where  $p_k - k^{\text{th}}$  a prime number. In this paper, we have improved estimates (1) and proved the inequalities

$$\frac{p_k}{k-1} + \frac{1}{2} < \ln(p_k) < \frac{p_k}{k} + \frac{3}{2}, \quad (2)$$

at  $k \geq 20$ .

In proving, we will use the well-known Rosser-Schoenfeld inequalities [2,3]

$$\frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)} < \pi(y) < \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)-1} \quad (3)$$

which are true at  $y \geq 67$ , where  $\pi(y)$  – a number of primes not exceeding  $y$ .

We introduce the following notations:

$$\varphi_0(p_k) = \prod_{i=1}^k p_i, \quad (4)$$

$$\varphi_1(p_k) = \prod_{i=1}^k (p_i - 1), \quad (5)$$

$$f_k(y) = \frac{y}{\ln(y/\sqrt{e})} - \frac{\varphi_1(p_k)}{\varphi_0(p_k)} y - k, \quad (6)$$

$$\sigma_k(y) = \eta_k(y) - \frac{\varphi_1(p_k)}{\varphi_0(p_k)} y, \quad (7)$$

where function  $f_k(y)$  is determined at  $y > e$ ,  $\eta_k(y)$  – a number of coprime numbers with a primorial  $\varphi_0(p_k)$ , not exceeding  $y$ , and  $p_k - k^{\text{th}}$  a prime number.

Consider the behavior of the function  $f_k(y)$  on the interval  $J = \left(1, e^{\frac{\varphi_0(p_k)}{\varphi_1(p_k)} + \frac{1}{2}}\right)$ . It is

increasing over the interval  $\left(\sqrt{e} \cdot e^{\left(\frac{1}{2} + \sqrt{\frac{1}{4} \frac{\varphi_1(p_k)}{\varphi_0(p_k)}}\right)}, \sqrt{e} \cdot e^{\left(\frac{1}{2} + \sqrt{\frac{1}{4} \frac{\varphi_1(p_k)}{\varphi_0(p_k)}}\right) \frac{\varphi_0(p_k)}{\varphi_1(p_k)}}\right)$  and at the end point

of this interval at  $y = \sqrt{e} \cdot e^{\frac{1 + \sqrt{1 - 4 \frac{\varphi_1(p_k)}{\varphi_0(p_k) \varphi_0(p_k)}}}{2} \frac{\varphi_0(p_k)}{\varphi_1(p_k)}}$  reaches its maximum value

$$f_{k,max} = \sqrt{e} \frac{1}{\left(1 + \sqrt{1 - 4 \frac{\varphi_1(p_k)}{\varphi_0(p_k)}}\right)^2} \left(\frac{\varphi_1(p_k)}{\varphi_0(p_k)}\right)^2 e^{\left(\frac{1}{2} + \sqrt{\frac{1}{4} \frac{\varphi_1(p_k)}{\varphi_0(p_k)}}\right) \frac{\varphi_0(p_k)}{\varphi_1(p_k)}} - k. \quad (8)$$

The function  $\sigma_k(y) = 1 - \frac{\varphi_1(p_k)}{\varphi_0(p_k)} y$  at  $y \in [1, p_{k+1})$ , further  $\sigma_k(y) = i + 1 - \frac{\varphi_1(p_k)}{\varphi_0(p_k)} y$  at  $y \in [p_{k+i}, p_{k+i+1})$  and  $y < p_{k+1}^2$ , since a set of the coprime numbers with a primorial  $\varphi_0(p_k)$  and a set of prime numbers on the interval  $(1, p_{k+1}^2)$  are coincided.

Let us use the well-known inequalities Dusart [3,4]

$$1,781 \ln(p_k) - 1,781 \frac{0.2}{\ln(p_k)} < \frac{\varphi_0(p_k)}{\varphi_1(p_k)} < 1,781 \ln(p_k) + 1,781 \frac{0.2}{\ln(p_k)},$$

from which it follows that

$$\frac{\varphi_0(p_k)}{e^{\frac{\varphi_0(p_k)}{\varphi_1(p_k)} + \frac{1}{2}}} < p_k^2. \quad (9)$$

Let us consider the difference on the interval  $J$

$$\sigma_{x_k}(y) - f_k(y) = i + 1 + k - \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)}, \quad (10)$$

where  $i+k$  – a number of primes, since interval  $J$  is contained within the interval  $J$ .

Lemma 1. At  $k \geq 20$  and  $p_{k+i} \leq y < p_{k+i+1}$ , and  $1 \leq i \leq i_{max}$ , a  $i_{max}$  such number that

$p_{k+i_{max}}$  is maximum prime number less than  $\sqrt{e} \cdot e^{\frac{1 + \sqrt{1 - 4 \frac{\varphi_1(p_k)}{\varphi_0(p_k)} \varphi_0(p_k)}}{2} \varphi_1(p_k)}$ , it is true

$$\ln\left(\frac{y}{\sqrt{e}}\right) - 1 < \frac{y}{k+i} < \ln\left(\frac{y}{\sqrt{e}}\right). \quad (11)$$

Proof. It follows from inequalities (3) that

$$1 < \sigma_{x_k}(y) - f_k(y) < 1 + \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right) - 1} - \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)}, \quad (12)$$

for all  $y > e^{\frac{5}{2}}$  and  $k \geq 20$ .

At a point  $p_{k+1}$   $\sigma_{x_k}(p_{k+1}) = 2 - \frac{\psi_1(x_k)}{\varphi(x_k)} p_{k+1}$ ,  $f_k(p_{k+1}) = \frac{p_{k+1}}{\ln\left(\frac{p_{k+1}}{\sqrt{e}}\right)} - \frac{\varphi_1(p_k)}{\varphi_0(p_k)} p_{k+1} - k$ .

Hence, by virtue of (12), it follows

$$\frac{p_{k+1}}{\ln\left(\frac{p_{k+1}}{\sqrt{e}}\right)} < k + 1 < \frac{p_{k+1}}{\ln\left(\frac{y}{\sqrt{e}}\right) - 1},$$

which is equivalent to

$$\ln\left(\frac{p_{k+1}}{\sqrt{e}}\right) - 1 < \frac{p_{k+1}}{k+1} < \ln\left(\frac{p_{k+1}}{\sqrt{e}}\right).$$

Further at  $p_{k+1} < y < p_{k+2}$   $\sigma_{x_k}(y) = 2 - \frac{\varphi_1(p_k)}{\varphi_0(p_k)} y$ ,  $f_k(y) = \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)} - \frac{\varphi_1(p_k)}{\varphi_0(p_k)} y - k$ , respectively

$$1 < k + 2 - \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)} < 1 + \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right) - 1} - \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)},$$

which is equivalent to

$$\ln\left(\frac{y}{\sqrt{e}}\right) - 1 < \frac{y}{k+1} < \ln\left(\frac{y}{\sqrt{e}}\right).$$

Continuing in the same way we have

$$1 < k + i + 1 - \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)} < 1 + \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right) - 1} - \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)}$$

at  $p_{k+i} \leq y < p_{k+i+1}$ , and  $i \leq i_{max}$ . The assertion of the lemma follows from the last inequality.

Lemma 1 is proved.

Lemma 2. At  $k \geq 20$  it is true

$$\ln\left(\frac{p_{k+1}}{\sqrt{e}}\right) - 1 < \frac{p_{k+1}}{k} < \ln\left(\frac{p_{k+1}}{\sqrt{e}}\right). \quad (13)$$

Proof. If  $e^2 < y < p_{k+1}$ , then  $\sigma_{x_k}(y) = 1 - \frac{\varphi_1(p_k)}{\varphi_0(p_k)} y$ , and  $f_k(y) = \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)} - \frac{\varphi_1(p_k)}{\varphi_0(p_k)} y - k$ .

Accordingly, due to (12),  $1 < \sigma_{x_k}(y) - f_k(y) = k + 1 - \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)} < 1 + \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right) - 1} - \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)}$  if

$k \geq 20$ , from which  $\frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right)} < k < \frac{y}{\ln\left(\frac{y}{\sqrt{e}}\right) - 1}$ . Let us pass to the limit at  $y \rightarrow p_{k+1}$  then we get

$\frac{p_{k+1}}{\ln\left(\frac{p_{k+1}}{\sqrt{e}}\right)} < k < \frac{p_{k+1}}{\ln\left(\frac{p_{k+1}}{\sqrt{e}}\right) - 1}$ , which is equivalent to the lemma.

Lemma 2 is proved.

Lemma 3. At  $k \geq 20$  the inequalities are true

$$\frac{p_k}{k} + \frac{1}{2} < \ln(p_k) < \frac{p_k}{k} + \frac{3}{2}, \quad (14)$$

$$\frac{p_k}{k-1} + \frac{1}{2} < \ln(p_k) < \frac{p_k}{k-1} + \frac{3}{2}. \quad (15)$$

Proof. The inequalities (14) follow from lemma 1 at  $i = 1$  and  $y = p_{k+1}$ , since  $p_{k+1} \in J$ . The inequalities (15) follow from lemma 2.

Lemma 3 is proved.

The validity of inequalities (2) automatically follows from Lemma 3.

## References

1. Safia Aoudjit, Djamel Berkane and Pierre Dusart : Explicit Estimates Involving the Primorial Integers and Applications. *Journal of Integer Sequences*, Vol. 24 (2021), Article 21.7.8.
2. Pierre Dusart : Estimates of  $\psi, \theta$  for large values of  $x$  without the Riemann hypothesis. *Math. Comput.* 85(298) (2016).
3. J. Barkley Rosser and Lowell Schoenfeld. Approximate formulas for some functions of prime numbers. *Illinois J. Math.* 6 (1962), 64–94.
4. Dusart, Pierre The  $k^{\text{th}}$  prime is greater than  $k(\ln k + \ln \ln k - 1)$  for  $k \geq 2$ . *Math. Comp.* 68 (1999), no. 225, 411--415. 11Y35 (11N05) MR1620223 (99d:11133)



## IT-ЖОБАЛАР МЕН СТАРТАПТАРДЫҢ ИНВЕСТИЦИЯЛЫҚ ТАРТЫМДЫЛЫҒЫН БАҒАЛАУ ЕРЕКШЕЛІКТЕРІН ЗЕРТТЕУ

*Кемелбек Майра Нариманқызы, 2 курс магистранты  
Алимжанова Лаура Муратбековна, т. э. к., профессоры  
Әл-Фараби атындағы Қазақ Ұлттық Университеті, Алматы қаласы  
e-mail: [kemelbekmaira.99@gmail.com](mailto:kemelbekmaira.99@gmail.com), [dimkim\\_01@mail.ru](mailto:dimkim_01@mail.ru)*

**Аннотация:** Қазіргі таңдағы қосымша табыс көздерінің бірі – инвестиция. Бұл процестің маңызды кезеңдерінің бірі – инвестициялық ресурстар салынатын жобаны таңдау. Осы мақсатта инвестициялық тартымдылыққа кешенді бағалау жүргізіледі. Бұл мақала IT-жобалар мен стартаптардың инвестициялық тартымдылығын бағалау ерекшеліктерін баяндауға арналады.

**Кілт сөздер:** IT-жоба, инвестициялық тартымдылық, IT-жоба мен стартапты бағалау.

**Кіріспе.** IT саласындағы инновациялық жобалар мен стартаптарды әзірлеу және іске асыру мәселелері ақпараттық-коммуникациялық технологиялардың кеңінен таралуы және цифрлық экономикаға көшу тұрғысынан өзекті бола түсуде. Осы кезекте, инвестициялық тартымдылықты қалыптастыру – компанияның ұзақ және тұрақты дамуына қол жеткізудің қажетті шарты. Инвестициялық тартымдылықтың жоғары деңгейі компаниялардың әртүрлі ресурстарға қол жеткізуін жеңілдетіп қана қоймайды, сонымен қатар компанияның сыртқы ортадағы мүдделерін сенімді қорғауды қамтамасыз етіп, менеджмент саласындағы шешімдердің негізділігіне әсер етеді.

**Негізгі бөлім.** Бүгінгі таңда бизнестің кез келген түрі бәсекелестіктің өте жоғары деңгейімен сипатталады. Осы ортада өмір сүріп қана қоймай, бәсекелестік позицияға ие болу үшін компаниялар үнемі дамып, ең жақсы әлемдік тәжірибелерді қабылдап, жаңа технологияларды игеріп, қызмет аясын кеңейтуі керек. Дәл осы динамикалық дамумен әрі қарай даму инвестициясыз мүмкін емес деген түсінік пайда болады. Осылайша, инвестициялар компанияға бәсекелестік артықшылық береді және көбінесе өсудің ең күшті құралы ретінде қызмет етеді. Инвесторлар үшін кәсіпорынның инвестициялық тартымдылығын талдау және бағалау өте маңызды, өйткені бұл дұрыс емес инвестициялау қаупін азайтуға мүмкіндік береді.

*Инвестициялық тартымдылық* – бұл "ұнату/ұнатпау" емес. Бұл қазіргі жағдай мен нарықтық жағдайды, ықтимал тәуекелдерді және белгілі бір объектінің кірістілігін бағалауды анықтайтын көрсеткіштер жиынтығы. Инвестициялық тартымдылық тек қаржылық құрылымдарға ғана емес, сонымен қатар нақты аймақтарға, елдерге, салаларға да байланысты екенін түсіну және есте сақтау қажет. Бұл көрсеткіш өзгермелі және көптеген факторлар оған әсер етуі мүмкін.

Инвестициялық тартымдылықты анықтайтын көрсеткіштердің бірі – *жобаның ерекшелігі*. Бұл IT-жобалар мен стартаптарды бағалаудың ең жалпы сипаттамасы. Бағалау кезінде сарапшы, біріншіден, жобаның немесе идеяның қаншалықты жаңа екендігіне, екіншіден, жобаның түпнұсқалығына назар аударады. Түпнұсқалық неғұрлым үлкен болса және оның дәрежесі неғұрлым жоғары болса, стартаптың бағасы соғұрлым жоғары болады. Алайда, ескеру керек: егер жоба тек бизнес немесе жеке тұлғалар тобы үшін түпнұсқа болса, онда ол инвестиция алуға құқылы, өйткені белгілі бір өнім немесе қызмет түрін құруға инвестиция салуға бағытталған немесе белгілі бір мақсатты аудиторияға бағытталған инвесторлар бар.

Келесі – *жобаның жаңалығы*. Бұл стартаптар мен IT-жобалардың негізгі сипаттамаларының бірі. Түпнұсқалықтан айырмашылығы, жоба жаңалығы – патенттік тазалықты, лицензиялық қорғауды, инновациялар бағытының басымдылығын болжайды.

*Жобаның пысықталуы*. IT-жобаның пысықталуы-стратегиялық және тактикалық мақсаттар қоя білу және оларға қол жеткізу; тәуекел факторларын есепке алу;

маркетингтік зерттеулердің болуы; бизнесті жүргізу тактикасын негіздеу және т.б. Егер жоба техникалық, экономикалық, ақпараттық, экологиялық, әлеуметтік және т. б. факторлардың алуан түрлілігін ескере отырып, көптеген элементтерді бір жүйеге біріктіре отырып, проблемаларды шешуде кешенді тәсілді іске асырса, пысықталған болып табылады.

Инвестициялық тартымдылық, ең алдымен, аумақта өнімді немесе қызметті құру үшін ресурстардың болуымен анықталады. Олар болмаған жағдайда жобаны тиімді іске асыру қаупі туындауы мүмкін. Арнайы шарттар – бұл жеке өнімдер мен қызметтерді жобалау мен өндіруге қойылатын талаптар жиынтығы. Ресурстарға мәлімделген қажеттілік жобаны іске асыру үшін қандай ресурстық әлеуеттің болуы (қаржылық, материалдық-техникалық, еңбек, ақпараттық және т.б.) негізделуге, көрсетілуге тиіс.

*Жобаның коммерциализациясы.* Инновацияларды коммерцияландыру – бұл табысты болған жағдайда болашақ пайдаға қатысу есебінен осы жаңалықты іске асыру бойынша қызметті қаржыландыру үшін инвесторларды тарту. Сонымен бірге, инновациялық жобаны нарыққа шығару процесі инновациялық қызметтің негізгі кезеңі болып табылады, содан кейін инновациялық өнімді әзірлеушінің (немесе иеленушінің) шығындары өтеледі және олардың қызметінен пайда алады. Коммерцияландыруды бағалау кезінде коммерциялық әлеуетті және игеруге дайындық дәрежесін ескеру қажет. Жобаның дайындық дәрежесі пысықталу сияқты критерий шеңберінде бағаланатындығына байланысты коммерциялық әлеует негізгі бағалауға ұшырауы тиіс.

**Дискуссия.** IT жобаларды бағалауға ықпал ететін заманауи үрдістерге тоқталсақ. Бизнес және IT әлеміндегі қазіргі тенденциялар бұл әдістің өзектілігін арттырады.

- Ақпараттық жүйелер күрделене түсуде, ал олардың компоненттері икемділікпен қайта пайдалануға мүмкіндік береді, нәтижесінде интеграцияланған жүйелер дамыды.

- Ақшаның барлық үлкен сомасы IT-жобалардың жұмысына бір уақытта тартылады.

- IT-жобаларға инвестициялар тарту, қазір стратегиялық инвестициялық портфельдің бөлігі ретінде қарастырылуда және осылайша ақша салудың басқа мүмкіндіктерімен бәсекелесуге тиіс.

Бұл заманауи үрдістер менеджмент арасында оң бағытта IT-жобаларды инвестициялық бағалаудың кешенді әдісіне деген теріс қатынасты бұрынғыдан да көбірек өзгертуді және тек IT-жобаның инвестициялық тартымдылығын бағалаудың кешенді құрылымдық әдісін қолданудың өзектілігін көрсетеді. Бұл әдістің бір нұсқасын IT саласының маманы Джон Харес және Дункан Роил ұсынған.

**Әдістер.** IT жобаны енгізудің инвестициялық тартымдылығын бағалау әдісі мынадай негізгі бөліктерден тұрады:

1. IT-ге қатысты компанияда жүзеге асырылуы мүмкін инвестициялар үшін мүмкіндіктерді анықтауға арналған біріктірілген тәсілдер жиынтығы.

2. Стратегиялық жоспарға кіретін IT жобаларды құндық бағалауға арналған интеграцияланған әдістемелер жиынтығы. Бағалау жобаның мынадай құрамдас бөліктеріне жүргізіледі:

- материалдық және материалдық емес құндылықтар түріндегі кірістер;
- тәуекелдің әр түрлі түрлері-нарықтық, күтпеген оқиғаларға байланысты жоба;
- жобаның иілісі арқылы тәуекелді азайту мүмкіндігі;
- көрінетін және жасырын шығындар.

**Қорытынды:** Бұл жұмыста оларды енгізер алдында IT-жобалар мен стартаптардың экономикалық тиімділігін бағалауды, бизнестің қолдану тәжірибесінің практикалық болмауы проблемасы атап өтілді. Сондай-ақ, IT-жобалардың инвестициялық тартымдылығын бағалау әдістерінің туындысының болуымен байланысты жағдай айтылды және бұл әдіс өзінің құрылымы мен пысықталуына қарамастан, бағдарламалық жасақтаманы әзірлеудің заманауи әдістері мен технологияларын ескермейді.

## Пайдаланылган әдебиеттер тізімі

1. Балабанов И.Т. Риск-менеджмент. М.: Финансы и статистика, 1996. 192 с.
2. Булгаков В.М., Спесивцев В.А., Баранова Л.А. Инвестиционная привлекательность проектных организаций: анализ и управление // Социально-экономические явления и процессы. 2013. № 4. С. 90–101.
3. Бухонова С.М., Дорошенко Ю.А. К оценке экономической эффективности инноваций разных типов // Экономический анализ: теория и практика. 2005. № 3. С. 13–18.
4. Инвестиционная привлекательность стартапа и её особенности / Статья [Электронный ресурс] Интернет-ресурс URL: <http://foykes.com/investitsionnaya-privlekatelnost-startapa-i-eyo-osobennosti/>
5. Факторы, определяющие инвестиционную привлекательность стартапа для инвестора, на макро и микроуровне / Статья [Электронный ресурс] URL: <https://pandia.ru/text/82/353/60724-10.php>
6. Карпов А.С., Сплендер В.А. Особенности формирования конкурентных преимуществ организаций наукоемких отраслей // Вестник Московского гуманитарно-экономического института. 2015. № 4. С. 16.
7. Полухин И.В. Необходимость и сущность коммерциализации инноваций предприятий ракетно-космической промышленности // Менеджмент социальных и экономических систем. 2016. №3. URL: <https://cyberleninka.ru/article/n/neobhodimost-i-suschnost-kommertsializatsii-innovatsiy-predpriyatiy-raketno-kosmicheskoy-promyshlennosti>
8. Сплендер В.А., Гребеник Т.В. Инновационные механизмы кредитования субъектов малого и среднего предпринимательства // В сборнике: Научные труды преподавателей МАЭП Московская академия экономики и права. Москва, 2014. С. 33–39.
9. Дайан А., Букерель Ф., Ланкар Р. Академия рынка: маркетинг / пер. с фр. М.: Экономика, 1993. — 572 с.
10. Каширин А. И., Семенов А. С. Инновационный бизнес: венчурное и бизнес-ангельское инвестирование: учеб. пособие. М.: Дело, РАНХиГС, 2012. — 260 с.

# SIMILARITIES OF HYBRIDS OF JONSSON QUASIVARIETIES

*Yeshkeyev A.R., Ulbrikht O.I., Mussina N.M.*

*Karaganda Buketov University, Karaganda, Kazakhstan*

*E-mail: [aibat.kz@gmail.com](mailto:aibat.kz@gmail.com), [ulbrikt@mail.ru](mailto:ulbrikt@mail.ru), [mussinazerke@gmail.com](mailto:mussinazerke@gmail.com)*

This abstract considers new approaches to the classification of classes of structures whose theories form a natural subclass of the class of inductive theories. This subclass satisfies the natural conditions of joint embedding and amalgam. Theorems related to the classical theorems of Model Theory are given in the framework of the study of Jonsson spectra of special classes of structures. These classes may contain expands in Jonsson manner, respectively, the class of varieties or the class of quasivarieties. We called such classes Jonsson varieties and, accordingly, Jonsson quasivarieties. The main result of this work is the following statement: within the framework of a fixed Jonsson spectrum with respect to the cosemanticness property for a hybrid of classes of the Jonsson spectrum of quasivarieties the fixed Jonsson theory of polygons (S-acts) is found, which defines this hybrid up to Jonsson syntactic similarity.

We will give the necessary definitions of concepts concerning the main result.

*Definition 1.* A theory  $T$  is called a Jonsson theory, if

1.  $T$  has an infinite model;
2.  $T$  is inductive, i.e.  $T$  equivalent to the set of  $\forall\exists$ -sentences;
3.  $T$  has the joint embedding property (JEP);
4.  $T$  has the amalgamation property (AP), that is if for any  $\mathcal{A}, \mathcal{B}, \mathcal{C} \models T$  such that  $f_1: \mathcal{A} \rightarrow \mathcal{B}, f_2: \mathcal{A} \rightarrow \mathcal{C}$  are isomorphic embeddings, there are  $\mathcal{D} \models T$  also isomorphic embeddings  $g_1: \mathcal{B} \rightarrow \mathcal{D}, g_2: \mathcal{C} \rightarrow \mathcal{D}$  such that  $g_1 f_1 = g_2 f_2$ .

*Definition 2.* Let  $k \geq \omega$ . A model  $\mathcal{M}$  of  $T$  is called  $k$ -universal for  $T$  if each model  $T$  of cardinality strictly less  $k$  is isomorphic ally embeddable in  $\mathcal{M}$ .

*Definition 3.* A semantic model  $\mathcal{C}_T$  of Jonsson theory  $T$  is called  $\omega^+$ -homogeneous-universal model of  $T$ .

*Definition 4.* Let  $T_1$  and  $T_2$  are an arbitrary Jonsson theories. We say that  $T_1$  and  $T_2$  are Jonsson syntactically similar if exists a bijection  $f: E(T_1) \rightarrow E(T_2)$  such that:

- 1) restriction  $f$  to  $E_n(T_1)$  is isomorphism of lattices  $E_n(T_1)$  and  $E_n(T_2)$ ,  $n < \omega$ ;
- 2)  $f(\exists v_{n+1}\varphi) = \exists v_{n+1}f(\varphi), \varphi \in E_{n+1}(T), n < \omega$ ;
- 3)  $f(v_1 = v_2) = (v_1 = v_2)$ .

*Definition 5.* The pure triple  $\langle C; Aut(C), Sub(C) \rangle$  is called the Jonsson semantic triple, where  $C$  is universe of semantic model  $\mathcal{C}$  of theory  $T$ ,  $Aut(C)$  is the automorphism group of  $C$ ,  $Sub(C)$  is a class of all subsets of  $C$  which are universe of the corresponding existentially closed submodels of  $C$ .

*Definition 6.* Two Jonsson theories  $T_1$  and  $T_2$  are called the Jonsson semantically similar if their Jonsson semantic triples are isomorphic as pure triples.

*Definition 7.* 1) Let  $T_1$  and  $T_2$  be some Jonsson theories of the countable language  $L$  of the same signature  $\sigma$ ;  $\mathcal{C}_1$  and  $\mathcal{C}_2$  are their semantic models, respectively. In the case of common signature of Jonsson theories  $T_1, T_2$ , let us call a hybrid of Jonsson theories  $T_1$  and  $T_2$  of the first type the following theory  $Th_{\forall\exists}(\mathcal{C}_1 \diamond \mathcal{C}_2)$  if that theory is Jonsson in the language of signature  $\sigma$  and denote it by  $H(T_1, T_2)$ , where the operation  $\diamond \in \{\times, +, \oplus, \prod_F, \prod_U\}$  and  $\mathcal{C}_1 \diamond \mathcal{C}_2 \in Mod\sigma$ . Here  $\times$  means cartesian product,  $+$  means sum,  $\oplus$  means direct sum,  $\prod_F$  – means reduced product and  $\prod_U$  – means ultraproduct of models. Herewith, the algebraic construction  $(\mathcal{C}_1 \diamond \mathcal{C}_2)$  is called a semantic hybrid of the theories  $T_1, T_2$ .

2) If  $T_1$  and  $T_2$  are Jonsson theories of different signatures  $\sigma_1$  and  $\sigma_2$ , then  $H(T_1, T_2) = Th_{\forall\exists}(\mathcal{C}_1 \diamond \mathcal{C}_2)$  will be called a hybrid of the second type, if that theory is Jonsson in the language of signature  $\sigma = \sigma_1 \cup \sigma_2$  where  $\mathcal{C}_1 \diamond \mathcal{C}_2 \in Mod\sigma$ .

Let  $K$  be the class of structures of countable signature  $\sigma$ . Let's introduce the notation:

$$\forall\exists(K) = Th(K) \cup \{\varphi \mid \varphi \text{ is a } \forall\exists\text{-sentence of considered language and}$$

$\varphi \cup Th(K)$  is a consistent}

*Definition 8.* A variety (quasivariety) of structures  $\mathcal{K}$  is called a Jonsson variety (quasivariety) if  $\forall\exists(K)$  is a Jonsson theory.

Consider the  $JSpV(K)$  be Jonsson spectrum of the Jonsson variety of class  $K$ , where  $K$  is the Jonsson variety:

$$JSpV(K) = \{T | T = \forall\exists(N) \text{ is Jonsson theory, } N \text{ is a subvariety of } K\}.$$

Then  $JSpV(K)/\approx$  is denoting the factor set of the Jonsson spectrum of Jonsson quasivariety of the class  $K$  by the relation  $\approx$ .

Similarly, we define the Jonsson spectrum of  $JSpQV(K)$  quasivariety:

$$JSpQV(K) = \{T | T = \forall\exists(N) \text{ is Jonsson theory, } N \text{ is a subquasivariety of } K\}$$

Then  $JSpQV(K)/\approx$  is denoting the factor set of the Jonsson spectrum of Jonsson quasivariety of the class  $K$  by the relation  $\approx$ .

The main results are the following theorems.

**Theorem 1.** Let  $K$  be some Jonsson quasivariety of structures of signature  $\sigma$ ,  $[T_1], [T_2] \in JSpQV(K)/\approx$ . For any perfect complete for  $\exists$ -sentences hybrid  $H([T_1], [T_2])$  there is a Jonsson  $\exists$ -complete theory of the polygon  $T_{\Pi}'$  such that  $H([T_1], [T_2]) \stackrel{s}{\sim} T_{\Pi}'$ .

All concepts that are not defined in this abstract can be extracted from [1-3].

*This work was supported by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (grant AP09260237).*

#### References

1. Yeshkeyev A.R., On Jonsson varieties and quasivarieties. Bulletin of the Karaganda university, Mathematics series 104:4 (2021), 151–157.
2. Yeshkeyev A.R., Mussina N.M., Properties of hybrids of Jonsson theories. Bulletin of the Karaganda university, Mathematics series 92:4 (2018), 99-104.
3. Yeshkeyev, A.R., Kassymetova, M.T., (2016), Jonsson theories and their classes of models, Karaganda, KarGU, 370 p.

## ON SMALL MODELS IN JONSSON THEORY

*Yeshkeyev A.R., Ulbrikht O.I., Issayeva A.K.*  
*Karaganda Buketov University, Karaganda, Kazakhstan*  
*E-mail: [aibat.kz@gmail.com](mailto:aibat.kz@gmail.com), [ulbrikt@mail.ru](mailto:ulbrikt@mail.ru), [isa\\_aiga@mail.ru](mailto:isa_aiga@mail.ru)*

This work is related to the well-known problems which were appeared regarding a description of algebraically prime models. In this paper, the properties of algebraically prime and various types of atomic models are considered in the frame of the study of Jonsson theories. One of the important concepts of this article is a definable set with various additional properties that define the corresponding types of prime and atomic models. The defining formula of this set makes it possible to define the pregeometry on definable subsets of the semantic model of the considered Jonsson theory. Also was adapted for fixed Jonsson theories the existence of a prime model over a basic set.

We will give the necessary definitions of concepts concerning the main result.

*Definition 1.* A set  $A$  is called the  $(\Gamma_1, \Gamma_2)$ -cl-atomic in the theory  $T$  if

1)  $cl(A) = M, M \in E_T$ , where  $E_T$  is the class of existentially closed models of the theory  $T$ ;

2)  $\forall \bar{a} \in A$ , there exists  $\varphi(\bar{x}) \in \Gamma_1$  such that  $M \models \varphi(\bar{a})$ ;

3)  $\varphi(\bar{x})$  generates  $t_{\Gamma_2 \cup \Gamma_2^*}^M(\bar{a})$

and obtained model  $M$  is called the  $(\Gamma_1, \Gamma_2)$ -cl-atomic model of the theory  $T$ .

*Definition 2.* A set  $A$  is called the weakly  $(\Gamma_1, \Gamma_2)$ -cl-atomic in the theory  $T$  if

1)  $cl(A) = M, M \in E_T$ , where  $E_T$  is the class of existentially closed models of the theory  $T$ ;

2)  $\forall \bar{a} \in A$ , there exists  $\varphi(\bar{x}) \in \Gamma_1$  such that  $M \models \varphi(\bar{a})$ ;

3)  $\varphi(\bar{x})$  generates  $t_{\Gamma_2}^M(\bar{a})$

and obtained model  $M$  is called the weakly  $(\Gamma_1, \Gamma_2)$ -cl-atomic model of the theory  $T$ .

*Definition 3.* A set  $A$  is called the almost  $(\Gamma_1, \Gamma_2)$ -cl-atomic in the theory  $T$  if for any  $\bar{a} \in A$  there exists a formula  $\varphi(\bar{x}) \in \Gamma_1$  such that:

1)  $(A) = M, M \in E_T$ , where  $E_T$  the class of existentially closed models of the theory  $T$ ;

2)  $\varphi(\bar{x}) \cup T$  is consistent;

3)  $\varphi(\bar{x})$  generates  $t_{\Gamma_2 \cup \Gamma_2^*}^M(\bar{a})$ ,

and obtained model  $M$  is called the almost  $(\Gamma_1, \Gamma_2)$ -cl-atomic model of the theory  $T$ .

*Definition 4.* A set  $A$  is called the almost-weakly  $(\Gamma_1, \Gamma_2)$ -cl atomic in the theory  $T$  if for any  $\bar{a} \in A$  there exists a formula  $\varphi(\bar{x}) \in \Gamma_1$  such that:

1)  $(A) = M, M \in E_T$ , where  $E_T$  the class of existentially closed models of the theory  $T$ ;

2)  $\varphi(\bar{x}) \cup T$  is consistent;

3)  $\varphi(\bar{x})$  generates  $t_{\Gamma_2}^M(\bar{a})$

and obtained model  $M$  is called the almost-weakly  $(\Gamma_1, \Gamma_2)$ -cl-atomic model of the theory  $T$ .

*Definition 5.* A set  $A$  is called the  $(\Gamma_1, \Gamma_2)$ -cl-algebraically prime in a theory  $T$  if  $cl(A) \in M, M$  is a  $(\Gamma_1, \Gamma_2)$ -cl-atomic model of the theory  $T, M \in E_T \cap AP_T$ , where  $AP_T \cap E_T \neq \emptyset$ , and the resulting model  $M$  is called the  $(\Gamma_1, \Gamma_2)$ -cl-algebraically prime model of the theory  $T$ .

*Definition 6.* A set  $A$  is called the weakly  $(\Gamma_1, \Gamma_2)$ -cl-algebraically prime in a theory  $T$  if  $cl(A) \in M, M$  is a weakly  $(\Gamma_1, \Gamma_2)$ -cl-atomic model of the theory  $T, M \in E_T \cap AP_T$ , where  $AP_T \cap E_T \neq \emptyset$ , and the resulting model  $M$  is called the weakly  $(\Gamma_1, \Gamma_2)$ -cl-algebraically prime model of the theory  $T$ .

*Definition 7.* A set  $A$  is called the almost  $(\Gamma_1, \Gamma_2)$ -cl-algebraically prime in a theory  $T$  if  $cl(A) \in M, M$  is an almost  $(\Gamma_1, \Gamma_2)$ -cl-atomic model of the theory  $T, M \in E_T \cap AP_T$ , where

$AP_T \cap E_T \neq \emptyset$ , and the resulting model  $M$  is called the almost  $(\Gamma_1, \Gamma_2)$ -cl-algebraically prime model of the theory  $T$ .

*Definition 8.* A set  $A$  is called the almost-weakly  $(\Gamma_1, \Gamma_2)$ -cl-algebraically prime in a theory  $T$  if  $cl(A) \in M$ ,  $M$  is an almost-weakly  $(\Gamma_1, \Gamma_2)$ -cl-atomic model of the theory  $T$ ,  $M \in E_T \cap AP_T$ , where  $AP_T \cap E_T \neq \emptyset$ , and the resulting model  $M$  is called the almost-weakly  $(\Gamma_1, \Gamma_2)$ -cl-algebraically prime model of the theory  $T$ .

The following definitions 9-16 were taken from [2].

*Definition 9.* The formula  $\Phi(\bar{x}; \bar{y})$  of the language  $L$  is called normal for  $\bar{x}$  if for any  $\bar{b}, \bar{c} \in C$  the sets  $\Phi(C; \bar{b})$  and  $\Phi(C; \bar{c})$  either coincide or do not intersect. A formula  $\Phi(\bar{x})$  is called normal if it is normal for any tuple of variables  $\bar{x}'$  whose elements belong to tuple  $\bar{x}$ .

*Definition 10.* A set of formulas BF is called basic if it consists of normal formulas, contains equalities of variables, is closed under renaming of variables, conjunction, and P - operation. If  $\Phi(\bar{x}; \bar{y}) \in \text{BF}$  is a basic formula of the language  $L$ , then a formula of the form  $\Phi(\bar{x}; \bar{a})$  for some tuple  $\bar{a} \in A$  is called a basic formula over  $A$  and is simply denoted by  $\Phi(\bar{x})$ .

*Definition 11.* A set of tuples  $D$  is said to be positive over  $A$  if  $D = \Phi(C)$  for some basic formula  $\Phi(\bar{x})$  over  $A$ .

*Definition 12.*

(a) The set  $t$  consisting of basic formulas over  $A$  with free variables from the set  $X$  and their negations is called the basic type over  $A$  from  $X$ ;

(b) If  $t$  is a type, then  $t^+$  denotes the set of all basic formulas in the type  $t$  and is called the positive part of the type  $t$ . We will denote by  $t^-$  the set of all negations of the basic formulas included in the type  $t$ ;

(c) A consistent type  $t$  in  $X$  variables is called basic complete over  $A$  in  $X$  if  $t \vdash \Phi$  or  $t \vdash \neg \Phi$  holds for any basic formula  $\Phi$  over  $A$  in  $X$ . A basis type  $t$  that is fundamentally complete over  $A$  from  $X$  and is closed under the deducibility of basic formulas over  $A$  from  $X$  and their negations is called maximal over  $A$  from  $X$ . The maximality of the basic type  $t$  means that among the consistent basic types over  $A$  from  $X$  there is no proper extension of the type  $t$ ;

(d) If for a basic type  $t$  over  $A$  in variables  $X$  there exists a type  $q \cup t^- \vdash t$  of cardinality less than  $\lambda$  and  $(q \cup t^+) \vdash t$  holds, then the type  $t$  is called a  $\lambda$ -positive type over  $A$  in variables  $X$ . Moreover, the type  $q$  will be called the  $\lambda$ -base (over  $A$  from  $X$ ) of the  $\lambda$ -positive type  $t$  over  $A$  from  $X$ . If  $\lambda = 1$ , then the  $\lambda$ -positive type  $t$  is called the positive type over  $A$  from  $X$ ;

(e) A tuple  $\bar{a}$  of length  $n$  is said to be  $\lambda$ -positively isolated over a set  $A$  if its basic type over  $A$  is  $\lambda$ -positive. For  $\lambda = 1$ , this notion will be called positive isolation;

(f) If  $t$  is a type over  $A$  and  $f$  is a mapping from  $A$  to  $C$ , then  $f(t)$  denotes the type obtained from the type  $t$  by replacing the parameters  $a$  in its formulas with their images  $f(a)$ .

*Definition 13.* A sequence  $S = \langle a_\alpha \mid \alpha < \kappa \rangle$  is called a  $\lambda$  positive construction over  $A$  if, for any  $\alpha < \kappa$ , the element  $a_\alpha$  is  $\lambda$ -positively isolated over the set  $(A \cup \{a_\beta \mid \beta < \alpha\})$ , where  $\lambda$  is from item (d) of Definition 12. Moreover, the sequence  $S$  is called the  $\lambda$ -positive construction (over  $A$ ) of the set  $B = \cup S$ .

*Definition 14.* A set  $B$  is called  $\lambda$ -positively constructible (or simply  $\lambda$ -constructible for  $\lambda=1$ ) over  $A$  if there exists a positive construction  $\langle a_\alpha \mid \alpha < \kappa \rangle$  over  $A$  such that  $B = (\cup \langle a_\alpha \mid \alpha < \kappa \rangle \cup A)$ .

*Definition 15.* A set  $A$  is called positively compact if any positive type over  $A$  is realized in  $A$ .

*Definition 16.* Let  $A \subseteq B$ . A set  $B$  is called the positive hull of a set  $A$  if it is positively compact and positively constructible over  $A$ .

The main results are the following theorems.

**Theorem 1.** Any set  $A$  has its positive hull.

**Theorem 2.** Let  $T$  be a perfect Jonsson theory complete for  $\Pi_1$  -sentences, then  $T$  has a nice almost-weakly  $(\Sigma_1, \Sigma_1)$ -cl-atomic model.

**Theorem 3.** Let  $T$  be a perfect Jonsson theory complete for  $\Pi_1$  –sentences. Then the following conditions are equivalent:

1)  $\mathfrak{A}$  –  $(\Sigma_1, \Sigma_1)$ -cl-algebraically prime model of the theory  $T$ .

2)  $\mathfrak{A}$  is a nice almost-weak  $(\Sigma_1, \Sigma_1)$ -cl-atomic model of  $T$ .

All concepts that are not defined in this abstract can be extracted from [1-3].

*This work was supported by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (grant AP09260237).*

#### References

1. Baldwin, J.T., Kueker, D.W., (1981), Algebraically prime models, *Ann. Math. Logic*, 20, pp.289–330.

2. Palutin, E.A., (2007), Positively prime models over a normal basic set, *Siberian Electronic Mathematical Reports*, 4, pp.596–604.

3. Yeshkeyev, A.R., Kassymetova, M.T., (2016), Jonsson theories and their classes of models, Karaganda, KarGU, 370 p.



## УПОРЯДОЧЕННАЯ СТАБИЛЬНОСТЬ И ОБОГАЩЕНИЯ ЧИСТОГО ЛИНЕЙНОГО ПОРЯДКА ОДНОМЕСТНОЙ ФУНКЦИЕЙ

В. В. Вербовский, А. Д. Еришгешова

Казахский национальный исследовательский университет имени К.И. Сатпаева, Алматы,  
[viktor.verbovskiy@gmail.com](mailto:viktor.verbovskiy@gmail.com)

Университет имени Сулеймана Демиреля, Каскелен

[aisha.yershigeshova@gmail.com](mailto:aisha.yershigeshova@gmail.com)

С момента появления понятия о-минимальности изучение линейно упорядоченных структур стало активно развиваться в теории моделей. Поскольку понятие о-минимальности оказалось весьма продуктивным, появились разнообразные его обобщения. Одним из них стало понятие упорядоченной стабильности, введённое в работе [1].

Следующие примеры показывают, что не всякое обогащение чистого линейного порядка одноместной функцией сохраняет упорядоченную стабильность.

**Пример 1.**  $(\mathbb{Q} \times \mathbb{Q}, <, f^1)$ , где порядок  $<$  лексикографический, а функция задана следующим образом:  $f((q_1, q_2)) = (q_2, q_1)$ . Тогда можно определить обратный лексикографический порядок:  $a < b \Leftrightarrow f(a) < f(b)$ . В работе [1] было доказано, что элементарная теория такой структуры  $(\mathbb{Q} \times \mathbb{Q}, <, <)$  является зависимой (не обладает свойством независимости), но не является упорядоченно стабильной. Заметим, что здесь функция не является локально монотонной, если под локальной монотонностью понимать, что ограничение на выпуклое подмножество является монотонной функцией.

**Пример 2.**  $\mathcal{M} = (\mathbb{Z} \times \mathbb{Z} + \mathbb{Q}, <, f^1)$ , где порядок на  $\mathbb{Z} \times \mathbb{Z}$  лексикографический, любой элемент из  $\mathbb{Q}$  больше любого элемента из  $\mathbb{Z} \times \mathbb{Z}$ , а функция задана, как описано ниже. Для начала определим функции  $g_1, g_2$  из  $\mathbb{Z} \times \mathbb{Z}$  в  $\mathbb{Q}$ , как  $g_i((z_1, z_2)) = z_i$ . Теперь определим функцию  $f$ :

$$\begin{aligned} f((z_1, z_2)) &= g_2((z_1, z_2)) + \frac{1}{g_1((z_1, z_2)) + 2}, \text{ если } z_1 > 0 \\ f((z_1, z_2)) &= g_2((z_1, z_2)), \text{ если } z_1 = 0 \\ f((z_1, z_2)) &= g_2((z_1, z_2)) + \frac{1}{g_1((z_1, z_2)) - 2}, \text{ если } z_1 < 0 \\ f(q) &= (0, 0) \in \mathbb{Z} \times \mathbb{Z} \end{aligned}$$

Очевидно, что множество  $\mathbb{Q}$  формульно:

$$Q(x) = \forall y, z (y < x < z \rightarrow \exists u, v (y < u < x < v < z))$$

Отсюда легко понять, что  $\neg Q$  выделяет множество  $\mathbb{Z} \times \mathbb{Z}$ .

Рассмотрим формулу  $\varphi(z, q) = f(z) > q$  и сечение  $s = (C, D)$ , где  $C = \mathbb{Z} \times \mathbb{Z}$ ,  $D = \mathbb{Q}$ . Поймём, что формула  $\varphi(z, q)$  обладает свойством строгого порядка внутри сечения  $s$ . Пусть  $q = n \in Q(\mathcal{M})$ . Тогда  $\varphi(\mathcal{M}, q)$  будет объединением следующих выпуклых множеств. Рассмотрим  $(z_1, z_2) \in \{z_1\} \times \mathbb{Z}$ , где  $z_1 > 0$ . Тогда  $f((z_1, z_2)) = g_2((z_1, z_2)) + \frac{1}{g_1((z_1, z_2)) + 2} = z_2 + \frac{1}{z_1 + 3}$ . Очевидно, что  $f((z_1, z_2)) > n$  тогда и только тогда, когда  $z_2 > n$ . Следовательно,

$$\varphi(\mathcal{M}, n) = \bigcup_{z_1 \in \mathbb{Z}} [(z_1, n), +\infty),$$

где  $[(z_1, n), +\infty) \subseteq \{z_1\} \times \mathbb{Z}$ .

По определению, формула  $\varphi(z, q)$  обладает свойством строгого порядка внутри сечения  $s$ , если существует последовательность  $q_i, i < \omega$ , такая что для любого  $c \in C$  и  $d \in D$  имеет место  $\varphi(\mathcal{M}, q_i) \cap (c, d) \subset \varphi(\mathcal{M}, q_j) \cap (c, d)$  тогда и только тогда, когда  $i < j$

[2]. В работе [2] было доказано, что упорядоченно стабильная теория не обладает свойством строгого порядка внутри никакого сечения.

Докажем, что  $\varphi(z, n)$  обладает свойством строгого порядка внутри сечения  $s$ . Пусть  $q_n = n \in P(\mathcal{M})$ , где  $n \in \omega$ . Пусть  $c = (\zeta_1, \zeta_2), d = q$ . Тогда

$$\varphi(\mathcal{M}, q_n) \cap (c, d) = \left( ((\zeta_1, \zeta_2), +\infty) \cap [(\zeta_1, n), +\infty) \right) \cup \bigcup_{z_1 \in \mathbb{Z}, z_1 > \zeta_1} [(z_1, n), +\infty)$$

Легко понять теперь, что  $\varphi(\mathcal{M}, q_n) \cap (c, d) \subset \varphi(\mathcal{M}, q_{n+1}) \cap (c, d)$  для каждого  $n$ . Таким образом, элементарная теория структуры  $(\mathbb{Z} \times \mathbb{Z} + \mathbb{Q}, <, f^1)$  не является упорядоченно стабильной, при этом легко понять, что  $f$  локально монотонна.

Рассмотрим сигнатуру  $\Sigma = \{=, <, f^1, P_i^1\}_{i \in I}$ . Мы будем рассматривать линейно упорядоченные структуры, где  $f$  — непрерывная монотонная функция, а  $P_i$  — одноместные предикаты.

Рассмотрим аксиомы теории  $T$ :

- 1)  $<$  — линейный порядок;
- 2)  $\forall x_1 \forall x_2 (x_1 < x_2 \rightarrow f(x_1) < f(x_2))$  ( $f$  возрастающая);
- 3)  $\forall x \forall y_1 \forall y_2 (y_1 < f(x) < y_2 \rightarrow \exists x_1 \exists x_2 (x_1 < x < x_2 \wedge \forall t (x_1 < t < x_2 \rightarrow y_1 < f(t) < y_2))$ ) ( $f$  непрерывна);
- 4)  $\forall x (x < f(x))$ .

Пусть  $R(y) := \exists x (y = f(x))$  и  $D(y) = \exists y (y = f(x))$ .

Если  $A$  — формульное подмножество, то  $A^n = f^n(A)$ :

$$y \in A^n \Leftrightarrow \exists x \in A (f^n(x) = y), \text{ если } n > 0$$

$$y \in A^0 \Leftrightarrow y \in A$$

$$y \in A^{-n} \Leftrightarrow \exists x \in A (f^n(y) = x), \text{ если } n > 0$$

Пусть  $\Sigma^+ = \{=, <, P_i^n, R^n, D^n\}_{i \in I, n \in \mathbb{Z}}$

**Факт 1. [1]** Пусть  $\mathcal{M} = (M, \Sigma^+)$  — линейно упорядоченная структура с одноместными предикатами. Тогда ее элементарная теория упорядоченно суперстабильна.

Рассмотрим структуру  $\mathcal{M}_1 = (M, \Sigma^+ \cup \{f^1\})$ . Поскольку факт 1 верен для любой сигнатуры  $\Sigma^+$ , состоящей из символа отношения порядка и одноместных предикатов, то мы можем считать, что образ любого одноместного предиката относительно  $f$  есть некоторый одноместный предикат из  $\Sigma^+$ , и то же самое можно сказать про прообраз любого одноместного предиката из  $\Sigma^+$ . Более того, будем считать, что  $\text{dom } f, \text{range } f \in \Sigma^+$ .

Пусть  $T = Th(\mathcal{M})$  и пусть  $\Sigma_{qe}^+ \supset \Sigma^+$  — такая сигнатура, что теория  $T_{qe}^+ = Th(M, \Sigma_{qe}^+)$  допускает элиминацию кванторов. Далее, сделаем еще одно предположение. Будем рассматривать только такие структуры  $\mathcal{M} = (M, \Sigma^+)$ , что  $\Sigma_{qe}^+$  состоит только из одноместных предикатов и символа отношения порядка. Для удобства будем считать, что если  $P, R \in \Sigma_{qe}^+$  — два одноместных предиката, то  $\neg P, P \wedge Q$  тоже лежат  $\Sigma_{qe}^+$ .

Пусть  $T_{qe}^f = Th(M, \Sigma_{qe}^+ \cup \{f\})$ . Обозначим,  $\Sigma_{qe}^f = \Sigma_{qe}^+ \cup \{f\}$

**Теорема 1.** Теория  $T_{qe}^f$  допускает элиминацию кванторов.

Доказательство. Рассмотрим термы сигнатуры  $\Sigma_{qe}^f$ . Они имеют вид: либо переменная, либо  $f^n(x)$ , где  $n \in \mathbb{Z}^+$ . Для удобства будем использовать  $f^n(x)$ , где  $n \in \mathbb{Z}$ , а в конце рассуждений покажем, как избавиться от отрицательных  $n$ . В силу критерия Тарского достаточно показать, как сокращать экзистенциальный квантор в формулах вида

$$\exists x \left( \bigwedge_{i=1}^k \varphi_i(x, \bar{y}_i) \right),$$

где  $\varphi_i(x, \bar{y}_i)$  — атомарная формула или ее отрицание, то есть формула одного из следующих видов:

- 1)  $f^n(x) = f^m(y)$ ;
- 2)  $f^n(x) \neq f^m(y)$ ;
- 3)  $f^n(x) < f^m(y)$ ;
- 4)  $f^n(x) \not< f^m(y)$ ;
- 5)  $f^n(x) > f^m(y)$ ;
- 6)  $f^n(x) \not> f^m(y)$ ;
- 7)  $U(f^n(x))$ , где  $U$  — одноместный предикат из  $\Sigma_{qe}^+$ ;

В силу введенных обозначений,  $U(f^n(x)) \Leftrightarrow U^{-n}(x)$ , таким образом, можно рассматривать среди формул вида (7) только формулы вида  $U(x)$ .

Заметим, что можно избавиться от формул вида с отрицанием (от формул вида 2, 4, 6). Действительно,

$$f^n(x) \neq f^m(y) \Leftrightarrow (f^n(x) < f^m(y)) \vee (f^n(x) > f^m(y)).$$

Обозначим конъюнкцию вида  $\exists x(\bigwedge_{i=1}^k \varphi_i(x, \bar{y}_i))$  как  $\theta$ . Тогда

$$\begin{aligned} \exists x(f^n(x) \neq f^m(y) \wedge \theta) &\Leftrightarrow \exists x \left( \left( (f^n(x) < f^m(y)) \vee (f^n(x) > f^m(y)) \right) \wedge \theta \right) \Leftrightarrow \\ &\Leftrightarrow \exists x \left( \left( (f^n(x) < f^m(y)) \wedge \theta \right) \vee \left( (f^n(x) > f^m(y)) \wedge \theta \right) \right) \Leftrightarrow \\ &\Leftrightarrow \exists x \left( (f^n(x) < f^m(y)) \wedge \theta \right) \vee \exists x \left( (f^n(x) > f^m(y)) \wedge \theta \right) \end{aligned}$$

Далее каждый дизъюнкт можно рассматривать по отдельности.

Так как  $f^n(x) \not< f^m(y) \Leftrightarrow (f^n(x) = f^m(y)) \vee (f^n(x) > f^m(y))$ , рассуждения аналогичны. Подобным образом можно избавиться и от формул вида 6.

Заметим, что в силу того, что  $f$  возрастающая,  $f^n(x) \geq f^m(y)$  эквивалентна  $f^{n+s}(x) \geq f^{m+s}(y)$  для любого  $s$ , где  $\geq \in \{=, <, >\}$  (то есть мы записали 3 эквивалентности). Поэтому не умаляя общности, можно считать, что формулы вида 1, 3 и 5 имеют вид:  $x = f^m(y)$ ;  $x < f^m(y)$ ;  $x > f^m(y)$ .

Если есть формула вида 1, то есть  $x = f^m(y)$ , то если  $m \geq 0$ , получаем:

$$\exists x \left( x = f^m(y) \wedge \bigwedge_{i=1}^k \varphi_i(x, \bar{y}_i) \right) \Leftrightarrow D(y) \wedge \bigwedge_{i=1}^k \varphi_i(f^m(y), \bar{y}_i)$$

Рассмотрим случай, когда  $m < 0$ . Пусть  $m = -n$ . Тогда

$$\begin{aligned} \exists x \left( x = f^{-n}(y) \wedge \bigwedge_{i=1}^k \varphi_i(x, \bar{y}_i) \right) &\Leftrightarrow \exists x \left( f^n(x) = y \wedge \bigwedge_{i=1}^k \varphi_i(x, \bar{y}_i) \right) \\ &\Leftrightarrow D^n(y) \wedge \bigwedge_{i=1}^k \varphi_i(f^{-n}(y), \bar{y}_i) \end{aligned}$$

Формула  $\varphi_i$  — атомарная, то есть лежит в  $\Sigma_{qe}^f$ .

А)  $\varphi_i = f^{-n}(y) < u$ . Тогда меняем  $f^{-n}(y) < u$  на  $y < f^n(u)$ .

Б)  $\varphi_i = f^{-n}(y) > u$ . Тогда меняем  $f^{-n}(y) > u$  на  $y > f^n(u)$ .

В)  $\varphi_i = U(f^{-n}(y))$ . Тогда меняем  $U(f^{-n}(y))$  на  $U^n(y)$ .

Поэтому мы теперь можем считать, что формул вида 1 у нас нет.

Без ограничения общности можно считать, что у нас имеется не более одной формулы вида 3. Действительно, предположим, что имеется хотя бы две формулы вида 3. Заметим, что  $(x < u \wedge x < v)$  эквивалентна  $(x < u \wedge u < v) \vee (x < v \wedge u \not< v)$ . Точно так же, как мы делали выше, можно будет прийти к дизъюнкции формул с кванторами существования и рассматривать каждый дизъюнкт отдельно. Аналогичным образом можно считать, что у нас имеется не более одной формулы вида 5.

Предположим, что у нас нет формулы вида (8). Тогда нам остается рассмотреть формулы трёх видов, а именно:

$$\begin{aligned} & \exists x \left( x < f^m(y) \wedge \bigwedge_i U_i(x) \right), \exists x \left( x > f^m(y) \wedge \bigwedge_i U_i(x) \right), \\ & \exists x \left( f^j(z) < x < f^m(y) \wedge \bigwedge_i U_i(x) \right) \end{aligned}$$

где  $U_i$  — одноместные предикаты.

Очевидно, что в силу определения имеет место

$$\begin{aligned} \exists x(x < f^m(y)) & \Leftrightarrow \exists x(f^{-m}(x) < y) \Leftrightarrow \exists z \exists x(f^{-m}(x) = z \wedge z < y) \Leftrightarrow \\ & \Leftrightarrow \exists z(\exists x(f^{-m}(x) = z) \wedge z < y) \Leftrightarrow \exists z(R^{-m}(z) \wedge z < y) \end{aligned}$$

Аналогично получаем, что

$$\exists x(x > f^m(y)) \Leftrightarrow \exists x(f^{-m}(x) > y) \Leftrightarrow \exists z(R^{-m}(z) \wedge z > y)$$

Однако формулы, написанные справа, являются формулами сигнатуры  $\Sigma_{qe}^+$ , а по предположению теория  $T_{qe}^+$  допускает сокращение кванторов, следовательно, эти формулы эквивалентны бескванторным.

Рассмотрим формулу последнего, третьего вида:

$$\begin{aligned} \exists x(f^j(z) < x < f^m(y)) & \Leftrightarrow \exists x(z < f^{-j}(x) \wedge f^{-m}(x) < y) \Leftrightarrow \\ & \Leftrightarrow \exists x \exists u \exists v(u = f^{-j}(x) \wedge z < u \wedge v = f^{-m}(x) \wedge v < u) \Leftrightarrow \\ & \Leftrightarrow \exists u \exists v(\exists x(u = f^{-j}(x) \wedge v = f^{-m}(x)) \wedge z < u \wedge v < u) \Leftrightarrow \end{aligned}$$

Пусть  $j \geq 0, m \geq 0$ . Тогда

$$\begin{aligned} \exists x(f^j(z) < x < f^m(y)) & \Leftrightarrow \exists x \exists u \exists v(f^j(z) = u \wedge f^m(y) = v \wedge u < x < v) \Leftrightarrow \\ & \Leftrightarrow \exists u \exists v(f^j(z) = u \wedge f^m(y) = v \wedge \exists x(u < x < v)) \end{aligned}$$

Заметим, что  $\exists x(u < x < v)$  — формула сигнатуры  $\Sigma_{qe}^+$ , следовательно эквивалентна бескванторной формуле, скажем,  $\Xi(u, v)$ . Тогда

$$\begin{aligned} \exists u \exists v(f^j(z) = u \wedge f^m(y) = v \wedge \exists x(u < x < v)) & \Leftrightarrow \exists u \exists v(f^j(z) = u \wedge f^m(y) = v \wedge \Xi(u, v)) \Leftrightarrow \Xi(f^j(z), f^m(y) = v) \end{aligned}$$

Итого, если среди в формуле  $\exists x(\bigwedge_{i=1}^k \varphi_i(x, \bar{y}_i))$  среди  $\varphi_i$ -тых нет формул вида (8), то эта формула допускает сокращение экзистенциального квантора. Теорема доказана.

**Теорема 2.** Теория  $T_{qe}^f$  упорядоченно суперстабильна.

Доказательство. Пусть  $\mathcal{M} \models T_{qe}^f$ . Рассмотрим произвольное сечение  $(C, D)$  линейно упорядоченной структуры  $\mathcal{M}$ . В доказательстве теоремы об элиминации кванторов мы заметили, что в любой формуле  $f^n(x)$  можно заменить на  $x$ . Грубо говоря, вместо  $\psi(f^n(x), a)$  рассматриваем  $\psi_1(x, f^{-n}(a))$ . Пусть  $b = f^{-n}(a)$ . Получаем формулу  $\psi_1(x, b)$ . Таким образом, тип полностью описывается сечением и формулами сигнатуры  $\Sigma_{qe}^+$ , которая является несущественным обогащением сигнатуры  $\Sigma^+$ . А как известно (см. Факт 1), любая теория сигнатуры  $\Sigma^+$ , где  $<$  — это линейный порядок, является упорядоченно суперстабильной. Следовательно, сечение  $(C, D)$  имеет самое большее  $2^{|\Sigma^+|}$  расширений, то есть,  $T_{qe}^f$  является упорядоченно суперстабильной.

**Вопрос.** Если элементарная теория обогащение чистого линейного порядка одноместной функцией является упорядоченно стабильной, будет ли она упорядоченно суперстабильной.

#### Список литературы

- 1) Б.С. Байжанов, В.В. Вербовский. Упорядоченно стабильные теории. Алгебра и логика, 2011, том 50, номер 3, С. 303–325.
- 2) V. V. Verbovskiy. O-Stable Ordered Groups. Siberian Advances in Mathematics. — 2012. — V. 22, N1. — P. 50–74.

# ВЫЧИСЛЕНИЯ НАД УПОРЯДОЧЕННЫМИ КОЛЬЦАМИ ПОСРЕДСТВОМ ОБОБЩЁННЫХ РЕГИСТРОВЫХ МАШИН

Латкин<sup>(1)</sup> И.В., Селиверстов<sup>(2)</sup> А.В.

<sup>(1)</sup>Восточно-Казахстанский технический университет им. Д. Серикбаева,

<sup>(2)</sup>Институт проблем передачи информации им. А.А. Харкевича

**Аннотация.** Рассматривается работа обобщённых регистровых машин над частично упорядоченными кольцами со вспомогательными бинарными операциями: над кольцом целых чисел, его бесконечной декартовой степенью и ультрастепенью. Обсуждается осуществимость некоторых алгоритмов и их сложность. Приводится пример нефакториального кольца элементарно эквивалентного кольцу целых чисел. Показывается, что недетерминированные вычисления над целыми числами можно реализовать как детерминированные над декартовой степенью кольца целых чисел. Используя такие машины можно также моделировать вычисления с оракулами. Это обеспечивает алгебраический подход к описанию некоторых классов вычислительной сложности. Однако эта модель вычислений существенно отличается от альтернирующих машин.

**Ключевые слова:** обобщённые регистровые машины, упорядоченные кольца, вычислительная сложность, недетерминированные вычисления, модель вычислений.

**Введение.** Рассматривается работа обобщённых регистровых машин (ОРМ) [1,2,3] над частично упорядоченным ассоциативным и коммутативным кольцом со вспомогательными бинарными операциями вычитания и  $\text{rest} : (R, 0, +, -, \text{rest}, \leq)$ , поскольку без подобного расширения вычислительные возможности таких машин весьма ограничены. В кольце целых чисел  $\mathbb{Z}$  для любого  $x$  и для  $y \geq 2$  значением  $\text{rest}(x, y)$  служит остаток от деления  $x$  на  $y$  из множества  $\{0, \dots, y - 1\}$ , а для  $y \leq 1$  полагаем  $\text{rest}(x, y) = 0$ . Далее также предполагается, что

$$\exists z((x = y \cdot z + \text{rest}(x, y)) \wedge (0 \leq \text{rest}(x, y) < y))$$

для  $y \geq 2$ , где под двойкой понимается сумма нейтрального элемента по умножению с самим собой, а при нарушении условия  $y \geq 2$  (в частности, если элемент  $y$  несравнимый с аналогом двойки) не обязательно будет выполняться равенство  $\text{rest}(x, y) = 0$ . Тем не менее, требуется, чтобы для любого  $y$  было верно  $\text{rest}(x, y) \geq 0$  и либо  $\text{rest}(x, y) < y$ , либо  $\text{rest}(x, y)$  несравним с  $y$ . Полагаем, что  $\text{rest}(x, y) = 0$ , когда элемент  $x$  делится на  $y$ , в частности, если  $y$  обратим. Функция  $\text{rest}$  может быть корректно определена и для неевклидовых колец, как мы это увидим далее. В случае совпадения кольца  $R$  с кольцом  $\mathbb{Z}$  получается консервативное расширение теории  $Th(\mathbb{Z})$ , так как отношение порядка и деление с остатком определяются над  $\mathbb{Z}$  формулами первого порядка.

**Методы:** применяются обычные методы алгебры и математической логики.

**Основная часть.** Напомним вкратце описание работы ОРМ над алгебраической системой с основным множеством  $A$  и заданными на нём операциями  $f_i$  местности  $k(i)$ , предикатами  $T_j$  местности  $l(j)$  и выделенными элементами  $c_t$ . Машина имеет бесконечное множество (рабочих) регистров  $R_j$ , содержащих элементы из  $A$ , и бесконечно много индексных регистров  $I(n)$ , содержащих натуральные числа. Константы соответствуют операциям записи соответствующего элемента в регистр. Программы представляют собой конечные списки команд, часть из которых может быть помечена (или пронумерована). Выполняя эти команды, машина может за один шаг скопировать элемент из регистра  $R_{I(s)}$ , индексированного содержимым индексного регистра  $I(s)$ , и переслать его в регистр  $R_{I(t)}$ :  $R_{I(t)} := R_{I(s)}$ . Также она может применить любую сигнатурную операцию  $f_j$  к элементам, из регистров  $R_{I(n_1)}, \dots, R_{I(n_k(j))}$ , и записать результат в  $R_{I(m)}$ . При проверке на истинность сигнатурного предиката  $T_j$ , применённого к элементам из  $R_{I(n_1)}, \dots, R_{I(n_l(j))}$ , машина переходит в новое состояние в зависимости от его истинности, т.е. в случае истинности

предиката  $T_j$  на наборе элементов из регистров  $R_{I(n_1)}, \dots, R_{I(n_{l(j)})}$ , машина переходит к выполнению, например, команды с меткой  $k$ , а в противном случае исполняет команду  $t$ .

Над индексными регистрами выполняются обычные операции регистровых машин. В начале работы в нулевом индексном регистре записано число регистров, занятых входными данными, а в остальных индексных регистрах записаны нули. Незанятые входными данными регистры содержат некоторый фиксированный элемент основного множества  $A$ , для частично упорядоченного кольца  $R$  там естественно записать нули.

Время работы машины полиномиальное, если существует такой многочлен  $p(n)$ , что если вначале ровно  $n$  регистров занято входными данными, то полное число шагов, выполняемых машиной до остановки, ограничено значением многочлена  $p(n)$ . Задача разрешима за полиномиальное время, если имеется ОРМ, решающая эту задачу за полиномиальное время. Это определение сложности вычисления ОРМ естественно в следующем смысле: набор значений аргументов  $x_1, \dots, x_n$  (элементов основного множества системы  $A$ ) отождествляется со словом  $x_1 \dots x_n$  в алфавите  $A$ , таким образом,  $n$  — это просто длина входной цепочки, которая распределена по  $n$  входным регистрам. Вычислительная сложность на рассматриваемых машинах не учитывает сложность выполнения отдельных арифметических операций, которые могут быть невычислимыми в обычном смысле. В частности, кольцо  $R$  может не быть счётным. Однако здесь учитывается время на операции над индексными регистрами.

Вслед за [1] мы не допускаем команд вида  $R_{I(t)} := a$ , где  $a$  — отличный от сигнатурной константы элемент основного множества системы. Поэтому если даже кольцо  $\mathbb{Z}$  вкладывается в  $R$  и  $\mathbb{Z}R$  — его образ, то при некоторых условиях на  $R$  и на входные данные  $x_1, \dots, x_n$ , в регистрах машины нельзя получить некоторые (или даже никакие) элементы из  $\mathbb{Z}R$  отличные от констант, при условии, что они не были заданы изначально. Пример такого сорта кольца и элементов  $x_1, \dots, x_n$  возникает при доказательстве теоремы 2.

В то же время, в индексных регистрах могут вычисляться любые рекурсивные (вычислимые) функции от натуральных чисел, так как имеются функции, позволяющие вычислять результат сложения, вычитания и умножения содержимого любого индексного регистра  $I(k)$  с натуральным числом  $a$ , а также функция, вычисляющая целую часть от деления на  $a$  числа из  $I(k)$ . Значит, можно находить значения и любых вычислимых функций над кольцом  $\mathbb{Z}$ , представляя целые числа в виде формальной разности двух натуральных. Для этого нужно расширить список команд обобщённых регистровых машин командами вида  $\text{HALT}(I(k))$ , которые означают, что результат вычислений содержится в индексном регистре  $I(k)$ , и у нас имеется возможность его прочитать. При этом всё равно остаётся проблема определения, представляет ли элемент, содержащийся в данном регистре, аналог натурального числа из некоторого индексного регистра, если кольцо  $\mathbb{Z}$  вкладывается в  $R$ , но его подмножеством не является. В лемме 1 мы увидим, что если в качестве дополнительного входа машины разрешить задавать константу 1 или она присутствует в сигнатуре, то эта проблема разрешима. В этом случае ОРМ могут вычислять любые рекурсивные (вычислимые в обычном смысле) функции внутри  $\mathbb{Z}R$ .

Рассмотрим вычисления над ультрастепенью кольца  $\mathbb{Z}$ . Пусть  $\omega$  — множество натуральных чисел, начиная с нуля, а  $D$  — некоторый его ультрафильтр, расширяющий фильтр коконечных подмножеств,  $U$  — ультрастепень линейно упорядоченного кольца  $\mathbb{Z}$  над ультрафильтром  $D$ . Это линейно упорядоченное кольцо — область целостности, и в нём корректно определен наибольший общий делитель, поскольку  $\mathbb{Z}$  и его ультрастепень  $U$  элементарно эквивалентны [4]. Но  $U$  обладает необычными свойствами, невыразимыми в языке первого порядка теории частично упорядоченных колец.

Элементы кольца  $U$  — классы эквивалентности бесконечных последовательностей целых чисел  $\mathbf{a} = (a_0, a_1, \dots)$ . Две последовательности эквивалентны, если они совпадают на множестве индексов, принадлежащем ультрафильтру  $D$ . В частности, эквивалентны

любые две последовательности, отличающиеся лишь в конечном числе позиций. Операции и отношение порядка в кольце  $U$  определяются покомпонентно. Кольцо  $\mathbb{Z}$  вложено в  $U$ , числу  $a$  соответствует класс постоянной последовательности  $\mathbf{a} = (a, a, \dots)$ . Ввиду элементарной эквивалентности колец  $\mathbb{Z}$  и  $U$ , кольцо  $U$  имеет ровно два обратимых элемента, а именно, классы последовательностей  $\mathbf{1} = (1, 1, \dots)$  и  $-\mathbf{1} = (-1, -1, \dots)$ .

**Теорема 1.** *Область целостности  $U$  не является факториальным кольцом, хотя в нём существует наибольший общий делитель любых двух ненулевых элементов  $\mathbf{a}$  и  $\mathbf{b}$ , а также неполное частное от деления элемента  $\mathbf{a}$  на  $\mathbf{b} \geq \mathbf{1}$ , то есть такой элемент  $\mathbf{q}$ , что  $\mathbf{a} = \mathbf{b} \cdot \mathbf{q} + \text{rest}(\mathbf{a}, \mathbf{b})$ .*

**Следствие 1.** *Упорядоченное кольцо  $U$  не является ни архимедовым, ни плотным, ни евклидовым.*

С другой стороны, в области целостности  $U$  операция  $\text{rest}(\cdot, \cdot)$  ведёт себя во многом одинаково с операцией вычисления остатка в целых числах.

**Теорема 2.** *Наибольший общий делитель и неполное частное от деления одного элемента на другой не вычислимы над кольцом  $(U, 0, +, -, \text{rest}, \leq)$  посредством ОРМ.*

Картина меняется, когда имеется возможность использовать константу 1. Пусть в области целостности  $R$  с нестрогим линейным порядком определена вычислимость так, что имеются алгоритмы для вычисления сложения, вычитания и умножения, вычислимыми являются константа 0 и отношение порядка. В утверждениях 1 и 2 предполагается выполнение этих свойств в кольце  $R$ . Модель вычислимости здесь никак себя не проявляет, будь то вычислимость, задаваемая некоторым абстрактным устройством, вроде ОРМ, или вычислимость, заданная подходящей нумерацией.

**Утверждение 1.** *Из наличия алгоритма для нахождения неполного частного любых двух элементов  $a$  и  $b \neq 0$  следует существование алгоритма для вычисления остатка от деления всякого элемента  $a$  на любой элемент  $b \geq 1$ .*

**Утверждение 2.** *Наоборот, наличие алгоритма для вычисления функции  $\text{rest}$  и возможность вычислять элемент 1 (или наличие его в сигнатуре) даёт алгоритмы для выяснения обратимости любых ненулевых элементов кольца и нахождения неполного частного во многих случаях.*

**Теорема 3.** *Вычисление неполного частного от деления любого элемента  $\mathbf{a}$  на элемент  $\mathbf{b} > \mathbf{0}$  производится подходящей ОРМ над кольцом  $(U, 0, +, -, \text{rest}, \leq)$  за время, ограниченное константой, если на вход машины подавать не только эти элементы, но также и запись элемента  $\mathbf{1}$  в регистре.*

Перейдём к вычислениям над декартовой степенью  $\mathbb{Z}^\omega$  кольца целых чисел, с покомпонентным определением сигнатурных операций и отношения порядка. Это кольцо имеет мощность континуума. отождествим кольцо  $\mathbb{Z}$  с образом диагонального вложения в  $\mathbb{Z}^\omega$ , когда целое число отождествляется с постоянной последовательностью. Кроме наличия делителей нуля и несравнимых элементов, у  $\mathbb{Z}^\omega$  имеются и другие существенные отличия от кольца  $U$ . Например, в кольце  $\mathbb{Z}^\omega$  для между элементами  $\mathbf{c} = (c_0, c_1, \dots)$  и  $\mathbf{c} + \mathbf{1} = (c_0 + 1, c_1 + 1, \dots)$  имеется бесконечно много попарно несравнимых друг с другом элементов. Однако, порядок в  $\mathbb{Z}^\omega$  тоже неплотный, так как между двумя элементами, у которых проекции на все множители, кроме одного, одинаковые, а особая координата второго элемента на единицу больше соответствующей проекции у первого, ничего нет. Здесь наблюдается эффект, отмеченный во введении: остаток от деления на элемент, у которого проекции на собственную часть множителей — минус единицы, а остальные проекции — положительные, может быть несравнимым с делителем.

Далее, наряду с элементом  $\mathbf{1}$ , мы будем использовать элемент  $\mathbf{d} = (0, 1, 2, \dots) \in \mathbb{Z}^\omega$ , у которого проекция на  $-i$ -й декартов множитель равна  $k$  и некоторые другие.

**Лемма 1.** *Пусть в одном из регистров ОРМ имеется элемент  $\mathbf{1}$ . Тогда по имеющейся записи элемента  $\mathbf{k} = (k, k, \dots) \in \mathbb{Z}^\omega$  можно найти запись в индексных регистрах представления числа  $k$  в двоичной системе счисления за сублинейное время от величины  $k$ . Наоборот, если в индексных регистрах имеется запись представления целого*

числа  $k$  в двоичной системе счисления или в одном из индексных регистрах записано само это число, то можно за линейное время от числа индексных регистров, содержащих цифры числа  $k$  или от величины  $k$ , соответственно, вычислить элемент  $\mathbf{k} \in \mathbb{Z}^\omega$ .

**Пример.** Рассмотрим критерий простоты числа, который основан на малой теореме Ферма: целое число  $p > 2$  простое тогда и только тогда, когда для каждого  $x \in \mathbb{Z}$  выполнено равенство  $x^p \equiv x \pmod{p}$ . Этот критерий лежит в основе вероятностного теста Рабина–Миллера для проверки простоты натурального числа в рамках обычной тьюринговой вычислимости, когда для достаточно большого количества натуральных чисел  $x$  проверяется сравнение  $x^p \equiv x \pmod{p}$ . Однако использование ОРМ над кольцом  $\mathbb{Z}^\omega$ , позволяет создать уже детерминированный тест для такой проверки.

Вместо перебора чисел  $x$  из  $\mathbb{Z}$  можно запустить ОРМ над  $\mathbb{Z}^\omega$  на независимых от входа  $\mathbf{p}$  последовательностях  $\mathbf{d}$  и  $\mathbf{1}$ . Целое число  $p > 2$  простое тогда и только тогда, когда  $\mathbb{Z}^\omega \models \mathbf{d}^p \equiv \mathbf{d} \pmod{\mathbf{p}}$ . Проверка этого условия завершается за конечное число шагов над  $\mathbb{Z}^\omega$ : остаток от деления на  $\mathbf{p} \in \mathbb{Z}^\omega$  вычисляется за один шаг, посредством функции  $\text{rest}(\cdot, \cdot)$ ; возведение в степень  $p \in \omega$  требует  $O(\log p)$  умножений, если нам известно это натуральное число. Но поскольку нам дано лишь  $\mathbf{p} \in \mathbb{Z}^\omega$ , то предварительно мы ищем число  $p$ , используя элемент  $\mathbf{1}$ , встроенные в машину операции и часть индексных регистров для хранения цифр в двоичном представлении числа  $p$ , опираясь на лемму 1. На это тратится тоже  $O(\log p)$  действий. При этом на вход подаётся только три элемента  $\mathbf{p}$ ,  $\mathbf{d}$  и  $\mathbf{1}$ . А число шагов зависит от значения числа  $p$  и может быть сколь угодно большим. Поэтому работа ОРМ не завершается за полиномиальное время относительно количества входных регистров, которых всего только три. Но время работы машины — линейное по отношению к величине числа  $p$ .

Напомним, что множество  $X$  из класса  $\mathbf{NP}$  называется  $\mathbf{NP}$ -полным, если каждое множество из класса  $\mathbf{NP}$  сводится по Карпу к  $X$ . Примером служит множество  $X_0$  таких линейных диофантовых уравнений от многих переменных, что каждое из этих уравнений имеет некоторое  $(0,1)$ -решение [5,6]. Коэффициентами уравнений служат обычные целые числа. Эту задачу можно интерпретировать и следующим образом. Можно ли среди нескольких целых чисел, которые задаются в качестве коэффициентов диофантова уравнения, выбрать такие, что их сумма равна данному числу — противоположному к свободному члену уравнения? Поэтому для краткости, мы будем называть задачу распознавания множества  $X_0$  задачей о сумме подмножества мультимножества (среди коэффициентов диофантова уравнения могут быть равные).

**Теорема 4.** *Задача о сумме подмножества над  $\mathbb{Z}$  детерминированно разрешима за полиномиальное время посредством ОРМ над  $\mathbb{Z}^\omega$ , использующей элементы  $\mathbf{d}$  и  $\mathbf{1}$ .*

Если позволить использовать не только элемент  $\mathbf{d}$ , проекции которого легко вычислимы, но и произвольные наперёд заданные элементы, то можно реализовать вычисление с оракулом.

**Теорема 5.** *Задача распознавания целых чисел, принадлежащих фиксированному непустому множеству  $Y \subset \mathbb{Z}$  разрешима за конечное время на ОРМ над  $\mathbb{Z}^\omega$ , использующей элемент  $\mathbf{1}$  и элемент  $\mathbf{f}$ , определяемый множеством  $Y$ .*

**Обсуждение.** Недетерминированное вычисление над  $\mathbb{Z}$  превращается в параллельное вычисление на неограниченном числе копий кольца  $\mathbb{Z}$ , которыми служат проекции декартовой степени на множители. Выигрыш достигается, если позволить машине использовать внутренние параметры из  $\mathbb{Z}^\omega$ , которые не принадлежат кольцу  $\mathbb{Z}$ , как дополнительные входы. Такая модель соответствует многопроцессорному вычислительному устройству с ограниченным обменом данными между процессорами, что существенно отличает эту модель от альтернирующих машин.

Принятая для ОРМ оценка вычислительной сложности оказывается неудобной, когда на вход подаётся одно число. При работе с многочленами, рациональными функциями или матрицами эта оценка лучше соответствует обычному понятию сложности. Но в общем случае полиномиально ограниченное число арифметических



операций нельзя выполнить за полиномиальное время на обычных машинах Тьюринга из-за возникновения неожиданно больших чисел. Рассмотрим, например, вычисление наибольшего общего делителя (в кольце  $\mathbb{Q}[x]$ ) двух многочленов с целыми коэффициентами от одной переменной на обобщённой регистровой машине над кольцом  $\mathbb{Z}$ . Пусть каждый многочлен задан набором коэффициентов, включая нулевые. Тогда запись одного многочлена степени  $d$  занимает  $d + 1$  регистров. Алгоритм Евклида требует линейного от суммы степеней числа операций. Однако возникающие на промежуточных шагах коэффициенты могут иметь очень большую длину записи [7,8,9, 10], что значительно увеличивает время вычислений при использовании многоленточных машин Тьюринга.

**Заключение.** Обобщённые регистровые машины – это мощное средство для изучения сложности вычислений над произвольными алгебраическими структурами, в первую очередь над кольцами и полями. Вычисления ОРМ над полем вещественных чисел подобны вычислениям на BSS-машине [11,12], а в случае линейно упорядоченных ассоциативных и коммутативных колец почти не отличаются от машин над списочной надстройкой Ашаева–Беляева–Мясникова [13] и S-машин Хеммерлинга [14].

### Литература

1. E. Neumann, P. Pauly, A topological view on algebraic computation models, *Journal of Complexity*, 44 (2018), 1–22.
2. A.V. Seliverstov, Heuristic algorithms for recognition of some cubic hypersurfaces, *Programming and Computer Software*, 47 (2021), 50–55.
3. A.V. Seliverstov, Binary solutions to large systems of linear equations, *Prikladnaya Diskretnaya Matematika*, no. 52 (2021), 5–15.
4. C.C. Chang, H.J. Keisler, *Model Theory*, Elsevier, 1990.
5. K. Koiliaris, C. Xu, Faster pseudopolynomial time algorithms for subset sum, *ACM Transactions on Algorithms*, 15:3 (2019), 40.
6. A.V. Seliverstov, On binary solutions to systems of equations, *Prikladnaya Diskretnaya Matematika*, no. 45 (2019), 26–32.
7. P.E. Alaev, V.L. Selivanov, Fields of algebraic numbers computable in polynomial time. I, *Algebra and Logic*, 58:6 (2020), 447–469.
8. A. Sinhababu, T. Thierauf, Factorization of polynomials given by arithmetic branching programs, *Computational complexity*, 30:15 (2021), 1–47.
9. W. Habicht, Eine Verallgemeinerung des Sturmschen Wurzelzählverfahrens, *Commentarii Mathematici Helvetici*, 21 (1948), 99–116.
10. A.G. Akritas, *Elements of Computer Algebra with Applications*, John Wiley and Sons, NY, 1989.
11. L. Blum, M. Shub, S. Smale, On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines, *Bulletin of the American Mathematical Society*, 21:1 (1989), 1–46.
12. L. Blum, F. Cucker, M. Shub, S. Smale, *Complexity and Real Computation*, Springer, New York, 1998.
13. I.V. Ashaev, V.Ya. Belyaev, A.G. Myasnikov, Toward a Generalized Computability Theory, *Algebra and Logic*, 32:4 (1993), 185–205.
14. A. Hemmerling, Computability of string functions over algebraic structures, *Mathematical Logic Quarterly* 44:1 (1998), 1–44.

# EXAMPLE OF A NON-LOCALLY FINITE NOVIKOV COALGEBRA

*Daniyar Kozybaev*

*L.N. Gumilyov Eurasian National University,*

*Faculty of Mechanics and Mathematics*

ORCID ID: <https://orcid.org/0000-0001-5509-8656>

*kozybayev@gmail.com*

## Abstract

Coalgebras have been considered by mathematicians for a long time as part of the structural theory of Hopf algebras [3]. At present, coalgebras have been actively studied in connection with the study of quantum groups. V. Drinfeld [4] introduced the notion of a Lie bialgebra, in which comultiplication determines the structure of a Lie coalgebra. The concept of Jordan and alternative coalgebras was defined in [2].

One of the main questions in the theory of coalgebras is the question of the local finiteness of a given variety of coalgebras. In [3], the local finiteness of associative coalgebras was proved. A similar result for Jordan and alternative coalgebras was established in [2]. An example of a non-locally finite Lie coalgebra was constructed by V. Michaelis [1].

In this paper, we introduce the concept of a left-symmetric coalgebra and Novikov coalgebra, and describe the identities of left-symmetric coalgebras and Novikov coalgebras. It is proved that the variety of Novikov coalgebras is not locally finite. For this, an example of a non-locally finite Novikov coalgebra was constructed. The results obtained can be used in information security algorithms and in elliptic cryptography.

**Definition.** A vector space  $A$  over a field  $F$  in which a linear mapping is given is  $\Delta : A \rightarrow A \otimes A$  called a coalgebra.

Let's put

$$\Delta(a) = \sum_a a_{(1)} \otimes a_{(2)},$$

where  $a \in A$  (Svidler's notation).

Let  $A^* = \{f : A \rightarrow F \text{ linear functions}\}$  be the space of functionals dual to  $A$ . Then we define pairing

$$\langle A^*, A \rangle \rightarrow F,$$

assuming  $\langle f, a \rangle = f(a)$ , where  $f \in A^*, a \in A$ .

We define the operation of multiplication on space  $A^*$  as follows

$$\langle f \cdot g, a \rangle = \langle f \otimes g, \Delta(a) \rangle = \sum_a f(a_{(1)}) \cdot g(a_{(2)}),$$

where  $f, g \in A^*, a \in A$ . It is easy to check what  $A^*$  is an algebra with respect to this multiplication.

We define linear mappings  $\tau : V \otimes W \rightarrow W \otimes V$ , where are  $V, W$  vector spaces, assuming

$$\tau(v \otimes w) = w \otimes v, v \in V, w \in V.$$

**Theorem 1.** A coalgebra  $\Delta : A \rightarrow A \otimes A$  is left-symmetric if and only if the identity

$$(1 - \tau \otimes 1)((\Delta \otimes 1 - 1 \otimes \Delta)\Delta) = 0.$$

**Theorem 2.** For a left-symmetric coalgebra  $\Delta : A \rightarrow A \otimes A$  to be a Novikov coalgebra it is necessary and sufficient that the following equalities hold:

$$((1-1 \otimes \tau)(\Delta \otimes 1)\Delta) = 0. \text{ (five)}$$

**Theorem 3.** The variety of left- symmetric coalgebras is not locally finite.

**Theorem 4.** The variety of Novikov coalgebras is not locally finite.

An example of a non-locally finite Novikov coalgebra

Let  $A$  be a vector space with basis  $e, f_1, f_2, \dots, f_n, \dots$ . Define a coalgebra  $\Delta : A \rightarrow A \otimes A$

$$\begin{aligned}\Delta(e) &= 0, \\ \Delta(f_{3n-2}) &= e \otimes f_{3n+1}, \\ \Delta(f_{3n-1}) &= e \otimes f_{3n}, \\ \Delta(f_{3n}) &= e \otimes f_{3n+2}.\end{aligned}$$

First, we check the left symmetry identity :

$$\begin{aligned}(1-\tau \otimes 1)(\Delta \otimes 1-1 \otimes \Delta)\Delta(e) &= 0, \\ (1-\tau \otimes 1)(\Delta \otimes 1-1 \otimes \Delta)\Delta(f_{3n-2}) &= (1-\tau \otimes 1)(\Delta \otimes 1-1 \otimes \Delta)(e \otimes f_{3n+1}) = (1-\tau \otimes 1)(-e \otimes e \otimes f_{3n+4}) = \\ &= -e \otimes e \otimes f_{3n+4} + e \otimes e \otimes f_{3n+4} = 0, \\ (1-\tau \otimes 1)(\Delta \otimes 1-1 \otimes \Delta)\Delta(f_{3n-1}) &= (1-\tau \otimes 1)(\Delta \otimes 1-1 \otimes \Delta)(e \otimes f_{3n}) = (1-\tau \otimes 1)(-e \otimes e \otimes f_{3n+2}) = \\ &= -e \otimes e \otimes f_{3n+2} + e \otimes e \otimes f_{3n+2} = 0, \\ (1-\tau \otimes 1)(\Delta \otimes 1-1 \otimes \Delta)\Delta(f_{3n}) &= (1-\tau \otimes 1)(\Delta \otimes 1-1 \otimes \Delta)(e \otimes f_{3n+2}) = (1-\tau \otimes 1)(-e \otimes e \otimes f_{3n+3}) = \\ &= -e \otimes e \otimes f_{3n+3} + e \otimes e \otimes f_{3n+3} = 0.\end{aligned}$$

This means that the given coalgebra is left-symmetric. Now let's check the Novikov identities :

$$\begin{aligned}(1-1 \otimes \tau)(\Delta \otimes 1)\Delta(e) &= 0, \\ (1-1 \otimes \tau)(\Delta \otimes 1)\Delta(f_{3n-2}) &= (1-1 \otimes \tau)(\Delta \otimes 1)(e \otimes f_{3n+1}) = 0, \\ (1-1 \otimes \tau)(\Delta \otimes 1)\Delta(f_{3n-1}) &= (1-1 \otimes \tau)(\Delta \otimes 1)(e \otimes f_{3n}) = 0, \\ (1-1 \otimes \tau)(\Delta \otimes 1)\Delta(f_{3n}) &= (1-1 \otimes \tau)(\Delta \otimes 1)(e \otimes f_{3n+2}) = 0.\end{aligned}$$

According to Theorem 2, a coalgebra is a Novikov coalgebra. Consider the subcoalgebra  $K$  generated by elements  $e, f_2$ . We have  $\Delta(f_2) = e \otimes f_3 \in K \otimes K$ , i.e.  $f_3 \in K$ ,  $\Delta(f_3) = e \otimes f_5$ , and  $f_5 \in K$ , i.e. continuing this process, we obtain that  $e, f_2, f_3, \dots, f_{3n-1}, f_{3n}, \dots \in K$ . Subcoalgebra  $K$  infinite-dimensional. Consequently, the variety of Novikov coalgebras is not locally finite.

**Keywords:** left-symmetric coalgebra, Novikov coalgebra, locally finite, identity.

## References

1. M. Sweedler. Hopf algebras// W.A. Benjamin Inc., New York. - 1969.
2. V. G. Drinfeld. Quantum Group// Proc. Int. Congress Math., Berkeley. - 1986.
3. J. Anquella, T. Cortes, F. Montaner. Nonassociative Coalgebras// Comm. in Algebra. - 1994. - Vol. 22, N12. - P. 4693-4716.
4. W. Michaelis. Lie Coalgebras// Adv. Math. - 1980. - Vol. 38. - P. 1-54.

## ФОРМУЛЬНО-ОПРЕДЕЛИМЫЕ КЛАССЫ АЛГЕБРАИЧЕСКИХ СИСТЕМ

Бекенов М., Касатова А., Нуракунов А.

<sup>1</sup>Евразийский Национальный университет имени Л.Н.Гумилева, Нур-Султан,  
Казахстан

E-mail: *bekenov50@mail.ru*, <sup>2</sup>НАО Медицинский университет Караганды

<sup>3</sup>Институт математики НАН КР, Бишкек, Кыргызстан

Рассматриваются формульно-определимые классы моделей (т. е. алгебраических систем) и элементарные теории счетного языка первого порядка этих моделей относительно произведений этих теорий, а также формульно-определимые квазимногообразия. Доказаны некоторые свойства, которым они удовлетворяют. В частности, если определяющая теория формульно-определимого класса  $K$  алгебраических систем будет универсальной (позитивной универсальной), то класс  $K$  будет квазимногообразием (многообразием).

Пусть  $A$  — модель сигнатуры  $\Omega$ ,  $\text{Th}(A)$  — полная (элементарная) теория модели  $A$ . В дальнейшем под теорией понимается полная теория.

На множестве  $T(\Omega)$ -множестве всех полных теорий счетной сигнатуры  $\Omega$  рассмотрим бинарную операцию  $\bullet$  по правилу  $T \bullet S = \text{Th}(\{A \times B \mid A \models T \text{ и } B \models S\})$ , для любых полные теории  $T, S \in T(\Omega)$ . Ясно, что  $\langle T(\Omega); \bullet \rangle$  — коммутативная полугруппа с единицей, которую мы называем полугруппой полных теорий.

Эту операцию можно найти, в частности, в [16 J. Wierzejewski, On stability and products, *Fundamenta Mathematicae*, 93(1976)], где показано, что множество всех стабильных (суперстабильных,  $\omega$ -стабильных) теорий образует подполугруппу полугруппы  $\langle T(\Omega); \bullet \rangle$ . А именно, доказано, что произведение двух стабильных (суперстабильных,  $\omega$ -стабильных) теорий является стабильной (суперстабильной,  $\omega$ -стабильной) теорией.

Пусть  $T$  некоторая теория, а  $M = \{S \mid S \bullet T = T, S \in T(\Omega)\}$ . Понятно, что  $\langle M; \bullet \rangle$  - коммутативная подполугруппа с единицей, полугруппы полных теорий.

**Определение.** Класс моделей  $K$  называется формульно-определимым классом моделей, если существует модель  $A$  такая, что  $K = \{B \mid \text{Th}(B) \bullet \text{Th}(A) = \text{Th}(A)\}$ . Теория  $\text{Th}(A)$  называется теорией, определяющей класс моделей  $K$ . Если квазимногообразие является формульно-определимым классом моделей, то такое квазимногообразие называется формульно-определимым квазимногообразием.

Каждой модели  $A$  соответствует элементарный класс (в общем случае неизоморфных) моделей  $K = \{B \mid \text{Th}(B) = \text{Th}(A)\}$ .

Каждой модели  $A$  соответствует формульно-определимый класс

$$K = \{B \mid \text{Th}(B) \bullet \text{Th}(A) = \text{Th}(A)\}.$$

Какими свойствами обладают формульно-определимые классы моделей?

**Теорема 1.** Формульно-определимый класс моделей - аксиоматизируемый класс моделей и замкнут относительно бесконечного произведения моделей.

Аксиоматизируемые классы структур, замкнутые относительно прямых произведений, изучались многими авторами (см. [8]). Проблема характеризации аксиоматизируемых классов структур, замкнутых относительно прямых произведений, остается открытой. Исследования, ссылки и закрытые проблемы, связанные с этой темой, хорошо представлены в книжной энциклопедии «Теория моделей» Р. Ходжеса [8].

Например, класс всех моделей, теории которых  $\omega$ -стабильны, не является формульно-определимым классом. То же самое верно как для классов суперстабильных моделей, так и для класса стабильных моделей. В [16] показано, что множество нестабильных теорий не является полугруппой. Это означает, что класс нестабильных теорий не формульно-определим.

Есть примеры аксиоматизируемого класса, который является формульно-определимым классом моделей, и аксиоматизируемого класса моделей, который не является формульно-определимым классом моделей.

Для квазимногообразий найдены примеры формульно-определимых квазимногообразий и не формульно-определимых квазимногообразий.

**Теорема 2.** Многообразие - формульно-определимый класс моделей.

**Теорема 3.** Класс коммутативных полугрупп, вложимых в группы, является формульно-определимым квазимногообразием, которое не является многообразием.

Теория  $T$  является идемпотентной теорией, если  $T \cdot T = T$ .

**Теорема 4.** Для любого формульно-определимого класса моделей существует идемпотентная теория, определяющая этот класс моделей.

Теорема 5. Формульно-определимый класс моделей является квазимногообразием, если идемпотентная теория, определяющая этот класс моделей, такова, что каждая подмодель ультрастепени модели этой теории лежит в этом классе.

Следствие. Формульно-определимый класс моделей будет квазимногообразием, если он определяется универсальной теорией.

Построены примеры, показывающие, что эти два условия (в теореме и следствии) не эквивалентны.

Также доказано, что множество всех идемпотентных полных теорий образует полную решетку относительно частичного порядка  $\leq$ , определяемого как  $T \leq S$ , тогда и только тогда, когда  $T \cdot S = S$  для любых  $T, S \in T(\Omega)$ .

Рассмотрены также некоторые свойства теорий моделей формульно-определимых классов, когда теория, определяющая этот класс, удовлетворяет некоторым свойствам (категоричности, существованию счетно-насыщенной модели или стабильности).

#### Список использованной литературы

1. Бекенов М.И., Нуракунов А.М. Полугруппы теорий и ее решетка идемпотентных элементов. Алгебра и логика 60, 1-14, 2021
2. G. Birkhoff, Lattice theory, Third Edition, Amer. Math. S., Providence, R.I. (1967).
3. S. Feferman and R. Vaught, The first order properties of algebraic systems. Fundamenta Mathematicae 47, p. 57-103, 1959.
4. T. E. Frayne, A. C. Morel and D. S. Scott, Reduced direct products, Fundamenta Mathematicae, 51 (1962), 195-228.
5. F. Galvin, Horn sentences. Ann. Math. Logic, 1(1970), 389-422.
6. V. A. Gorbunov, Algebraic Theory of Quasivarieties, Plenum Publ. Co., New York (1998).
7. W. Hodges, Model Theory (Encyclopedia of Mathematics and its Applications). Cambridge: Cambridge University Press. (1993).
8. H. J. Keisler, Ultraproducts and elementary classes, Indagationes Mathematicae 23 (1961) 477-495.
9. M. Machover, A note on sentences preserved under direct products and powers, Bull. Acad. Polon. Sci. Math., 8(1960), 519-523.
10. A. Macintyre, Direct powers with distinguished diagonal, Conference in Mathematical Logic London 1970
11. А. И. Мальцев, Алгебраические системы, М. Наука (1970), с.392
12. D. Rees, On semigroups, Proc. Camb. Phil. Soc. 36(1940), 387-400.
13. S. Shelah, Every two elementarily equivalent models have isomorphic ultrapowers, Israel Journal of Mathematics, 10 (1971), 224-233.
14. R. Vaught, On sentences holding in direct products of relational system, Proc. Intern. Congr. of Mathematicians, Amsterdam, (1954), Noordho\_, Groningen, 409.

15.J. M. Weinstein, First order properties preserved by direct product, Ph.D. thesis. Univ. Wisconsin (1965), Madison,

16.J. Wierzejewski, On stability and products, *Fundamenta Mathematicae*, 93(1976), 81-95.

17.М.И.Бекенов, Решетка формульно-определимых подквазимногообразий полных теорий квазимногообразия полных теорий. Мальцевские чтения. Новосибирск. 2018.

# MODEL THEORY OF FINITE AND PSEUDOFINITE GRAPHS

Nurlan D. Markhabatov

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

[nur\\_24.08.93@mail.ru](mailto:nur_24.08.93@mail.ru)

We are dealing with graphs defined by their automorphisms. This work is a continuation of the works [4-6].

*Definition.* [1] An infinite  $\Sigma$ -structure  $M$  is *pseudofinite* if for all  $\Sigma$ -sentences  $\varphi$ ,  $M \models \varphi$  implies that there is a finite  $M_0$  such that  $M_0 \models \varphi$ . The theory  $T = \text{Th}(M)$  of the pseudofinite structure  $M$  is called *pseudofinite*.

*Definition.* [2] A graph is *symmetric (flag-transitive or arc-transitive)* if its automorphism group acts transitively on ordered pairs of adjacent vertices. A graph is *vertex-transitive* if its automorphism group acts transitively on its vertices. A graph is *edge-transitive* if its automorphism group acts transitively on its edges.

A *t-arc* is defined to be a sequence of  $t + 1$  vertices, such that any two consecutive vertices in the sequence are adjacent, and with any repeated vertices being more than 2 steps apart. A *t-transitive* graph is a graph such that the automorphism group acts transitively on  $t$ -arcs, but not on  $(t + 1)$ -arcs.

*Definition.* [3] A *distance-transitive* graph is a graph such that, given any two vertices  $v$  and  $w$  at any distance  $i$ , and any other two vertices  $x$  and  $y$  at the same distance, there is an automorphism of the graph that carries  $v$  to  $x$  and  $w$  to  $y$ .

**Theorem 1.** Any theory  $T$  of a  $t$ -transitive graphs with an infinite model is pseudofinite.

**Corollary 1.** Any theory  $T$  of a vertex-transitive graphs with an infinite model is pseudofinite.

**Corollary 2.** Any theory  $T$  of a edge-transitive graphs with an infinite model is pseudofinite.

**Corollary 3.** Any theory  $T$  of a distance-transitive graphs with an infinite model is pseudofinite.

**Corollary 4.** Any theory  $T$  of a symmetric graphs with an infinite model is pseudofinite.

*This research was partially supported by Committee of Science in Education and Science Ministry of the Republic of Kazakhstan (Grant No. AP08855497, AP08855544).*

## References

1. Ax J. The elementary theory of finite fields // Ann. Math., 1968, vol. 88, no. 2, pp. 239–271.
2. Godsil, C. and Royle, G. "Arc-Transitive Graphs." Ch. 4 in Algebraic Graph Theory. New York: Springer-Verlag, 2001, P. 59-76.
3. Biggs, N. L.; Smith, D. H. (1971), "On trivalent graphs", Bulletin of the London Mathematical Society, 3 (2): 155–158
4. N. D. Markhabatov, S. V. Sudoplatov, "Approximations of Regular Graphs", Herald of the Kazakh-British Technical University, 19:1 (2022), 44–49  
<https://vestnik.kbtu.edu.kz/jour/article/view/452>
5. N. D. Markhabatov, "Approximations of Acyclic Graphs", Bulletin of Irkutsk State University, Series Mathematics, 40 (2022), 104–111  
<https://mathizv.isu.ru/ru/article/file?id=1414>
6. N. D. Markhabatov, "On Smoothly Approximable Acyclic Graphs", Proceedings of the International Scientific Conference «Actual Problems of Mathematics, Mechanics and Informatics» dedicated to the 80th anniversary of professor T.G. Mustafin (8-9 September, Karaganda), Karaganda Buketov University, 2022, 36–37

# КРИТЕРИЙ ВЫЧИСЛИМОСТИ $\langle p_\omega, N \rangle$ – РАЗЛОЖИМОЙ АБЕЛЕВОЙ ГРУППЫ БЕЗ КРУЧЕНИЯ

Н.Г.Хисамиев, Д.А.Тусупов, С.Д.Тыныбекова  
Евразийский национальный университет им. Л.Гумилева  
E-mail: [hisamiev@mail.ru](mailto:hisamiev@mail.ru)

Изучение конструктивных, т.е. вычислимо нумерованных и вычислимых абелевых групп, начато А.И.Мальцевым в статье [1], где он поставил общую задачу: «определить какие конструктивные (т.е. вычисляемые) нумерации допускают те или иные абстрактно заданные группы».

В [2]-[4] введены понятия эффективной и сильной разложимости абелевой группы и получены критерии таких разложимостей и вычислимости групп вида  $A = \bigoplus \{ A(p(i)) \mid i \in \omega \}$ , где  $A(p(i))$  - аддитивная группа рациональных чисел, знаменателями которых являются степенями простого числа  $p(i)$ . В [5] получены критерии вычислимости и разрешимости для такого класса групп.

В данном сообщении введен класс  $\langle p_\omega, N \rangle$  - разложимых абелевых групп, где  $N \in \mathbb{N}$ , и получен критерий вычислимости для таких групп.

Пусть группа

$$A = \bigoplus \{ A_i \mid A_i \leq \langle Q, +, 0 \rangle, i \in \omega \}, \quad (1)$$

где  $\langle Q, +, 0 \rangle$  - аддитивная группа множества всех рациональных чисел.

Для любого элемента  $a \in A \setminus \{0\}$  введем следующие множества простых чисел, положив:

$$P_{<\omega}(a) \Leftrightarrow \{ p \mid A \models \exists n(p) \exists a_{p,n(p)} \left( (p^{n(p)} a_{p,n(p)} = a) \ \& \ \forall x (p^{n(p)+1} x \neq a) \right) \}, \quad (2)$$

$$P_\omega(a) = \{ p \mid A \models \forall n \exists a_{p,n} (p^n a_{p,n} = a) \}. \quad (3)$$

Пусть для любого элемента  $a_i \in A_i \setminus \{0\}$  справедливы следующие условия:

$$\alpha_1) \quad \text{множество} \quad P_{<\omega}(a_i) \quad \text{конечно}; \quad (4)$$

$\alpha_2)$  существует натуральное число  $N$  такое, что справедливы:

$$\alpha_{21}) \text{ мощность множества } P_\omega(a_i) \text{ более } N, \quad (5)$$

$$\alpha_{22}) \text{ для любых } i, j, i \neq j \text{ мощность множества} \quad (6)$$

$$P_\omega(a_i) \cap P_\omega(a_j)$$

не более  $N$ , где  $a_i \in A_i \setminus \{0\}$ ,  $a_j \in A_j \setminus \{0\}$ .

Тогда группу  $A$  назовем  $\langle P_\omega, N \rangle$  - разложимой.

Пусть на множестве  $\omega = \{0, 1, 2, \dots\}$  определен предикат  $D(i, p, n, x)$ ,  $i, n, x \in \omega$ ,  $p \in P$ , где  $P$  – множество всех простых чисел, удовлетворяющее следующим условиям:

$$1. \forall i \forall p \forall n \forall x_0 \forall x_1, \left( (D(i, p, n, x_0) \ \& \ D(i, p, n, x_1)) \rightarrow x_0 = x_1 \right). \quad (7)$$

$$2. \forall i \forall p \forall n \forall m \exists x_{i,p,n} \left( (D(i, p, n, x_{i,p,n}) \ \& \ m < n) \rightarrow \exists x_{i,p,m(i,p)} D(i, p, m, x_{i,p,m(i,p)}) \right). \quad (8)$$

Введем следующие множества, положив:

$$P_\omega(i) \Leftrightarrow \{ p \mid \forall n \exists x_{i,p,n} D(i, p, n, x_{i,p,n}) \}, \quad (9)$$

$$P_{<\omega}(i) \Leftrightarrow \{ p \mid \exists n(i, p) \exists x_-(i, p, n(i, p)) \left( n_-(i, p) \geq 0 \ \& \ D(i, p, n(i, p), x_-(i, p, n(i, p))) \ \& \ \forall x \neg D(i, p, n(i, p) + 1, x) \right) \}. \quad (10)$$

Пусть для этих множеств справедливы следующие условия:

$\alpha_1)$  для любого  $i \in \omega$  множества  $P_{<\omega}(i)$  конечно;

$\alpha_2)$  существует такое число  $N \in \omega$ , что справедливы:

$\alpha_{21})$  для любых  $i, j \in \omega$ ,  $i \neq j$  мощность множества  $P_\omega(i) \cap P_\omega(j)$  не более  $N$ ;

$\alpha_{22})$  для любого числа  $i \in \omega$  мощность множества  $P_\omega(i)$  более числа  $N$ .

Тогда предикат  $D$  назовем  $\langle p_\omega, N \rangle$  - предикатом.

По  $\langle p_\omega, N \rangle$  - предикату  $D(i, p, n, x)$  определим следующие группы, положив:



$$B_{i,p} \simeq qr \left\{ \frac{i}{p^n} \mid \exists x D(i, p, n, x), n, x \in \omega \right\}, +, 0 \}, \quad (11)$$

$$B_i \simeq qr \{ B_{i,p} \mid p \in P \}. \quad (12)$$

ЛЕММА 1. Пусть группа  $A$ , определенная равенством (1),  $\langle P_\omega, N \rangle$  разложима. Тогда для любого  $i$  существует такой элемент  $a_{i,\omega}$ , что для любого простого числа  $p \in P$  справедлива эквивалентность:

$$p \in P_\omega(a_i) \Leftrightarrow \exists x (px = a_{i,\omega}), \quad (13)$$

где множество  $P_\omega(a_i)$  определено равенством (3).

СЛЕДСТВИЕ 1. Пусть для элемента  $a_i \in A_i \setminus \{0\}$  и множества  $P_\omega(a_i)$ , определенное равенством (3), где  $a = a_i$ , справедливо равенство:

$$P_\omega(a_i) = \emptyset.$$

Тогда группа  $A_i$  изоморфна циклической группе, порожденной элементом  $a_{i,\omega}$ , определенное равенством (19).

По  $\langle p_\omega, N \rangle$  - предикату  $D(i, p, n, x)$ , определенный условиями  $\alpha_1), \alpha_2)$ , введем следующие множества и группы, положив:

$$P_\omega(i) = \{ p \mid \forall n \exists x_{i,p,n} D(i, p, n, x_{i,p,n}) \}, \quad (14)$$

$$A_i(\leq N) = qr \left\{ \frac{i}{p^n} \mid p \in P_\omega(i), [P_\omega(i)] \leq N \right\}, \quad (15)$$

$$A_i(> N) = qr \left\{ \frac{i}{p^n} \mid p \in P_\omega(i), [P_\omega(i)] > N \right\}, \quad (16)$$

$$A(\leq N) \simeq \bigoplus \{ A_i(\leq N) \mid i \in \omega \}, \quad (17)$$

$$A(> N) \simeq \bigoplus \{ A_i(> N) \mid i \in \omega \}. \quad (18)$$

ТЕОРЕМА 1. Пусть абелева группа  $A$ , определенная равенством (1),  $\langle p_\omega, N \rangle$  - разложима. Тогда она вычислима, если и только если существует вычислимый  $\langle p_\omega, N \rangle$  - предикат  $D(i, p, n, x)$  такой, что группа  $A$  изоморфна группе  $A(> N)$ , определенной равенствами (14), (16), (18), где  $p \in P$ ,  $i, n, x \in \omega$ ,  $P$  - множество всех простых чисел,  $\omega$  - множество всех натуральных чисел.

ЛЕММА 2. В паре  $(A, \mu)$  существует вычислимо перечислимая и максимально 2- линейно независимая последовательность элементов  $\langle \mu t_i \mid i \in \omega \rangle$ .

$$(19)$$

По этой последовательности определим предикат  $D(i, p, n, x)$ , положив

$$D(i, p, n, x) = \text{истинен} \Leftrightarrow p^n \mu x = \mu t_i. \quad (20)$$

По предикату  $D(i, p, n, x)$  определим следующие множества и группы, положив:

$$P_\omega(i) = \{ p \mid \forall n \exists x_{i,p,n} D(i, p, n, x_{i,p,n}) \}, \quad (21)$$

$$E_i(\leq N) = qr \left\{ \frac{1}{p^n} \mid p \in P_\omega(i), |P_\omega(i)| \leq N \right\}, \quad (22)$$

$$E_i(> N) = qr \left\{ \frac{1}{p^n} \mid p \in P_\omega(i), |P_\omega(i)| > N \right\}, \quad (23)$$

$$E(\leq N) = \bigoplus \{ E_i(\leq N) \mid i \in \omega \}, \quad (24)$$

$$E(> N) = \bigoplus \{ E_i(> N) \mid i \in \omega \}, \quad (25)$$

$$E = E(\leq N) \oplus E(> N). \quad (26)$$

ЛЕММА 3. Группа  $E_{i,p}$ , определенная равенством (24), изоморфна:

а) либо циклической группе, порожденной элементом, отмеченной меткой (\*).

в) либо группе

$$A_p \cong qr \left\{ \frac{1}{p^n} \mid n \in \omega \right\}. \quad (27)$$

Для любого числа  $i \in \omega$  определим группу  $E_i$  положив:

$$E_i \simeq qr \{ e_{i,p,n} \mid p \in P, n \in \omega \}, \quad (28)$$

где элементы  $e_{i,p,n}$  определены при доказательстве леммы 3.

ЛЕММА 4. Пусть для числа  $i \in \omega$  справедливо равенство

$$P_\omega(i) = \emptyset, \quad (29)$$

где множество  $P_\omega(i)$  определено равенством (9).

Тогда группа  $E_i$ , определенная равенством (28), изоморфна циклической группе, порождающие элементы которой отмечены меткой (\*).

ЛЕММА 5. Пусть для числа  $i \in \omega$  справедливо равенство

$$P_\omega(i) \neq \emptyset, \quad (30)$$

где множество  $P_\omega(i)$  определено равенством (9).

Тогда группа  $E_i$ , определенная равенством (25) изоморфна группе

$E_{\omega,i} \text{ qr}\{e_{i,p,n} \mid \text{истинен предикат } D(i, p, n, e_{i,p,n}), n \in \omega, p \in P_\omega(i)\}$ , (31) где элементы  $e_{i,p,n}$  не отмечены меткой (\*).

Из лемм (4) и (5) следует

ЛЕММА 6. Для любого числа  $i \in \omega$  группа  $E_i$ , определенная формулой (38), изоморфна

(а) либо циклической группе  $E_{<\omega}(i)$ , порождающие элементы которой, отмечены меткой (\*),

(в) либо группе  $E_\omega(i) \cong \text{qr}\{e_{i,p,n} \mid P_\omega(i) \neq \emptyset, p \in P_\omega(i), n \in \omega\}$ , (32) где множество  $P_\omega(i)$  определено формулой (9).

Определим группу  $E$  равенством  $E = \bigoplus \{E_i \mid i \in \omega\}$ , (33)

где группа  $E_i$  определена формулой (28).

ЛЕММА 7. Подгруппа  $E_i$  группы  $E$ , определенные формулами (28) и (33) по вычислимому предикату  $D(i, p, n, x)$ , как в доказательстве леммы 3, равномерно по  $i$  вычислимо перечислимо определена.

СЛЕДСТВИЕ 2. Группа  $E$ , определенная формулой (33), вычислимо перечислимо определена.

ЛЕММА 8. Для любого числа  $i \in \omega$  подгруппа  $E_i(> N)$  группы  $E_i$ , определенные формулами (23) и (28) соответственно по вычислимому предикату  $D(i, p, n, x)$ , как в доказательстве достаточности условий теоремы 1, равномерно по  $i$  вычислимо перечислимо определена.

СЛЕДСТВИЕ 3. Группа  $E(> N)$  вычислимо перечислимо определена.

В [4] доказана следующая

ТЕОРЕМА 2. Любая вычислимо перечислимо определенная абелева группа без кручения вычислима.

Отсюда и следствия 3 следует, что группа  $E(> N)$ , определенная формулой (25), вычислима.

Из теоремы 1, где группа  $A = \bigoplus \{A_{p(i)} \mid i \in \omega\}$  и  $A_{p(i)}$  – аддитивная группа рациональных чисел, знаменателями которых являются степенями некоторого простого числа  $p(i)$ , и  $N=1$ , следует теорема, доказанная в [5].

#### Список литературы

- [1] Мальцев А.И. О рекурсивных абелевых группах // Доклады АН СССР. 1962. Т 146, №5. С.1009-1012.
- [2] Хисамиев Н.Г., Крыкпаева А.А. Эффективно вполне разложимые абелевы группы // Сиб. Мат. журнал. 1997. Т.38, №6. С.1410-1412.
- [3] Хисамиев Н.Г. Об одном классе сильно разложимых абелевых групп // Алгебра и логика. 2002. Т.41, №4. С.493-509.
- [4] Khissamiev N.G. Constructive abelian groups // Studies in Logic and the Foundation of Mathematics. Elsevier, 1998. V.139. P. 1177-1231.
- [5] Downey, R., Goncharov, S., Kach, A.M., Knight, J., Kudinov, O., Melnikov, A.G., Turetsky, D. Decidability and Computability of Certain Torsion-Free Abelian Groups // Notre Dame J. Formal Log. 2010. V. 51, С. 85-96.

## ABOUT GENERATED FINITE POINTED ABELIAN GROUP

Basheyeva A.O.  
L.N. Gumilyov ENU, Astana  
[basheeva@mail.ru](mailto:basheeva@mail.ru)

A *pointed group* is an algebra  $\langle G; +, -, 0, C \rangle$  of signature  $\langle +, -, 0, C \rangle$ , where  $C$  is a finite set of constants, which  $\langle +, -, 0 \rangle$ -reduct is a group. And a pointed group is *Abelian* if its group's reduct is Abelian group. Also we say that an algebra  $\langle G; +, -, 0, C \rangle$  is a *pointed enrichment* of a group  $\langle G; +, -, 0 \rangle$  and denote it by  $G^c$ . A *quasivariety* is a class  $\mathcal{K}$  of similar algebras that closed with respect to subalgebras, direct products and ultraproducts. Equivalently, a quasivariety is the same thing as a class of similar algebras axiomatized by a set of quasi-identities.

A *quasi-identity* means a universal Horn sentence with the non-empty positive part, that is of the form

$$\forall \bar{x} [p_1(\bar{x}) \approx q_1(\bar{x}) \wedge \dots \wedge p_n(\bar{x}) \approx q_n(\bar{x}) \rightarrow p(\bar{x}) \approx q(\bar{x})]$$

where  $p_i(x), q_i(x), p(x), q(x)$  - are terms of signature  $\sigma$ . A variety is a *quasivariety* which is closed under homomorphisms.

A quasivariety  $\mathcal{R} \subseteq \mathcal{K}$  is *finitely axiomatizable* with respect to  $\mathcal{K}$  if the number of upper covers  $\mathcal{R}$  in the lattice of quasivariety  $\mathcal{K}$  is finite.

It follows from the main theorem on the structure of a finite Abelian group that the quasivariety generated by a finite Abelian group is a variety; consequently, by the Oates-Powell theorem, it has a finite basis of (quasi)identities. At the same time it is easy to construct a finite pointed Abelian group that generates a proper quasivariety, that is, a quasivariety that is not a variety. Thus, the study of the above mentioned problem for finite pointed groups is natural to start within the finite pointed Abelian groups.

We prove that every pointed enrichment of finite Abelian group generates finitely axiomatizable variety and quasivariety.

### References

1. R. M. Bryant, The laws of finite pointed groups, Bull. London Math. Soc., 14 (1982), 119–123.
2. M.I. Korgapolov, Y.I. Merzlykov Basis of group theory, (Moscow, Nauka 1982),
3. A. M. Nurakunov, Quasivariety lattices of pointed Abelian groups, Algebra and Logic, 53 (2014), 238–257.
4. S. Oates and M. B. Powell, Identical relations in finite groups, Journal of Algebra, 1 (1964), 11–39.

# ON SOME PROPERTIES OF THE QUASIVARIETY LATTICE OF LUKASIEWICZ ALGEBRAS

*S.M. Lutsak<sup>1</sup>, O.A. Voronina<sup>2</sup>*

*M. Kozybayev North Kazakhstan University, Petropavlovsk, Kazakhstan  
e-mail: <sup>1</sup>sveta.lutsak@mail.ru, <sup>2</sup>oavyu@mail.ru*

**Abstract.** We study Birkhoff-Maltsev problem for Lukasiewicz algebras. The main purpose of this work is to identify nontrivial properties of the quasivariety lattice  $Lq(\mathbf{L})$  of the variety  $\mathbf{L}$  generated by the set of all finite Lukasiewicz algebras. We prove that the investigated quasivariety lattice  $Lq(\mathbf{L})$  is  $Q$ -universal, the quasivariety  $\mathbf{L}$  contains continuum many  $Q$ -universal subquasivarieties, continuum many subquasivarieties with no upper covers in the lattice  $Lq(\mathbf{L})$ , continuum many subclasses  $\mathbf{K} \subseteq \mathbf{L}$  with the property (N), continuum many subclasses  $\mathbf{K} \subseteq \mathbf{L}$  with the property (N) but which are not  $Q$ -universal, continuum many non-standard subquasivarieties without an independent basis of quasi-identities, and continuum many non-standard subquasivarieties with the so-called finitely split basis of quasi-identities. We show the highest complexity of considered quasivariety lattice  $Lq(\mathbf{L})$ .

**Introduction, relevance and literature review.** The present research is related to the fundamental Birkhoff-Maltsev problem of describing lattices that are isomorphic to the quasivariety lattices. This problem was posed in 1945 by G. Birkhoff and, independently, in 1966 by A.I. Maltsev and is relevant to this day. It is one of the wellknown problems of universal algebra and lattice theory. A large number of papers are devoted to the study of the properties of lattices isomorphic to the quasivariety lattices of algebraic structures. Among them are the works of A.I. Maltsev, V.A. Gorbunov and V.I. Tumanov, M. Adams and W. Dziobiak, M.V. Sapir, A.M. Nurakunov, A.V. Kravchenko, V.K. Kartashov, M.V. Schwidefsky and other. Despite the fact that the problem has been posed for a long time and the study of quasivariety lattices has been intensively carried out for several decades in a number of countries (Kyrgyzstan, Poland, Russia, USA), effective approaches to its solution have not yet been developed. Moreover, the results obtained so far demonstrate the exceptional structural and algorithmic complexity of such lattices [1-13]. Thus, the progress in the study of the properties of quasivariety lattices is a very topical issue. The study of the Birkhoff-Maltsev problem in specific classes of algebraic structures is of undoubted interest and deserves attention.

Recall that a quasivariety is a class of structures of the same type which is closed under substructures, direct products (including the direct product of an empty family) and ultraproducts. A variety is a quasivariety which is closed under homomorphic images. A quasivariety  $R$  which is contained in a quasivariety  $M$  is called a subquasivariety of  $M$ . The set  $Lq(M)$  of all subquasivarieties of a given quasivariety  $M$  forms a complete lattice (under inclusion) which is called a lattice of quasivarieties of  $M$  or a quasivariety lattice of  $M$ .

Currently, there are several approaches to the concept of complexity. Among others, two measures of the complexity of the structure of quasivariety lattices are well known:  $Q$ -universality and the property (N) or noncomputability of the set of all types of isomorphism of the class of finite sublattices of the quasivariety lattice. The presence in the quasivariety lattices continuum many elements which have no covers also tells about the complexity of the structure of these lattices; in this case, there is continuum many subquasivarieties of this quasivariety  $M$  which have no independent basis of quasi-identities relatively  $M$ .

The concept of  $Q$ -universality was introduced by M.V. Sapir [2] in 1985. A quasivariety  $M$  is  $Q$ -universal if, for any quasivariety  $R$  of a finite type, the quasivariety lattice  $Lq(R)$  is a homomorphic image of some sublattice of the quasivariety lattice  $Lq(M)$ . In this case the quasivariety lattice  $Lq(M)$  of a quasivariety  $M$  is called  $Q$ -universal. Note that  $Q$ -universality indicates maximum complexity in the lattice-theoretic sense. To date, we know a lot of different  $Q$ -universal classes and the number of such examples is constantly growing. In [2] M.V. Sapir established  $Q$ -universality of the quasivariety generated by one particular semigroup. A.M.

Nurakunov proved  $Q$ -universality of the quasivariety of pointed Abelian groups [3]. In the paper of M. Adams and W. Dziobiak [4] found sufficient conditions for  $Q$ -universality. These conditions have obtained generalization in the paper of M.V. Schwidefsky [5].

Another approach to the complexity of the quasivariety lattices was suggested by A.M. Nurakunov [6] in 2012. We say that a class  $\mathbf{M}$  of algebraic structures of a fixed type has the property (N) if the set of all isomorphism types of finite sublattices of the quasivariety lattice  $Lq(\mathbf{M})$  is not computable. The presence of the property (N) indicates algorithmic complexity of the quasivariety lattice. In the papers [3,6] A.M. Nurakunov built the first examples of classes which have the property (N).

Based on the ideas of A.M. Nurakunov, in the paper of M.V. Schwidefsky and A. Zamojska-Dzienio [7] a connection between  $Q$ -universality and the property (N) was established. Namely, it has been proved that class  $\mathbf{K}$  of all structures of type  $\sigma$  is  $Q$ -universal if and only if it contains a subclass possessing the property (N) [7]. In this regard, the following problem arose [5,7]. Is it true that any  $Q$ -universal class of structures  $\mathbf{K}$  of a fixed type contains a subclass having the property (N)? Is there a class  $\mathbf{K}$  that is not  $Q$ -universal but nevertheless it has the property (N)?

M.V. Schwidefsky has been given a positive answer to the first question for almost all currently known  $Q$ -universal quasivarieties [5]. That is, it was proved that almost all known  $Q$ -universal quasivarieties contain classes having property (N) [5]. S.M. Lutsak has been given a positive answer to the second question, namely, proved that for different types there is continuum many classes  $\mathbf{K}$  which have the property (N) but are not  $Q$ -universal [8].

According to [4,5], the presence of an AD-class (an Adams-Dziobiak class) in a quasivariety  $\mathbf{R}$  is a sufficient condition for the  $Q$ -universality of  $\mathbf{R}$ . In [8] it is proved that if a class  $\mathbf{K}$  of algebraic structures contains an AD-class then there exists continuum many subclasses  $\mathbf{M}$  that have the property (N) but which are not  $Q$ -universal. For this subclass  $\mathbf{M}$  the set of all types of isomorphism of the class of finite sublattices of  $Lq(\mathbf{M})$  is not computable, which means that there is no algorithm that would determine whether a finite lattice is embeddable in the lattice  $Lq(\mathbf{M})$  or not.

In the paper of A.V. Kravchenko, A.M. Nurakunov, and M.V. Schwidefsky [9] the concept of a B-class (with respect to a quasivariety  $\mathbf{K}$ ) was introduced and it was proved that any B-class (with respect to a quasivariety  $\mathbf{K}$ ) is an AD-class, which immediately implies the  $Q$ -universality of a quasivariety  $\mathbf{K}$  containing such a class. In addition, the presence of a B-class in a quasivariety  $\mathbf{K}$  is a sufficient condition for the existence of continuum many subquasivarieties of quasivariety  $\mathbf{K}$  which have no covers in the lattice  $Lq(\mathbf{K})$ , and therefore have no an independent basis of quasi-identities with respect to  $\mathbf{K}$ , which also indicates the complexity of the structure of such a lattice [9].

The paper [13] investigated the questions of the standardness of quasivarieties and found sufficient conditions (the presence of a B-class in a quasivariety  $\mathbf{K}$ ) under which a quasivariety contains a continuum of non-standard subquasivarieties without an independent basis of quasi-identities. Note that a quasivariety  $\mathbf{K}$  is called standard if every Boolean topological structure whose algebraic reduct belongs to  $\mathbf{K}$  is profinite in  $\mathbf{K}$  [13].

**Methods, main results and discussion.** In this work we study Birkhoff-Maltsev problem for Lukasiewicz algebras. We consider the following measures of the highest complexity of the structure of quasivariety lattices that were introduced in [2,6,9], respectively:  $Q$ -universality, the property (N), and existence of continuum of quasivarieties without covers in a given quasivariety lattice. The presence of continuum many non-standard subquasivarieties without an independent basis of quasi-identities and continuum many non-standard subquasivarieties with the so-called finitely split basis of quasi-identities also indicates the complexity of such a lattice.

In our research the methods of lattice theory and universal algebra were used, in particular, the method of constructing classes (quasivarieties) with certain properties, developed in the papers of A.M. Nurakunov, lattice-theoretic methods from the works of M.V.

Schwedfsky, the method of AD-classes, developed in the papers of M. Adams and W. Dziobiak, the method of B-classes from [9,13].

The main purpose of this work is to identify nontrivial properties of the quasivariety lattice  $Lq(\mathbf{L})$  of the variety  $\mathbf{L}$  generated by the set of all finite Lukasiewicz algebras.

We prove that the quasivariety lattice  $Lq(\mathbf{L})$  of the variety  $\mathbf{L}$  generated by the set of all finite Lukasiewicz algebras is Q-universal, the quasivariety  $\mathbf{L}$  contains continuum many Q-universal subquasivarieties, continuum many subquasivarieties with no upper covers in the lattice  $Lq(\mathbf{L})$ , continuum many subclasses  $\mathbf{K} \subseteq \mathbf{L}$  with the property (N), continuum many subclasses  $\mathbf{K} \subseteq \mathbf{L}$  with the property (N) but which are not Q-universal, continuum many non-standard subquasivarieties without an independent basis of quasi-identities, and continuum many non-standard subquasivarieties with the so-called finitely split basis of quasi-identities. We show the highest complexity of the quasivariety lattice  $Lq(\mathbf{L})$  of the variety  $\mathbf{L}$  generated by the set of all finite Lukasiewicz algebras.

Recall that the Lukasiewicz algebra is called the algebra

$$\mathcal{L}_p = \left( \left\{ 0, \frac{1}{p}, \dots, \frac{p-1}{p}, 1 \right\}, \rightarrow, \neg \right), p \geq 1,$$

with operations defined as follows: for all  $x, y$ ,  $x \rightarrow y = \min\{1, 1 - x + y\}$  and  $\neg x = 1 - x$ .

The main result of this work is the following theorem.

**Theorem 1.** Let  $\mathbf{L}$  be the variety generated by the set of all finite Lukasiewicz algebras. Then  $\mathbf{L}$  is Q-universal and contains continuum many

- (1) Q-universal quasivarieties;
- (2) quasivarieties having no covers in the quasivariety lattice  $Lq(\mathbf{L})$ ;
- (3) subclasses  $\mathbf{K} \subseteq \mathbf{L}$  having the property (N);
- (4) subclasses  $\mathbf{K} \subseteq \mathbf{L}$  having the property (N) but which are not Q-universal;
- (5) non-standard subquasivarieties without an independent basis of quasi-identities;
- (6) non-standard subquasivarieties with the so-called finitely split basis of quasi-identities.

Note that in [12] it was proved that the considered quasivariety lattice  $Lq(\mathbf{L})$  does not satisfy to any non-trivial lattice's identity.

**Conclusion.** The results obtained demonstrate the structural and algorithmic complexity of the quasivariety lattice of the variety generated by the set of all finite Lukasiewicz algebras.

This work is theoretical. The results obtained will be used for further study of quasivariety lattices. The result similar to Theorem 1 holds for some other quasivarieties of algebraic structures.

This research is funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP09058390).

## References

1. Gorbunov V.A. Algebraic theory of quasivarieties. – Novosibirsk: Scientific book, 1999. – 368 p.
2. Sapir M.V. The lattice of quasivarieties of semigroups // Algebra Univers. – 1985. – Vol. 21, № 2/3. – P. 172-180.
3. Nurakunov A.M. Quasivariety lattices of pointed Abelian groups // Algebra and Logic. – 2014. – Vol. 53, № 3. – P. 372-400.
4. Adams M.E., Dziobiak W. Q-universal quasivarieties of algebras // Proc. Amer. Math. Soc. – 1994. – Vol. 120, № 4 – P. 1053-1059.
5. Schwedfsky M.V. On the complexity of quasivariety lattices // Algebra and Logic. – 2015. – Vol. 54, № 3. – P. 381-398.
6. Nurakunov A.M. Unreasonable lattices of quasivarieties // Internat. J. Algebra Comput. – 2012. – Vol. 22, № 3. – P. 1-17.

7. Schwidefsky M.V., Zamojska-Dzienio A. Lattices of subclasses. II // *Internat. J. Algebra Comput.* – 2014. – Vol. 24, № 8. – P. 1099-1126.
8. Lutsak S.M. On the complexity of quasivariety lattices // *Sib. Electron. Math. Rep.* – 2017. – Vol. 14. – P. 92-97.
9. Kravchenko A.V., Nurakunov A.M., Schwidefsky M.V. Structure of quasivariety lattices. I. Independent axiomatizability // *Algebra and Logic.* – 2019. – Vol. 57, № 6. – P. 445-462.
10. Schwidefsky M.V. On sufficient conditions for Q-universality // *Sib. Electron. Math. Rep.* – 2020. – Vol. 17. – P. 1043-1051.
11. Schwidefsky M.V. Existence of independent quasi-equational bases // *Algebra and Logic.* – 2020. – Vol. 58, № 6. – P. 514-537.
12. Dziobiak W. On subquasivariety lattices of semi-primal varieties // *Algebra Universalis.* – 1985. – Vol. 20. – P. 127-129.
13. Kravchenko A.V., Nurakunov A.M., Schwidefsky M.V. Structure of quasivariety lattices. IV. Nonstandard quasivarieties // *Siberian Math. J.* – 2021. – Vol.62, № 5. – P. 850-858.

METHODOLOGICAL FOUNDATIONS OF INTEGRATED INFORMATION SECURITY OF THE ORGANIZATION'S INFORMATION AND COMMUNICATION INFRASTRUCTURE

*Idrissova I.A., Mukhanova A.A.*

*Eurasian National University named after L.N. Gumilyov, Nur-Sultan,  
[indira\\_idrisova@mail.ru](mailto:indira_idrisova@mail.ru)*

**Annotation**

Organizations implement information security for a wide range of reasons. The main objectives of InfoSec are typically related to ensuring confidentiality, integrity, and availability of company information. Since information security covers many areas, it often involves the implementation of various types of security, including application security, infrastructure security, cryptography, incident response, vulnerability management, and disaster recovery. This methodological foundation provides an in-depth look into the field of information security, as well as roles and responsibilities.

**Keywords:** information and communication infrastructure, information security, functional model of complex systems, system of integrated information security, IDEF0.

**Introduction**

Today, the problem of protecting computer networks from unauthorized access has become particularly acute. The amount of information stored in modern computers, and its importance makes it increasingly necessary to strengthen the protection of information, like storing cash in banks. Copyright, national security, legal issues, private life - all these aspects of human activity require special attention to internal control in commercial and government organizations.

Security is a complex concept and cannot be considered as a simple sum of its constituent parts. Each part is critically important, these parts are interconnected and interdependent, therefore, when solving the problems of integrated information security, a synergistic effect is clearly manifested: information security tools, on the one hand, are an integral part of the system, on the other hand, they themselves organize the system by carrying out protective measures. In the absence of individual components of the system or their inconsistency with each other, the emergence of vulnerabilities in information security technology is inevitable. As a consequence, one of the basic principles in the creation and development of an effective information security system should be the principle of consistency [1].

Integrated information security of the information and communication infrastructure allows to obtain the most complete and objective assessment of the security of information resources, localize existing problems and develop an effective program for building an organization's information security management system. The development, implementation and operation of an integrated information security system is a multifunctional dynamic system, which is characterized by features inherent in complex systems. A complex system with functional decomposition is divided into simpler subsystems, each of which, in turn, represents a certain set of even less complex structural units that perform independent functions.

**Methods**

To describe the elements of a decomposed system, its representation as a stochastic system with incomplete information about states can be used. At the same time, since the tasks facing the system of integrated information security (hereinafter referred to as IIS) are difficult to formalize, as the analysis shows, the most preferred methods of description are such as expert assessments, simulation modeling, methods of fuzzy sets theory, linguistic variables, soft



measurements, etc. It is advisable to build a functional model of complex systems, such as the IIS system, in accordance with the IDEF standards (IDEF0, IDEF3, IDEF5, DFD, etc.). The development of models in these standards allows to visually and effectively display the entire mechanism of creation, implementation and operation of an integrated information security system in the required context. The results of the IDEF0 analysis can be used in conducting research using IDEF3 models and DFD data flow diagrams. A distinctive feature of the IDEF0 language is the use of experts' natural language as a basis, which is structured using graphical tools. This allows the expert to freely describe the functioning of the system using familiar and convenient terminology, and then easily transfer the description in natural language to the graphical representation of the IDEF0 language. In the IDEF0 standard, the description of the system is organized in the form of hierarchically ordered and interconnected diagrams. The top of this tree structure represents the most general description of the system and its interaction with the external environment. At the base of the structure are the most detailed descriptions of the functions performed by the system. Diagrams contain functional blocks connected by arcs. Arcs represent interactions and relationships between blocks. The functional block in the diagrams is represented by a rectangle and represents a function or an active part of the system. The names of the blocks are verbs or verb phrases. Each side of the block has a special, well-defined purpose. The arcs of the inputs fit to the left side of the block, the control arcs fit to the upper side, the arcs of the mechanisms for implementing the function performed fit to the lower side, and the arcs of the outputs are directed from the right side. Such an agreement assumes that using the control information about conditions and restrictions, as well as the mechanism implementing it, the block function converts its inputs into the corresponding outputs. In the diagram, the blocks are ordered starting from the upper-left corner and ending with the lower-right corner. To provide clarity and a better understanding of the simulated processes, it is recommended to use from 3 to 6 blocks in one diagram. This representation of the model eliminates the ambiguity inherent in natural language, and due to this, the conciseness and accuracy of the description necessary for understanding and analysis are achieved, without loss of detail and quality [2].

### **Discussion**

The general controlling influence for the IIS processes are the provisions of relevant laws, GOST standards, and guidance documents (hereinafter referred to as GD). The implementation of various processes displayed in the models is carried out by employees of the information security department, subjects of information relations, experts in the field of information security. Since, in accordance with the IDEF0 standard, the system is represented as a set of nested functions forming a hierarchy, first of all, a function describing the system as a whole must be defined, that is, a so-called context diagram should be constructed. Then the hierarchical decomposition of each block can be continued to the required level of detail. The general context diagram of integrated information security is shown in Figure 1. The corresponding decomposition diagrams are given below. The process of Development, implementation and operation of the Integrated Information Security System (IISS) is depicted in Figure 2, Identification of information security threats and identification of information security vulnerabilities is depicted in Figure 3, Definition of the concept of protection of information assets and development of information security policy (ISP) is depicted in Figure 4. IS (Information Security) planning as a management function is a process of consistently removing uncertainty about the structure and composition of security tools at the management facility [3].

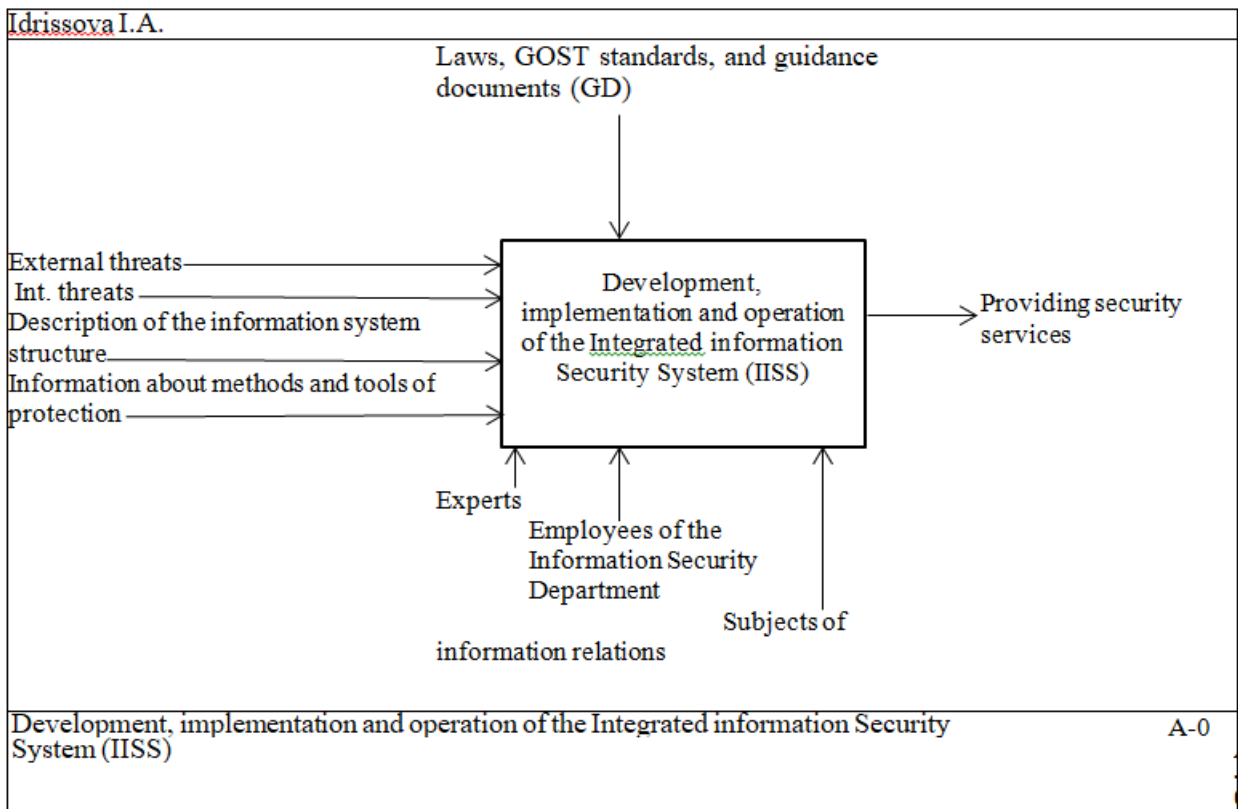


Figure 1. Context diagram of the IIS model

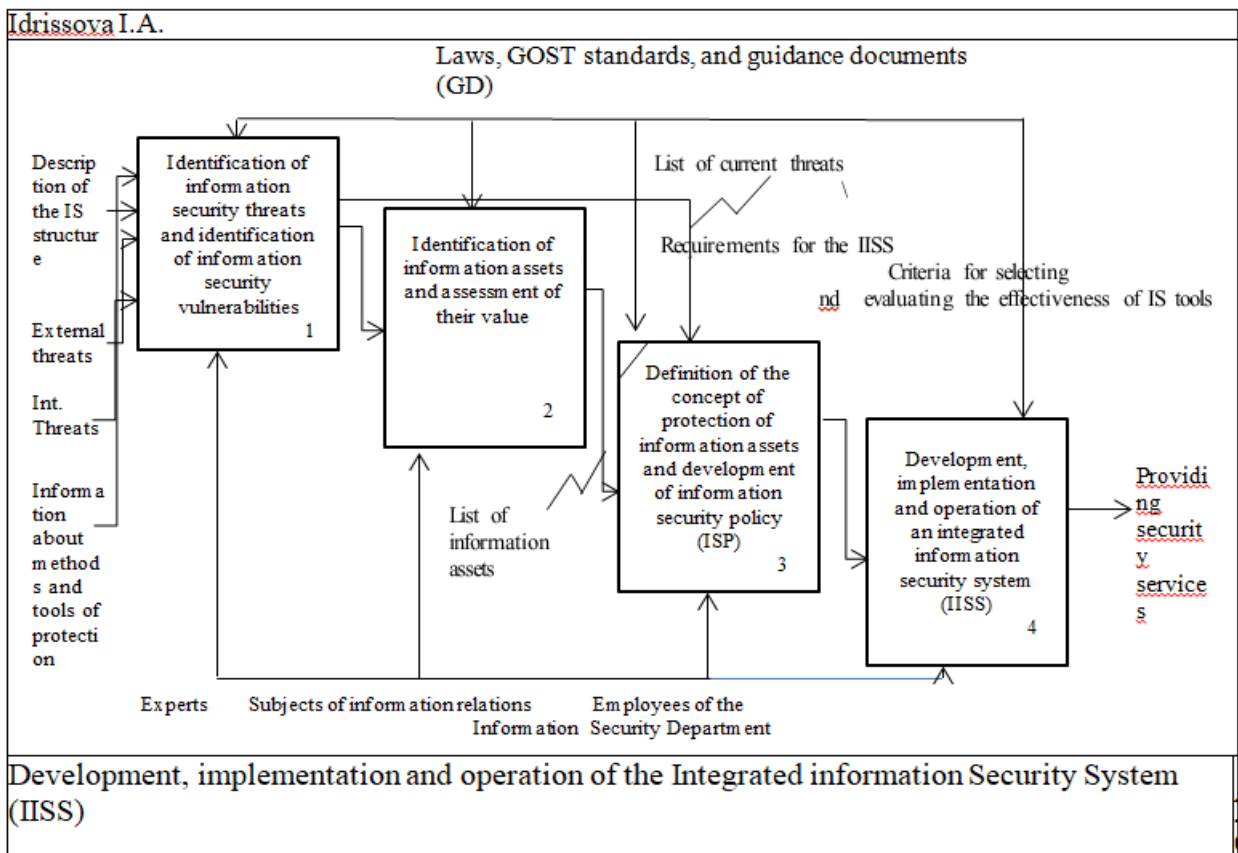


Figure 2. Decomposition diagram of the IIS model (A-0)

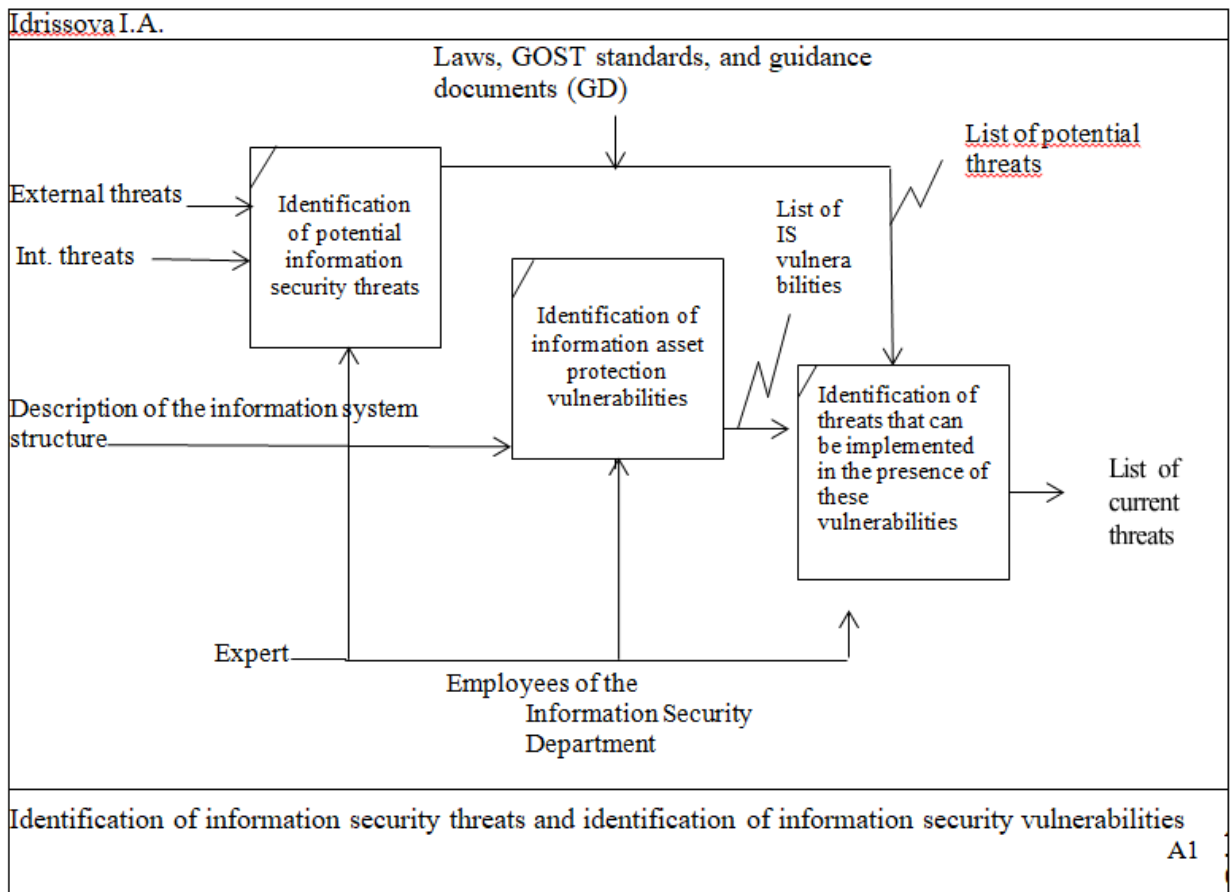


Figure 3. Decomposition diagram of the IIS model A1

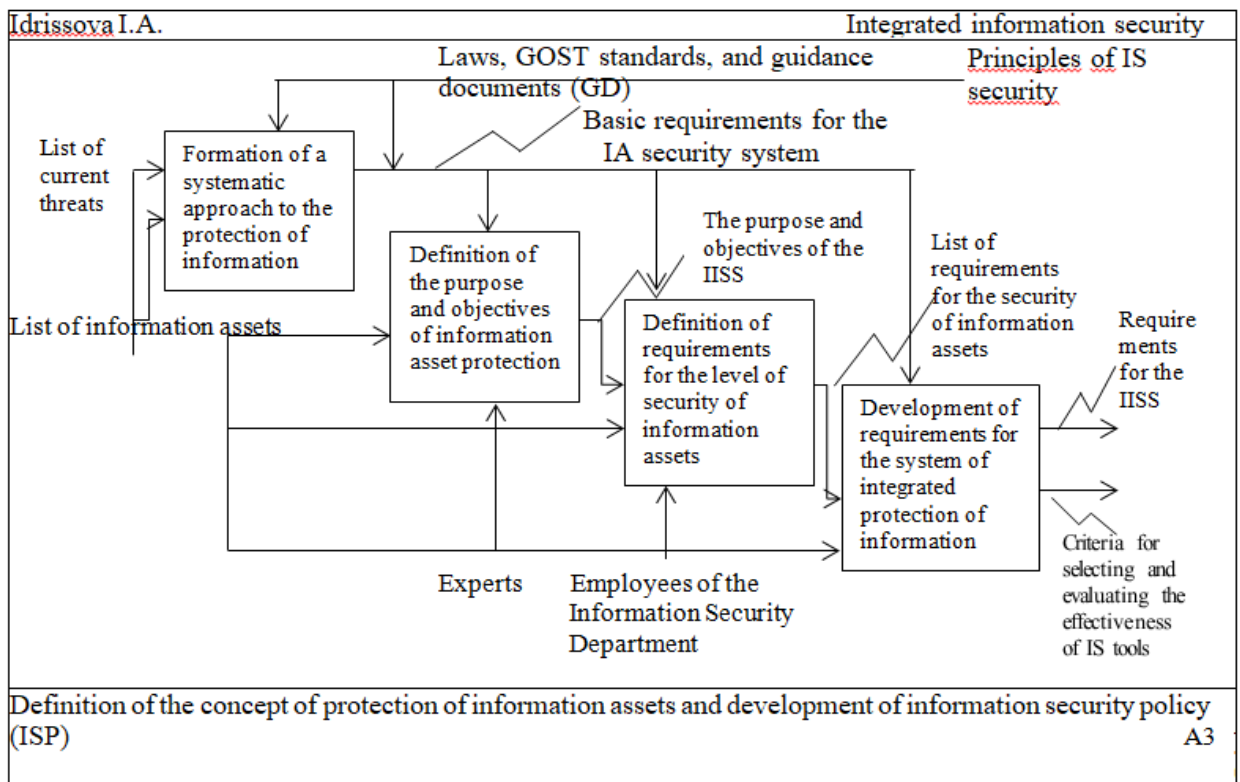


Figure 4. Decomposition diagram of the IIS model A3

## Conclusion

The functional model of the IIS developed within the framework of the IDEF0 standard makes it possible to establish which functions are performed by the system, in what sequence, who is responsible for carrying out specific work, which is the result of the implementation of a particular procedure. Thus, the IDEF0 tools allows to effectively simulate the process of integrated information security in order to optimize the management of information asset protection.

The level of details in the description of information handling process based on IDEF0 standard allows constructing a functional model, which will adequately describe real processes and account for significant factors affecting information. The factors of risk that pose threats to information security are estimated to formulate final values of the factors of risk that characterize the total level of threat, potential damage and vulnerability of the protected object (vulnerability scales) [10].

## Literature Review

1. Azhmukhamedov I.M. Conceptual model of complex system security management. Bulletin of ASTU. Series: "Management, computer technology and informatics" 1/2010, p.62-66
2. Cheremnykh S.V., Semenov I.O. Ruchkin V.S. Modeling and analysis of systems. IDEF-technologies:workshop. M.: Finance and statistics, 2006
3. Azhmukhamedov I.M. Solution of information security problems based on system analysis and fuzzy cognitive modeling // Monograph, Astrakhan 2012, p. 100-130
4. Boranbayev, S., Goranin, N. and Nurusheva, A. (2018). The methods and technologies of reliability and security of information systems and information and communication infrastructures. Journal of Theoretical and Applied Information Technology, 96(18), 6172–6188.
5. ISO/IEC 27001:2013 Information technology — Security techniques — Information security management systems — Requirements
6. Vernikov G. The main methodologies for survey organizations. IDEF0 standard. url: <https://www.cfin.ru/vernikov/idef/idef0.shtml>
7. Konev A Approach to creation protected information model 2012 Reports of Tomsk State University of Control Systems and Radioelectronics pp 34–39
8. Grant T, Burke I and van Heerde R 2012 Comparing Models of Offensive Cyber Operations Proceedings of the 7th International Conference on Information Warfare and Security p 108
9. Jouinia M, Ben Arfa Rabaia L, Ben Aissa 2014 Classification of security threats in information systems Procedia Computer Science 32 489–496
10. Mikov D 2016 Identification of information risk factors of the telemedicine system Polytechnic youth magazine MSTU. N.E. Bauman p 1.

## ТЕХНОЛОГИЯ БЛОКЧЕЙН И БЕЗОПАСНОСТЬ КОНЕЧНЫХ УСТРОЙСТВ ИНТЕРНЕТА ВЕЩЕЙ

*А.Т.Тохметов<sup>1</sup>, Г.Б. Абдикеримова<sup>1</sup>, В.Э. Ли<sup>1</sup>, Н.К. Токашев<sup>2</sup>*

<sup>1</sup> *Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан*

<sup>2</sup> *Esil University, Астана, Казахстан*

[attohmetov@mail.ru](mailto:attohmetov@mail.ru)

На сегодняшний день наблюдается рост использования различных информационных технологий, а спрос на безопасность передаваемых данных пропорционально растет. Одной из таких технологий является интернет вещей (IoT), который очень активно и быстро развивается.

Всего восемь лет назад количество подключенных к сети интернета вещей практически равнялось числу людей, имеющих доступ к мировой паутине. Но уже к 2030 году, по прогнозам футурологов, инфраструктура интернета вещей распространится на 1 трлн. устройств и 100 млн. приложений. Аналитики фирмы McKensey оценивают потенциальный вклад интернета вещей в глобальную экономику в 11 трлн. долларов уже к 2025 году. В идеале общество ближайшего будущего – общество, в котором практически к любому бытовому и производственному процессу можно будет прибавить приставку smart [1].

Развитие интернета вещей предоставляет огромный потенциал для их повсеместного использования как в повседневной, так и в профессиональной деятельности человека. Устройства, которые почти не потребляют энергию, а вследствие этого имеют возможность длительно работать без подключения к сети электроэнергии и отправлять сигнал на внушительные расстояния, всегда будут востребованными. Часть подобных устройств работают на технологии LPWAN (различные метеорологические датчики, счетчики ЖКХ и подобные устройства, не требующий большого объема передачи данных).

Безопасность устройств IoT в течение некоторого времени вызывала беспокойство и имела неизбежные последствия, допуская как мелкие, так и крупномасштабные атаки. Большинство этих атак происходят из-за простых проблем с безопасностью. Устройства IoT могут доверять локальной сети до такого уровня, что дальнейшая аутентификация или авторизация не требуются. Любое другое устройство, подключенное к той же сети, также считается доверенным. Это особенно является проблемой, когда IoT-устройство подключено к Интернету: теперь любой человек в мире потенциально может получить доступ к функциям, предлагаемым устройством.

Когда устройство обменивается данными в виде простого текста, вся информация, которой обмениваются с клиентским устройством или серверной службой, может быть получена «человеком посередине» (MitM). Любой, кто может получить позицию на сетевом пути между устройством и его конечной точкой, может проверить сетевой трафик и потенциально получить конфиденциальные данные, такие как учетные данные для входа.

Большинство IoT-устройств фактически представляют собой компьютеры общего назначения, на которых можно запускать определенное программное обеспечение. Это позволяет злоумышленникам устанавливать собственное программное обеспечение, функциональность которого не является частью нормального функционирования устройства. Например, злоумышленник может установить программное обеспечение, которое выполняет DDoS-атаку. Ограничение функциональности устройства и программного обеспечения, которое оно может запускать, ограничивает возможности злоупотребления устройством.

Когда устройство взломано, оно часто продолжает нормально работать с точки зрения пользователя. Любая дополнительная пропускная способность или потребление

энергии обычно не обнаруживаются. Большинство устройств не имеют функции ведения журнала или оповещения, чтобы уведомить пользователя о любых проблемах с безопасностью. Если они есть, их можно перезаписать или отключить при взломе устройства. В результате пользователи редко обнаруживают, что их устройство подверглось атаке или было скомпрометировано, что не позволяет им принять меры по смягчению последствий.

Использование IoT устройств требует хорошей защиты передаваемой информации, для минимизации вероятности перехвата данных, а также спуффинга и других подобных атак, кроме того, обеспечить надежное хранение уже полученных данных. Одним из возможных решений данных проблем в данном докладе будет рассматриваться блокчейн – еще одна современная быстроразвивающаяся технология.

Блокчейн — это многофункциональная и многоуровневая информационная технология, предназначенная для надежного учета различных активов. Потенциально эта технология охватывает все без исключения сферы любой деятельности и имеет множество областей применения. В их число входит также и интернет вещей. [2]

Интернет вещей – это глобальная сеть компьютеров, датчиков (сенсоров) и исполнительных устройств (актуаторов), связывающихся между собой с использованием интернет протокола IP (Internet Protocol). Например, для решения определенной задачи компьютер связывается через публичный интернет с небольшим устройством, к которому подключен соответствующий датчик (например, температуры), как это показано на рисунке 1. [3]



Рисунок 1. Связка умных устройств через интернет вещей.

Технологии LPWAN (Low Power Wide Area Network) предназначены для M2M (Machine-to-Machine)-приложений, которые требуют низкоскоростной передачи данных по радиоканалу и работы без присмотра в течение длительного периода времени, возможно, в отдаленных или труднодоступных местах.

Особенности LPWAN — низкое энергопотребление (low-power) и широкий территориальный охват (wide-area).

Очень небольшая потребность в энергии для передачи данных и, следовательно, большой срок службы батареи, имеет решающее значение для встраиваемой техники. Прогнозируется, что сети LPWAN будут применяться в широком спектре приложений интернета вещей, таких, как отслеживание производственных активов, мониторинг безопасности, учет потребления воды и газа, а также в интеллектуальных сетях, городских парковках, торговых автоматах и городском освещении. Технология может использоваться и для подключения носимых устройств, например, в трекерах для людей или животных. Модули LPWAN должны иметь низкую стоимость, то же касается и их подключения. Это поможет IoT-рынку расти. [4]

Актуальная задача на сегодняшний день – связать эти три, казалось бы, несочетаемые технологии в единое целое, что позволит обеспечить надежное, стабильное и защищенное хранение и передачу данных в среде интернета вещей.

На данный момент блокчейн технологии широко используются в сфере цифровых активов – криптовалют. Ярким примером цифрового актива на рынке криптовалют

является “Биткойн”. Сеть цифрового актива использует систему распределенного реестра, цифровые подписи, временные метки и смарт-контракты.

С помощью технологии блокчейн мы можем решить задачу безопасного администрирования сети, исключая хакерские атаки типа “человек-посередине”, тем самым снимая проблему главной точки отказа – “единого администратора” [5]. В тоже время цифровые сертификаты позволяют осуществлять авторизацию специалистов в системе управления устройствами.

Передавая IoT радиосигналы в режиме “точка-точка”, мы имеем уязвимость в виде подделки передаваемых данных взломанным устройством, не имея на руках закрытые ключи для подписи транзакций, злоумышленник не сможет компрометировать работу IoT устройства [6]. Блокчейн технология позволяет использовать закрытые и открытые ключи для подписей осуществляемых транзакций, тем самым устройства подтверждают “валидность” передаваемой информации. Стоит отметить что защита закрытого ключа на устройства должна осуществляться аппаратно, на данный момент крупные производители кремниевых чипов, имеют соответствующие решения, не поддающиеся даже физическому взлому [7].

Фиксация времени играет крайне важную роль в передаче информации и фиксации происходящих событий, замеров с датчиков IoT устройств. На текущий момент стандарты времени и частоты, располагающиеся в инфраструктуре сервера, уязвимы к атакам, при дестабилизации работы устройства, вся инфраструктура и связанные с ней внешние сервисы, прекращают свою работу. Блокчейн решает проблему фиксации времени с помощью “временных меток”, подключая стандарт времени и частоты к сети блокчейна, на основании подписей и сертификатов мы можем быть уверены в том, что он является надёжным источником.

IoT системы имеют необходимость в безопасной передаче команд на устройство, для корректировки его настроек и типа передаваемой информации. Даже если среда передачи была скомпрометирована. блокчейн предлагает решение в виде смарт-контракта, специальной двухсторонней цифровой транзакции. Смарт-контракт позволяет нам определить, получателя и отправителя команд, метку времени и цифровые подписи обеих сторон, после исполнения, запись сохраняется в блокчейне, данная информация не подлежит изменению в реестре и является основной транзакцией блокчейна [8].

Контроль доступа к информации. Блокчейн технологии позволяют контролировать доступ к информации путем программного добавления уполномоченных лиц, на основе закрытых/открытых ключей и записи их в общий реестр. Контроль доступа осуществляется всеми членами сети, а несанкционированный доступ к информации путем взлома конечной точки как уже упоминалось ранее является невозможным.

Хэширование информации. Для проверки целостности в блокчейне могут использоваться различные алгоритмы хэширования. На данный момент, распространённым алгоритмом является SHA-2, однако недавние работы показали критические уязвимости в алгоритме, поэтому на замену устаревшему алгоритму, придёт новый – SHA-3.

Главные проблемы безопасности, без сомнения, связаны с контролем доступа и незащищенными сервисами. Поставщики могут способствовать безопасному использованию своих продуктов, предоставляя документацию и взаимодействуя с пользователями и специалистами по безопасности. Чтобы усложнить задачу злоумышленникам, устройства должны быть физически защищены. Наконец, если устройство скомпрометировано, оно должно отклонить программы, предоставленные злоумышленником, и уведомить своего пользователя о том, что что-то не так.

Использование блокчейна как инфраструктуры для современных сетей является вполне естественным шагом развития в сфере ИБ. Блокчейн как технология только набирает обороты, использование её в сетях IoT является лишь одним из многих решений для применения. Стоит отметить что IoT также является новой технологией, но уже нашла

применение во многих областях, однако и она имеет существенные проблемы безопасности, которые без особых усилий решаются блокчейном. Объединив обе технологии, мы получим практически совершенную технологию для автоматизации процессов во многих сферах деятельности человека.

#### Список литературы

- [1] Интернет вещей и его перспективы развития в Казахстане 2017 // URL: [https://forbes.kz/process/internet/что\\_такое\\_internet\\_вещей\\_и\\_каковы\\_перспективы\\_его\\_развития\\_в\\_казахстане](https://forbes.kz/process/internet/что_такое_internet_вещей_и_каковы_перспективы_его_развития_в_казахстане)
- [2] Блокчейн: Схема новой экономики / Мелани Свон : [перевод с английского]. — Москва : Издательство «Олимп–Бизнес» 2017 // URL: <https://kniga.biz.ua/pdf/5351-Blockchain.pdf>
- [3] А.В. Росляков, С.В. Ваняшин, .Ю. Гребешков ИНТЕРНЕТ ВЕЩЕЙ 2015 // URL: <https://iotas.ru/files/documents/internet%20of%20things>
- [4] LPWAN и интернет вещей 2020 // <http://1234g.ru/novosti/standart-lpwa-dlya-interneta-veshchej>
- [5] Blockchain-Based Man-in-the-Middle (MITM) Attack Detection for Photovoltaic Systems 2020 // URL: <https://ieeexplore.ieee.org/document/9529949>
- [6] IoT security: challenges and countermeasures 2020 // URL: <https://www.sciencedirect.com/science/article/pii/S1877050920323395>
- [7] Case Study: On the Security of Key Storage on PCs 2013 // URL: [https://www.academia.edu/11729584/Case\\_Study\\_On\\_the\\_Security\\_of\\_Key\\_Storage\\_on\\_PCs](https://www.academia.edu/11729584/Case_Study_On_the_Security_of_Key_Storage_on_PCs)
- [8] Blockchain smart contracts: Applications, challenges, and future trends 2021 // URL: <https://link.springer.com/article/10.1007/s12083-021-01127-0>



# ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ РАБОТЫ ИНФОРМАЦИОННОЙ СИСТЕМЫ МОНИТОРИНГА ИНЦИДЕНТОВ БЕЗОПАСНОСТИ В БАНКОВСКОЙ ДЕЯТЕЛЬНОСТИ

Омирбекова Ильфа Сержановна

Магистрант 2 курса специальности «Информационные системы»

Научный руководитель: к.п.н., ст.преподаватель Баймулдина Н.С.

Казахский национальный университет им. Аль-Фараби

г. Алматы, Республика Казахстан

email: omerbekova.ilfa@gmail.com

**Аннотация:** Основывается на необходимости разработки алгоритма и программного обеспечения для системы прогноза конфликтов информационной защищенности, позволяющих ей расценивать воздействие определенных моментов на эффективность прогноза и на обеспечение безопасности секретных данных в банке.

**Ключевые слова:** банк, информационная безопасность, эффективность, мониторинг, инцидент.

## 1. Введение

Мониторинг инцидентов информационной безопасности (МИИБ) содержит собственную цель выявления всех конфликтов, происходящих в компьютерной сети банка. Впрочем, в мощь воздействия всевозможных моментов, настоящий итог прогноза имеет возможность выделяться от желанного и собственно чем меньше отличие между реальным и нужным итогом, тем прогноз эффективнее. Для увеличения производительности МИИБ нужно вначале обнаружить моменты, а после этого, при необходимости убрать их плохое воздействие.

## 2. Методы исследования

Структура мониторинга инцидентов информационной безопасности:

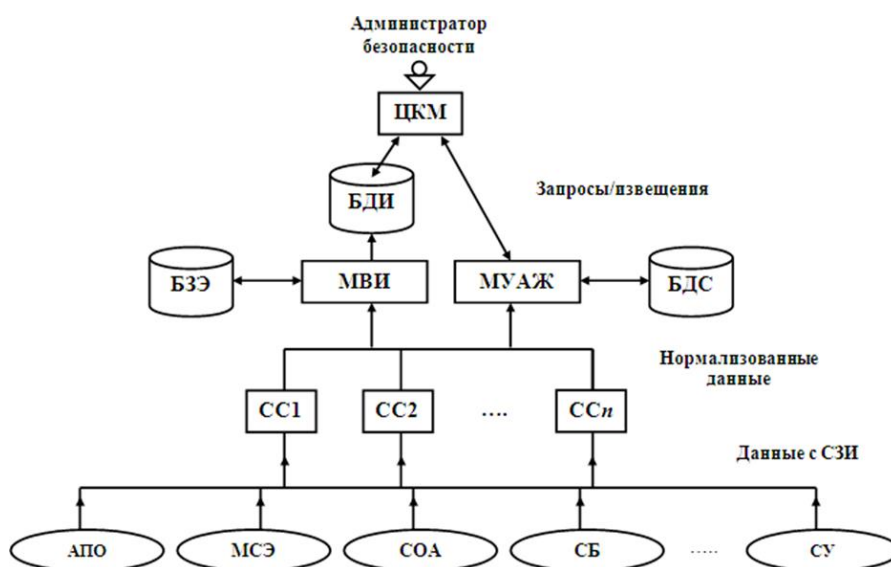


Рис 1. Структура МИИБ

Ведущие компоненты системы:

– центральная консоль мониторинга (ЦКМ), предназначенная для управления и опции системы МИИБ, отражения данных о состоянии информационной безопасности (ИБ) в автоматизированной банковской системе (АБС), вывода сообщений об инцидентах и выдачи назначений по их устранению [1];

– база данных инцидентов (БДИ);

– база данных событий (БДС), служащая для сбережения всех поступивших от систем защиты информации (СЗИ) мероприятий информационной безопасности (ИБ), что в количестве не являющихся инцидентами;

– база знаний экспертов (БЗЭ), имеющая экспертные познания, применяемые МВИ при выявлении конфликтов информационной безопасности (ИБ);

– модуль выявления инцидентов (МВИ) – автоматизированное ядро системы, коррелирующее и анализирующее информацию, поступающую от систем защиты информации (СЗИ);

– модуль управления архивными журналами (МУАЖ), отвечающий за запись мероприятий в БДС, а еще за их извлечение по запросу с ЦКМ;

– сборщики мероприятий (СС1, СС2,...,ССn ), которые применяются для нормализации данных, приобретенных от СЗИ, приведения их к единственному формату, применяемому МВИ [2].

При нижнем уровне системы имеют все шансы находится всевозможные программно- аппаратные способы обороны данных (СЗИ), собирающие данные об инцидентах информационной безопасности (ИБ):

– АПО – антивирусное ПО;

– МСЭ – межсетевые экраны;

– СОА – системы обнаружения атак;

– СБ – серверы безопасности;

– СУ – сетевые приборы (маршрутизаторы, концентраторы и т.д.).

В последующем под мероприятием информационной защищенности (СИБ), сообразно, понимается идентифицированное возникновение конкретного состояния системы, обслуживания или же сети, которое говорит или о вероятном несоблюдении политической деятельности СИБ или же отказе защитных мер, или до этого неведомой истории, которая имеет возможность владеть отношением к защищенности. [3] В соответствии с этим, инцидент информационной безопасности (ИИБ) – это возникновение 1-го или же нескольких ненужных или же внезапных СИБ, с которыми связана важная возможность сотворения опасности ИБ.

### **3. Обсуждение**

При оперативном выявлении конфликтов система МИИБ за применимое время обрабатывает важный размер данных о СИБ, получаемых от СЗИ, при этом на процесс выявления между ними их конфликтов оказывают воздействие характеристики как программного и аппаратного обеспечения всевозможных компонентов системы [4].

В большей степени на итог прогноза воздействуют надлежащие моменты:

1. Своевременность обнаружения инцидента;

2. Наличие и вероятность пополнения познаний о вероятных инцидентах.

Разглядим 1-ый из данных моментов, т.е. своевременность обнаружения конфликта. Он ориентируется периодом, в направление которого не будет нанесен вещественный или же репутационный вред для банка, связанный с запаздыванием в адекватном реагировании на сообразный инцидент [5]. При данном явлении нужно выделять время поступления данных с СЗИ на обработку МВИ и время их обработки в МВИ.

Время поступления данных с СЗИ на обработку находится в зависимости, ключевым образом, от данных определенных СЗИ и каналов передачи данных:

– исправности СЗИ (сбои имеют все шансы быть вызваны как физиологическими или же программными отказами приборов, например и намеренными деяниями злоумышленников);

– корректности опции СЗИ (неправильно настроенное СЗИ регистрирует важный размер неверных срабатываний, т.е. обнаруживает СИБ, когда их на самом деле нет, собственно что приводит к увеличенной нагрузке на каналы передачи данных и МВИ, а еще к игнорированию персоналом инфы о СИБ, приобретенной от предоставленного

СЗИ)[6];

– пропускной возможности каналов передачи данных (системы МИИБ использует для передачи данных о СИБ и ИИБ те же телекоммуникационные каналы, собственно что и составляющие АБС, в следствие этого недостающая пропускная дееспособность каналов способна замедлить работу не лишь только системы МИИБ, но и самой АБС) [7].

Время обработки данных, поступивших в МВИ с СЗИ находится в зависимости от:

– исправности МВИ (физический или же программный отказ МВИ способен не сблности процесс прогноза, вплоть до совершенного остановки обработки поступающих данных о СИБ);

– присутствия очереди из извещений о СИБ (вновь пришедшие на обработку сообщения о СИБ имеют все шансы оказаться сообщениями об ИИБ, собственно что имеет возможность быть выявлено очень поздно для предотвращения вероятного ущерба).

Дальше разглядим 2 из упомянутых моментов, в большей степени влияющих на итог прогноза. Речь идет о наличии и способности пополнения познаний, важных для выявления ИИБ из большого количества извещений о СИБ, поступивших в МВИ на обработку. При данном случае стоит говорить о сигнатурах – особых симптомах и свойствах определенных ИИБ и скопленном персоналом эксперименте по выявлению ИИБ [8].

Присутствие и пополнение сигнатур в системе МИИБ находится в зависимости от:

– существования сигнатур для определенных ИИБ (появление сигнатур для раньше неведомых ИИБ занимает кое-какое время, в направление которого обнаружить аналогичный конфликт ИБ невозможно) [9];

– исправности процедуры обновления сигнатур (ошибки в настройке или же сбой программно-аппаратных компонент системы МИИБ имеют все шансы приводить к нарушениям в процессе обновления);

– единства сигнатур (в случае конфигурации злоумышленником или же вредным ходом содержимого сигнатур, хранящихся на строгих дисках или же в оперативной памяти использующих их СЗИ, система МИИБ не может обнаружить надлежащие ИИБ).

Накапливаемый персоналом в процессе эксплуатации АБС навык (экспертные познания по идентификации свежих, неведомых раньше конфликтов, для коих сигнатурное выявление было невозможно), реализуется в системе МИИБ при:

– наличии в ней подсистем получения и скопления знаний;

– наличии когнитологов - знатоков по познаниям, способных извлечь нужные познания из полученной от профессионалов инфы и избрать подходящую форму ее представления в БЗЭ [10].

В процессе эксплуатации АБС банка админ защищенности обязан периодически расценивать воздействие, оказываемое разными причинами на эффективность МИИБ, а это значит, и на обеспечение совместной безопасности секретных данных, собственно что настоятельно просит конкретных умственных усилий и времени [11]. При надобности, в случае если оказываемое моментом воздействие плохое, т.е. снижающее эффективность прогноза (например, по причине сбоя в МСЭ в систему МИИБ не станут попадать сообщения о сканировании портов), админ защищенности совершает корректирующие воздействия, приводящие к восстановлению обычного функционирования системы МИИБ (например, перенастраивает ПО МСЭ).

#### **4. Заключение**

Таким образом рекомендуется в больших банках требуются роли нескольких админов защищенности, потому что существует большое наличие численности СЗИ, для которых нужно проводить похожую оценку.[5] На наш взор, в этих случаях, видится целесообразным исполнять данную оценку автоматизировано, собственно, что несомненно поможет в значимой степени разгрузить работу админов защищенности и уменьшить их численность. Нужно создать методы оценки и отвечающее ПО для системы МИИБ, которые дозволили бы ей исполнять надлежащие функции:

- контроль единства и доступности собственных компонентов;
- оценка корректности опции данных компонентов;
- резервирование данных;
- выдача назначений админу защищенности по принятию решений;
- совершение под корректирующих поступков (автоматически или же по команде админа безопасности), приводящих к восстановлению обычного функционирования системы.

#### Список литературы

1. Кумар, Прадип и Мохаммад Хан. «Системы, методы и машиночитаемые носители для передачи платежных и неплатежных виртуальных карт между мобильными устройствами», патент США 10 026 076, выданный 17 июля 2018 г.
2. Акхисар, Ильяс, К. Бату Тунай и Некла Тунай, «Влияние инноваций на деятельность банка: пример электронных банковских услуг», *Procedia-Social and Behavioral Sciences*, no. 195, стр. 369-375, 2015.
3. Теохари, Калифорния, Использование Интернета террористами: информационные операции в киберпространстве, Diane Publishing, 2011.
4. Гроуди, Аллан Д., Фотиос Харманцис и Грегори Дж. Капл, «Операционный риск и справочные данные: изучение затрат, требований к капиталу и снижения рисков», *Требования к капиталу и снижение рисков*, том. 1, нет. 3, стр. 130-187, 2006.
5. Бамбергер, Кеннет А., «Технологии соответствия: риск и регулирование в эпоху цифровых технологий», *Tech. L. Rev.*, № 88, стр. 669, 2009 г.
6. Нами, Мохаммад Реза, «Электронный банкинг: проблемы и проблемы», 10-я Международная конференция ACIS по программной инженерии, искусственному интеллекту, сетям и параллельным/ распределенным вычислениям, 2009 г., стр. 263-266. 2009.
7. Джерман-Блажич, Борка, «Подход к экономическому моделированию управления рисками информационной безопасности», *Международный журнал управления информацией*, том. 28, нет. 5. С. 413-422, 2008.
8. Ангелакопулос, Георгиос и Афанассиос Михиотис, «Электронный банкинг: проблемы и возможности в греческом банковском секторе», *Исследование электронной коммерции*, том. 11, нет. 3, стр. 297-319, 2011.
9. Нитсуре, Рупа Рече, «Электронный банкинг: вызовы и возможности», *Еженедельник по экономическим и политическим вопросам*, том. 38, нет. 51, стр. 5377-5381, 2003.
10. Ахмад, Алаэддин Мохд Халаф и Хасан Али Аль-Зуби, «Функциональность электронного банкинга и результаты удовлетворенности клиентов: эмпирическое исследование», *Международный журнал маркетинговых исследований*, том 3, № 1, стр. 50-65, 2011.
11. О, Юнсанг и Такаши Оби, «Идентификация угроз фишинга в государственных веб-службах», *Международный журнал информации и сетевой безопасности*, том. 2, нет. 1, стр. 32, 2013.

## АҚПАРАТТЫҚ ҚАУІПСІЗДІКТІҢ ҚАТЕРЛЕР ИНЦИДЕНТІН МОДЕЛЬДЕУ

Муратхан Р., Туткуше А.Е.

Академик Е. А. Бөкетов атындағы Қарағанды университеті, Қарағанды, Қазақстан

E-mail: [raikhan.muratkhan@mail.ksu.kz](mailto:raikhan.muratkhan@mail.ksu.kz)

Ақпараттық қауіпсіздіктің стандартты жобалауы бизнес аналитиктің (бизнес саласының маманы) және қауіпсіздік аналитигінің (қауіпсіздік облысының маманы) тығыз байланыста жұмыс істеуін талап етеді. Бизнес аналитиктер тек ғана бизнестік көрсеткіштердің жақсаруына ғана көңіл бөлетіндіктен, дер кезінде жасалған қауіпсіздіктің талдауы, жеткілікті деңгейде қауіпсіздік талаптарын қанғаттандырмайтын жобаларды анықтауға және оларды жүйеден алып тастауға мүмкіндік береді.

Бизнес процестерді графикалық көрсету үшін әртүрлі модельдеу тілдері қолданылады. Біз MisuseCases [1, 2] модельдеу тілін таңдап алдық. Өйткені ол семантикалық модельмен тығыз байланысты. Яғни әрбір форманың анықталған бір мәні бар және ол объектіге нақты бір ережелерді қоса алады. MisuseCases тілінде қауіпсіздіктің қажетті талаптарын анықтауға мүмкіндік беретін, қауіпсіздікті қамтамасыз ету процесі бар. Қауіпсіздікті қамтамасыз ету процесі келесі бес қадамнан тұрады [1]:


- 1) критикалық активтерді анықтау;
- 2) қауіпсіздік мақсаттарын анықтау;
- 3) қатерлерді анықтау;
- 4) тәуекелдерді анықтау және талдау;
- 5) қауіпсіздік талаптарын анықтау.

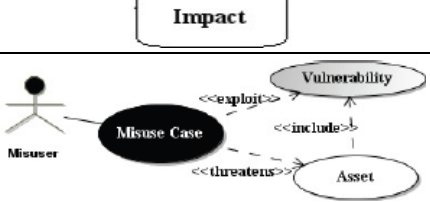




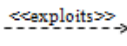
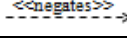
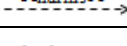
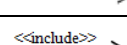

ISSRM онтологиялық моделінің [3] түсініктерін графикалық кескіндеу үшін Misuse Case нотациясының синтаксисі (кесте 1-3) қолданылған [4, 5]. Бұл кестелерде бірінші бағанда ISSRM моделіндегі түсініктер, екінші бағанда олардың басқа әдебиеттерде қолданылған синонимдері және үшінші бағанда түсініктер мен қатынастарды анықтайды. Ал соңғы бағанда Misuse Case нотациясының графикалық синтаксисі берілген.

Кесте 1 – Активке байланысты түсініктер

ISSRM моделіндегі түсінік	Синонимдері	MisuseCase-тегі көрінісі
Актив	Құнды ресурс	 actor
Бизнес актив	Қызметтік қолдану жағдайы	
АЖ активі	АЖ қолдану жағдайы	
Қауіпсіздік критерийі	Қауіпсіздік шектеулері	
Қолдау	Енгізу	
Шектеу	Шектеу	

Кесте 2 – Тәуекелмен байланысты түсініктер

ISSRM моделіндегі түсінік	Синонимдері	Белгіленуі	MisuseCase-тегі көрінісі
1	2	3	4
Тәуекел	Қауіп-қатер	R	

Ықпал ету	Шығын	AV	
Оқиға	Инцидент		
Шабуыл әдісі	Зорлық зомбылық (Violence)	-	
Осалдық	Әлсіз жері	V	
Қатер бастауы	Шабуылдаушы		
Қатер	Қауіп-қатер	$P(T)$	
Пайдалану	Пайдалану		
Мойындамау	Мойындамау		
Шығын	Шығын		
Алып келеді	Алып келеді		
Ерекшеліктері	Енгізу		

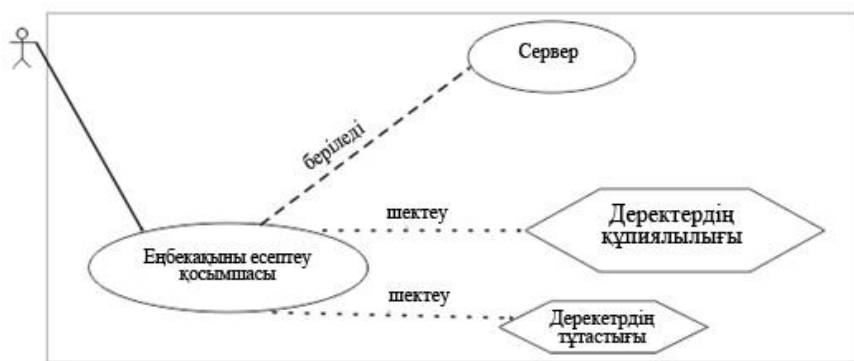
Кесте 3 – Тәуекелді өңдеуге байланысты түсініктер

ISSRM моделіндегі түсінік	Синонимдері	MisuseCase-тегі көрінісі
Тәуекелді өңдеу		
Қорғаныс талаптары	Контршара	
Бақылау		-
Тазалау		-
Жұмсарту	Азайту	

АҚ бағалаудың мысалын қарастырайық.

Қызметкердің еңбекақысы туралы дерек еңбекақыны есептеу қосымшасы арқылы орталықтан қолданушыға (қызметкерге) жіберіледі.

Жіберілетін деректердің екі қауіпсіздік критерийі бар: құпиялылық және тұтастық. Яғни, жіберілетін деректер құпия және адресатқа тұтас күйінде жету керек. 1- суретте қауіпсіздік критерийілері алтыбұрышпен белгіленген.



Сурет 1 – Активке байланысты түсініктерді модельдеу



## ХЕШ-ФУНКЦИЯ ҰҒЫМЫ ЖӘНЕ БАҒДАРЛАМАЛЫҚ КӨРІНІСІ

*Қинойт Дина Дидарқызы  
Ақпаратты қорғаудың математикалық әдістері  
мамандығының 4-ші курс студенті  
Дәулет Серікбаев атындағы  
Шығыс Қазақстан Техникалық Университеті  
Өскемен, Қазақстан*

**Аннотация:** Мақала хеш-функциялар және SHA алгоритмі туралы түсініктеме, және хешингті жай ғана сөзбен түсіндіріп қоймай, бағдарламалық түрде жүзеге асыралыған нұсқасын көрсетіледі.

**Кілттік сөздер:** хеш, хешинг, хеш-функция, SHA, SHA-1, SHA-256, SHA-384, SHA-512, ақпараттық қауіпсіздік, C#, бағдарлама.

Ақпараттық технологиялар қарқынды түрде дамып келе жатқан бүгінгі күні деректерді қорғаудың күрделі әдістерінің талап етілетіні айтпаса да белгілі. Осы ретте криптографиялық әдістерді қолдану дұрыс шешім болып табылатыны анық. Сол әдістердің бірі ретінде хеш-функцияларды айтуға болады.

Хэш немесе хэш-функциясы қазіргі криптографияның және блокчейн алгоритмінің негізгі компоненттерінің бірі болып саналады. Хэшинг дегеніміз – ақпараттың кез келген көлемін кіріс ақпараттың осы массивіне ғана тән бірегей таңбалар жиынына түрлендіру болып табылса, ал бұл таңбалар жинағы хэш деп аталады.

Хэш-функция немесе жинақтау функциясы дегеніміз – белгілі бір алгоритммен орындалатын еркін ұзындықтағы кіріс деректерінің массивін белгіленген ұзындықтағы шығыс бит жолына түрлендіруді жүзеге асыратын функция болып табылады. Хэш-функциясы арқылы жасалған түрлендіру хэшинг деп аталады. Бастапқы ақпарат кіріс массиві, "кілт" немесе "хабарлама" деп, ал түрлендіру нәтижесі "хэш", "хэш-код", "хэш сомасы", "хабарламаның қысқаша мазмұны" деп аталады.

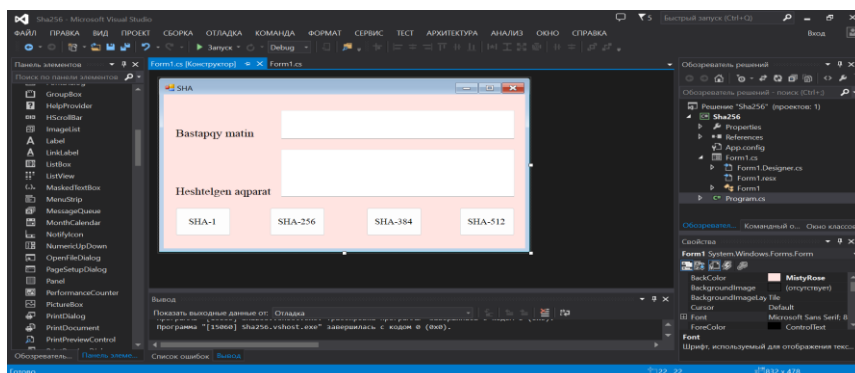
Secure Hash Algorithm – криптографиялық хэш алгоритмі. Көптеген криптографиялық қосымшалар мен хаттамаларда қолданылады. Сондай-ақ, АҚШ-тағы мемлекеттік мекемелер үшін негізгі ретінде ұсынылған. SHA-1 негізіндегі принциптер Рональд Ривесттің MD4 дизайнында қолданған принциптеріне ұқсас. Хэш функциялары нұсқаларды басқару жүйелерінде, электрондық қолтаңба жүйелерінде, сондай-ақ аутентификация кодтарын құру үшін қолданылады.

SHA-1 бүкіл SHA отбасының ең кең тараған түрі болып табылады және кеңінен қолданылатын әртүрлі криптографиялық қолданбалар мен алгоритмдерде пайдаланылады.

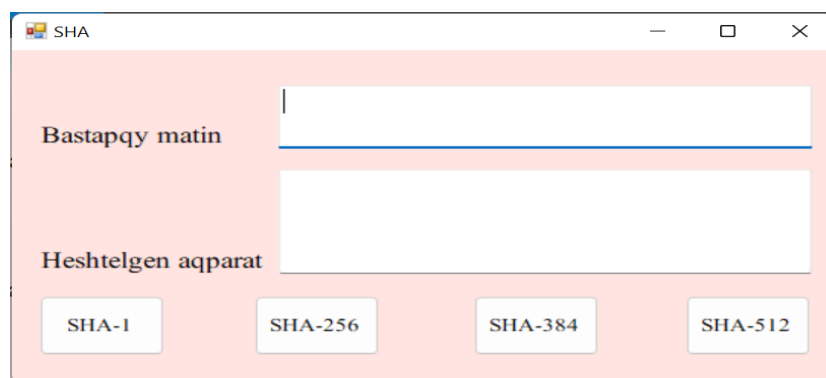
Хэш-функцияға арналған қосымшаны жасап, хештеуді жүзеге асыру үшін C# бағдарламалау тілі таңдалды. C# – объектіге бағытталған бағдарламалау тілі. Бұл тілдің артықшылығы – оның объектіге бағытталған бағдарламалау тәсілін пайдаланатынында. Демек, бастапқыда абстрактілі конструкцияларды сипаттауды жүзеге асырып, содан кейін ғана олардың арасында өзара іс-қимылды орындау қажет. Осы әдіс танымалдылыққа ие, және де бағдарламашының ауыр өмірін жеңілдітеді деп айтуға болды.

C# объектіге бағытталған программалау тілінде SHA хэш-функцияларының алгоритмін қолдана отырып хештеу операциясын жүргіземін. Бағдарламаны жасау үшін label, textbox және button басқару элементтерін қолданамын.



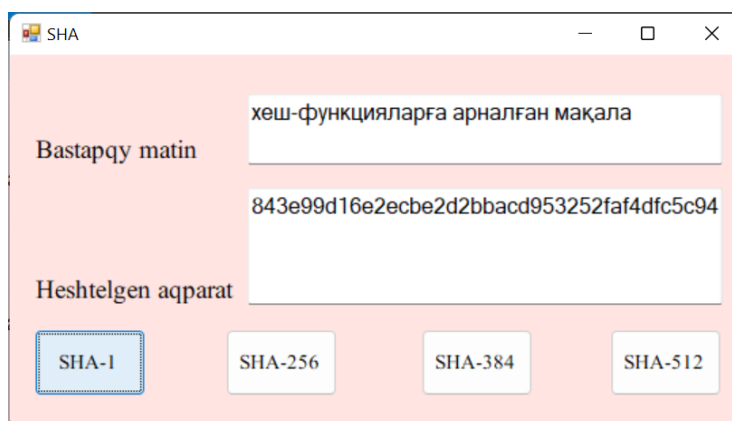


Сурет 1 - Бағдарламаны қосқан кездегі көрініс.  
Бағдарламаның орындалу барысы.



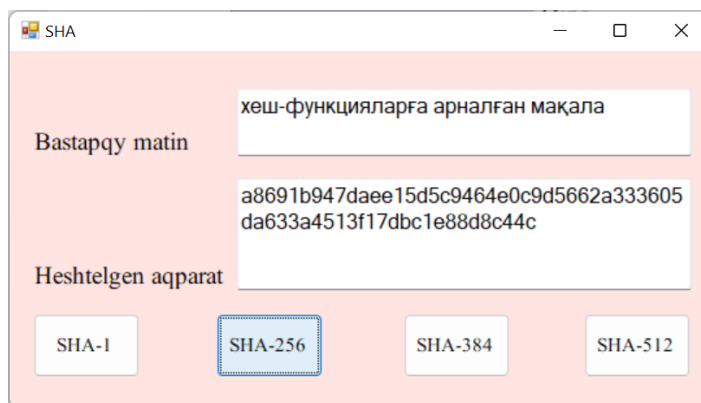
Сурет 2 - Бағдарлама терезесі

Бағдарламаны қосқаннан кейін сурет 2-де көрсетілгендей терезе ашылады. Сол терезеде көрсетілген “Bastapqy matın” сөзінің жанындағы ұяшыққа сурет 3-те көрсетілгендей тиісті, хештелуді керек бастапқы мәтінді, яғни ақпаратты енгіземін.



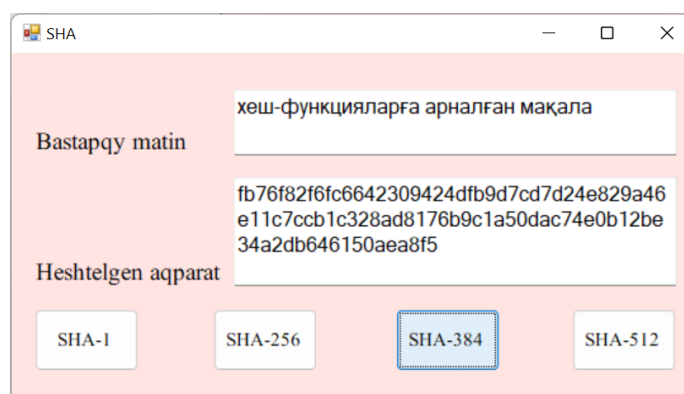
Сурет 3 – SHA-1 алгоритмі арқылы хештеу

“SHA-1” батырмасын басқаннан кейін, сурет 3-те көрсетілгендей тиісті жолда SHA-1 алгоритмі арқылы хештелген ақпарат шығады.



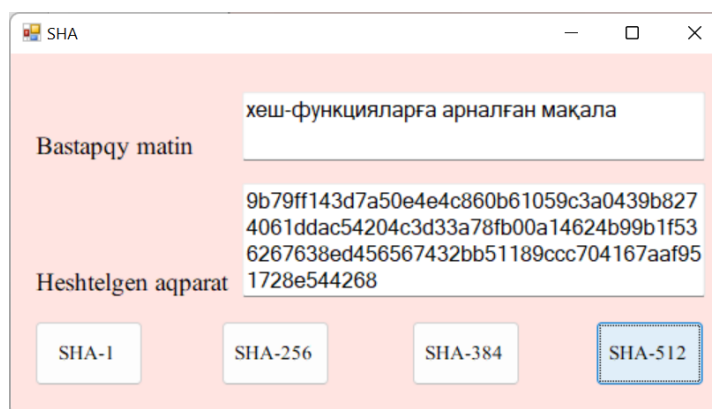
Сурет 4 - SHA-256 алгоритмі арқылы хештеу

“SHA-256” батырмасын басқаннан кейін, сурет 4-те көрсетілгендей тиісті жолда SHA-256 алгоритмі арқылы хештелген ақпарат шығады.



Сурет 5 - SHA-384 алгоритмі арқылы хештеу

“SHA-384” батырмасын басқаннан кейін, сурет 5-те көрсетілгендей тиісті жолда SHA-384 алгоритмі арқылы хештелген ақпарат шығады.



Сурет 6 - SHA-512 алгоритмі арқылы хештеу

“SHA-512” батырмасын басқаннан кейін, сурет 6-да көрсетілгендей тиісті жолда SHA-512 алгоритмі арқылы хештелген ақпарат шығады.

Қорытындылай келе, мақалада хеш-функциялар жайлы сөз қозғала отырып, C# тілін қолдану арқылы жасалған қосымшада SHA алгоритмдерінде хештеудің қалай жүзеге асырылатынына бағдарламалық түрде көрініс берілді, демек мақсат орындалды. Осы ретте айта кететін жайт – бұл хеш-функцияларының ақпараттық қауіпсіздік саласына қосатын үлесінің айтарлықтай көп екені белгілі, және бұл әдіс болашақта да аса қолданыста болатыны анық!

## Пайданылған әдебиеттер тізімі

1. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си = Applied Cryptography. Protocols, Algorithms and Source Code in C. — М.: Триумф, 2002. — 816 с.
2. Нильс Фергюсон, Брюс Шнайер. Практическая криптография = Practical Cryptography: Designing and Implementing Secure Cryptographic Systems. — М.: Диалектика, 2004. — 432 с.
3. Брюс Шнайер. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си. — М.: Триумф, 2002.

## ТЕЗАУРУСЫ И ОНТОЛОГИИ В НАУЧНО-ОБРАЗОВАТЕЛЬНЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ

*Самбетбаева М.А. Баймырзаева Г.Ж.*

*Евразийский национальный университет имени Л.Н. Гумилева, г. Астана,*

*Таразский региональный университет имени М.Х. Дулати, г.Тараз*

*E-mail: [madina\\_jgtu@mail.ru](mailto:madina_jgtu@mail.ru), [gaukhar.baymyrzaeva@mail.ru](mailto:gaukhar.baymyrzaeva@mail.ru)*

Одним из основных результатов созидательной, социальной и интеллектуальной человеческой деятельности является создание и накопление информационных ресурсов с целью их дальнейшего использования и недопущения утраты опыта предыдущих поколений.

Не будет преувеличением сказать, что уровень развития технологий накопления информации и эффективности использования накопленной ранее информации на протяжении всей истории человечества значительно влиял на уровень развития производительных сил. Утеря информации приводила к отбрасыванию цивилизации на века назад.

Однако, чтобы эффективно пользоваться накопленной ранее информацией, необходимы специальные инструменты и специальные технологии, при помощи которых могут быть реализованы специальные приемы работы с информацией.

Проблема поиска информации - одна из вечных проблем человечества. Чтобы решить проблему доступа к информации, человечество создало библиотеки - универсальную систему хранения, систематизации и каталогизации «информации и знаний».

В процессе научной, а особенно образовательной, деятельности много времени и сил отнимает работа с литературными источниками, разного рода материалами и документами: поиск необходимых документов, систематизация и классификация документов в соответствии с поставленной задачей.

Для удовлетворения информационных потребностей современных пользователей необходима поддержка сложных функций поиска и классификации информации, а также просмотр ресурсов по категориям (рубрикам) и словарям-классификаторам.

Наиболее важной задачей является задача систематизации ресурсов, для решения которой необходимо четко определить состав логико-семантических категорий (фасетов) и ключевых терминов (понятий), покрывающих избранную достаточно узкую предметную область, интересующую пользователя. Как правило предметная область ограничивается изучаемым учебным курсом или конкретной темой курса.

Стандартным подходом к систематизации информации является классификация документов с помощью таксономий.

Таксономия – это предметная классификация, которая группирует термины в виде управляемого словаря (тезауруса) и упорядочивает их (словари) в виде иерархических структур.

Основу классификации составляет выделение понятий (ключевых терминов), установление парадигматических отношений (например, типа родитель – потомок) между ними и сопоставление анализируемого документа выделенным понятиям.

Для описания какой-либо предметной области обычно используется определенный набор ключевых терминов, каждый из которых обозначает или описывает какое-либо понятие из данной предметной области.

По определению международной организации по стандартизации (ISO), тезаурус является словарем, управляемым языком индексации, формально организованным для того, чтобы установить явные априорные отношения между понятиями.

Индексация (Индексирование) — это сортировка (систематизация) информации по каким-либо формальным принципам. Различают два вида индексирования: классификационное и координатное.

Это определение устанавливает лексические единицы (термины) и семантические отношения между этими единицами в качестве элементов, составляющих тезаурус.

Тезаурусные (парадигматические) отношения (род-вид, часть-целое, комплекс-элемент, причина-следствие) налагаются на структуру таксономии, то есть идентифицируются основные таксономии предметной области.

Тезаурус это словарь (идеографический или семантический словарь), в котором слова (лексические единицы, термины, понятия, дескрипторы) размещаются их по смысловой близости.

Словари, в которых лексика располагается не алфавитном порядке, а на основании этого критерия, получили название идеографических (от греч. *idea* — понятие, идея, образ и *grapho* — пишу) или семантических (от др.-греч. *σημαντικός* — обозначающий) — раздел лингвистики, изучающий смысловое значение единиц языка).

Тезаурус является инструментом концептуального (понятийного) описания отдельных предметных областей.

В отличие от толкового словаря, тезаурус определяет понятия (ключевые термины) не только с помощью описания, данного в словарной статье, но и посредством соотнесения понятия с другими понятиями и их группами, благодаря чему может использоваться как система знаний, отраженных языком (словами), тогда тезаурус становится интересным сам по себе, а не только как вспомогательный инструмент.

#### **Основные отношения в тезаурусе**

Синонимия — связь между терминами, различными по написанию, но имеющих одинаковое или очень близкое понятийное (лексическое) значение, например: кавалерия — конница, смелый — храбрый.

Антонимия — связь между терминами, различными по написанию, имеющих прямо противоположные значения: правда — ложь, добрый — злой.

Гипонимия/гиперонимия Гипероним — термин (понятие, слово) с более широким значением, выражающее общее, родовое понятие, название класса (множества) предметов (свойств, признаков). Гипоним — термин с более узким значением, называющее предмет (свойство, признак) как элемент класса (множества). Эти отношения транзитивны и несимметричны. Гипоним наследует все свойства гиперонима.

Меронимия/партонимия — отношение типа «часть-целое», с выделенными отношениями «быть элементом» и «быть сделанным из».

Для описания какой-либо предметной области всегда используется определенный набор терминов, каждый из которых обозначает или описывает какое-либо понятие или концепцию из данной предметной области. Совокупность терминов, описывающих данную предметную область, с указанием семантических отношений (связей) между ними является тезаурусом. Такие отношения в тезаурусе всегда указывают на наличие смысловой (семантической) связи между терминами.

Основным отношением (связью) между терминами в тезаурусе является связь между более широкими (более выразительными) и более узкими (более специализированными) понятиями. Часто выделяют 2 подвида этого отношения:

- Один термин обозначает понятие, являющееся частью понятия, обозначаемого другим термином (например, «наука» и «математика», «математика» и «теория чисел»).
- Один термин обозначает элемент класса, обозначаемого другим термином («горные районы» и «Кавказ»).

Существуют также и другие связи между терминами. Например, одно понятие или концепция может быть обозначено несколькими терминами, являющимися синонимами. Некоторые термины могут быть антонимами для других. Часто среди терминов, относящихся к одному понятию, выделяют единственный (для каждого языка тезауруса)

наиболее предпочтительный (наиболее подходящий) термин, который наиболее хорошо характеризует, или обозначает данное понятие. Остальные термины являются менее предпочтительными (менее подходящими).

Помимо вышеописанных, между терминами могут существовать также и другие, ассоциативные связи, если понятия, обозначаемые этими терминами, как-либо связаны между собою по своему смыслу, за исключением описанных выше иерархических связей.

#### **Тезаурусы в описании информации**

Термины могут иметь следующие атрибуты (основные):

- ID – Identifier. Уникальный идентификатор термина.

- SN – Scope Note. Комментарий к термину. Например, представляет вербальное пояснение термина, или правила его использования.

- TT – Top Term. Признак, выделяющий термины на самом верхнем уровне иерархии (термины наиболее общих понятий в данной иерархии понятий).

- HN – History Note. История модификации связей и атрибутов данного термина.

Термины могут иметь следующие отношения (основные):

- USE – Связывает термин с наиболее предпочтительным (на том же языке) термином для данного понятия.

- UF – Used For. Обращение связи USE. Связывает наиболее подходящий термин с синонимами и квазисинонимами (менее подходящими терминами).

- BT – Broader Term. Связь термина с термином более общего понятия.

- NT – Narrower Term. Обращение связи BT.

- RT – Related Term. Ассоциативная связь. Связывает семантически связанные между собою термины, не находящиеся при этом в одной иерархии, и не являющиеся синонимами или квазисинонимами.

- LE – лексический эквивалент термина (на другом языке)

#### **Тезаурусы в информационном мире**

Следует подчеркнуть, что первые тезаурусы составлялись без всякой связи с особенностями информационной деятельности, они были органически связаны с фундаментальными проблемами познания, отображая представление о мире в целом и закономерностях его постижения средствами естественного языка.

Тезаурусы в информационном поиске

Коллежский советник Семён Николаевич Корсаков (1787—1853) ставил задачу усиления возможностей разума. В 1832 году С. Н. Корсаков опубликовал описание пяти изобретённых им механических устройств, так называемых «интеллектуальных машин», для частичной механизации умственной деятельности в задачах поиска, сравнения и классификации. В конструкции своих машин Корсаков впервые в истории информатики применил перфорированные карты. В первой половине XIX века Семён Николаевич Корсаков выдвинул концепцию усиления возможностей разума посредством разработки научных методов и устройств. В 1832 году он опубликовал описание изобретенных им механических устройств, так называемых «интеллектуальных машин».

В своих машинах Корсаков впервые предложил использовать перфорированные карты для задач информационного поиска и классификации. В работах Корсакова содержится целая плеяда новых для того времени идей, как-то: многокритериальный поиск с учетом относительной степени важности различных критериев (весовых коэффициентов), способ обработки больших массивов данных – предтеча современных экспертных систем – попытка определить понятие алгоритма.

В работах Корсакова содержится целый ряд новых для того времени решений, как-то:

- многокритериальный поиск с учетом относительной степени важности различных критериев (весовых коэффициентов),

- способ обработки больших массивов данных,

- предтеча современных экспертных систем,

- попытка определить понятие алгоритма.

Он предложил общий принцип поиска или сравнения идей (записей, фактов) на основе их деталей (признаков).

Таким образом, Корсаков, выражаясь современным языком, определяет информационную запись, набором ее признаков – координатное индексирование по Солтону. Словарь Корсакова представлял перфорированные таблицы, где каждый столбец определяет некоторую идею, а в строках перфорируются признаки этой идеи. Перфорация отверстий обеспечивает возможность механического поиска и сравнения идей на основе их признаков. Исторически тезаурусы создавались для ручного индексирования документов и при их создании не принимались во внимание вопросы, связанные с автоматической индексацией. Трудность построения тезауруса, соответствующего всему тематическому многообразию индексируемой информации, является основной причиной его непопулярности в современных информационных системах. Но эффективность информационно-поисковых систем для поддержки научно-образовательной деятельности напрямую зависит от использования специализированных тезаурусов.

Идеографическая классификация.

Один из первых в истории и наиболее известных на сегодня тезаурусов (идеографических словарей) составлен британским лексикографом Питером Марком Роже и опубликован в 1852 г. тезаурус Роже (Roget's Thesaurus) — оригинальное название Thesaurus of English Words and Phrases – «Тезаурус английских слов и фраз». Непосредственным и предшественниками словаря П. М. Роже были уже упоминавшиеся работы Д. Далгарно и Д. Уилкинса, санскритский словарь «Амаракоша», переведенный на английский язык в 1808 году.

Роже в полной мере использовал опыт своих предшественников, он пишет: «Принцип, которым я руководствовался, классифицируя слова, является тем же самым, который используется при классификации особей в различных областях естественной истории. Поэтому разделы, выделенные мной, соответствуют естественным семьям ботаники и зоологии, а ряды слов сцементированы теми же отношениями, которые объединяют естественные ряды растений и животных».

В предисловии к своему словарю он пишет: ...Какой бы живостью ни обладало наше воображение, как бы ни переполняли нас чувства, мы часто попадаем в такое положение, когда нам не хватает слов, чтобы точно выразить свою мысль. Единственно необходимое слово зачастую бежит нашей памяти, и мы вынуждены обходиться словами слишком сильными или слишком слабыми, слишком общими или излишне конкретными. Помощь, которую оказывает этот словарь, состоит в предоставлении богатейшего набора слов и выражений, исчерпывающих все оттенки и нюансы каждой общей идеи...

Исходя из общих методологических принципов классификации понятий можно говорить о трех типах идеографических словарей: 1. Идеографический словарь-тезаурус — основан на логической рубрикации всего понятийного содержания лексики. Главной задачей является идентификация и последующая рациональная классификация понятийных групп, реально представленных в лексике языка, характер и количество лексико-семантических групп определяется смысловой емкостью языка.

2. Аналогический словарь-тезаурус — основан на выделении тематических лексико-семантических групп, которые располагаются в порядке алфавитного следования тематических доминант (слов-центров). 3. Тематический (учебный) словарь-тезаурус — главной задачей выбор тем, характеризующий выбранную предметную область (актуальных для учебных целей). Основные требования, которые предъявляются к классификационной системе, сводятся к следующему: а) классифицироваться должны не слова определенного языка, а понятия, что обеспечивает универсальность системы; б) классифицируются понятия исходные, лежащие в основе языка; в) направляющим принципом классификации является осознание системы понятий как определенным

образом организованного единства, расчленение которого должно вестись в естественной последовательности.

## Полисемия в информатике.

### Словарные статьи в коллекции: (public\_cat = Thesaurus of Information Technology: Dictionary Articles )

**информация**

*Информация* - сведения (сообщения, данные) независимо от формы их представления (Федеральный закон «Об информации, информатизации и защите информации»).

*Информация* (в процессах ее обработки) - любой факт, понятие или значение, полученные из данных, а также контекст, выбранный из знаний, или контекст, ассоциированный со знаниями (стандарт ISO/IEC 2382-1).

*Информация* - сведения, воспринимаемые человеком и (или) специальными устройствами как отражение фактов материального или духовного мира в процессе коммуникации (ГОСТ 7.0-99. Информационно-библиотечная деятельность, библиография).

В настоящий момент в литературе нет четкого определения, что такое "информация" - считается, что это понятие является интуитивно понятным, как понятие точки в геометрии. В обиходе "информацией" называют любые данные или сведения, которые кого-либо интересуют. Например, сообщение о каких-либо событиях, о чьей-либо деятельности и т.п. "Информировать" в этом смысле означает "сообщить нечто, неизвестное раньше". С середины 20 века понятие "информация" стало общенаучной категорией, что было связано с всеобщим распространением ЭВМ и становлением науки "информатика", а так же введением количественной меры информации, разработкой количественной теории информации (по Шеннону и Колмогорову).

Ниже приводятся некоторые определения понятия информация:

- ИНФОРМАЦИЯ** (от лат. informatio - разъяснение, осведомление) - любые сведения и данные, отражающие свойства объектов в природных (биологических, физических и др.), социальных и технических системах и передаваемые звуковым, графическим (в том числе письменным) или иным способом без применения или с применением технических средств.
- ИНФОРМАЦИЯ** — это обозначение содержания сообщения, полученного из внешнего мира в процессе нашего приспособления к нему и приспособления к нему наших чувств.
- ИНФОРМАЦИЯ** - (по Колмогорову) «Информация есть характеристика не сообщения, а соотношения между сообщением и его потребителем. Без наличия потребителя, хотя бы потенциального, говорить об информации бессмысленно».
- ИНФОРМАЦИЯ** - часть знаний, которая используется для ориентирования, активного действия, управления, т.е. в целях сохранения, совершенствования, развития системы (Н. Винер).

Важность (ценность) какой-либо информации зависит от множества обстоятельств и, по существу, не поддается формализации. Однако можно четко определить основные свойства информации исходя из подхода к ее определению. Условно

Рисунок 1. Словарные статьи в коллекции

### Словарь терминов в коллекции: Thesaurus of Information Technology (zthes\_cat)

**Тезаурус [ru]**

*Тезаурус* - совокупность терминов, описывающих данную предметную область, с указанием семантических отношений (связей) между ними.

— это словарь, в котором слова и словосочетания с близкими значениями сгруппированы в единицы, называемые понятиями, концептами или дескрипторами, и в котором явно (в виде отношений, иерархии) указываются семантические отношения между этими понятиями (концептами, дескрипторами).

**(NT) Тезаурус (add) [ru]**

**Головные термины:**  
[BT] Система организации знаний [ru]

**Дочерние термины:**  
[NT] Аппаратный-издеу тезаурус [kz]  
[NT] Информационно-поисковый тезаурус [ru]  
[NT] Лексическая единица ШПЯ [ru]  
[NT] Тезаурус WordNet [ru]  
[NT] Тезаурус Роже [ru]  
[NT] Тезаурус РуТез [ru]  
[NT] Тезаурус ЮНЕСКО [ru]  
[NT] Управленческий словарь [ru]

**Ключевые термины, связанные с термином**  
[RT] База знаний [ru]  
[RT] Дескрипторный словарь [ru]  
[RT] Иерархический указатель информационно-поискового тезауруса [ru]  
[RT] Искусственный интеллект [ru]  
[RT] Лексико-семантический указатель информационно-поискового тезауруса [ru]  
[RT] Лексическая единица [ru]  
[RT] Многоязычный информационно-поисковый тезаурус [ru]  
[RT] Одноязычный информационно-поисковый тезаурус [ru]  
[RT] Онтология [ru]  
[RT] Парадигматические отношения [ru]

Рисунок 2. Словарь терминов в коллекции

В информационной системе тезаурус является не только самостоятельным информационным ресурсом, но и инструментом для классификации или индексации ресурсов. Пользователь информационной системы должен иметь возможность:

- Осуществлять просмотр тезауруса.
- Осуществлять поиск ресурсов по ассоциированным с ними терминам или понятиям. о Поиск ресурсов может вестись двумя способами:
  - Поиск, по ключевым словам, используя тезаурус.
  - Навигация по тезаурусу. То есть поиск сначала нужного понятия в тезаурусе с последующим запросом ресурсов, соответствующих этому понятию.

При поиске ресурсов, по ключевым словам, поисковая система может, используя тезаурус, расширять результаты поиска, выдавая пользователю не только ресурсы, соответствующие введенным пользователем ключевым словам, но и ресурсы, соответствующие связанным с ними терминам, или терминам, обозначающим также более узкие понятия относительно исходного термина. Например, если пользователь ищет



ресурсы, соответствующие термину «туннель», в результатах поиска необходимо выдать также все ресурсы, соответствующие термину «тоннель», поскольку оба они являются разными правильными вариантами написания одного и того же слова.

Или если ищутся ресурсы, соответствующие понятию дифференциальные и функциональные уравнения, имеет смысл включить в результаты поиска также ресурсы, соответствующие рубрике системы функциональных уравнений и неравенства. Система поиска может также, используя тезаурус, подсказать пользователю, по каким еще словам ему стоит осуществить поиск (например, квазисинонимы, связанные термины, более широкие термины, и т.д.). Оба этих варианта использования тезауруса широко применяются, например, в поисковых машинах.

Интерфейс просмотра тезауруса должен:

- Показывать все атрибуты данного термина или понятия.
- Показывать, с какими терминами и понятиями связан данный термин или понятие.
- Достаточно наглядно показывать пользователю место термина или понятия в иерархии понятий тезауруса.

Научная новизна проводимых исследований заключается в том, что впервые в фокусе казахстанской лингвистической науки находятся интеграционные и трансдисциплинарные подходы в изучении научно-теоретических и методических основ разработки полиязычных управляемых тезаурусов отраслевых терминов, которые будут эффективно использованы в компьютерной лингвистике, отраслевой терминологии, в области машинного и автоматизированного перевода с использованием NLP (Natural Language Processing). Впервые проводится систематизация и рубрификация отраслевых терминов по предметным областям на основе казахского языка. Это позволяет упорядочить и классифицировать отраслевые термины по предметным областям. Исследования такого плана помогут укрепить и расширить функции научного стиля казахского языка, а также его интеграцию с английским и русским языками на электронных платформах для информационного поиска (Information Retrieval). Тематика исследования актуальна и вносит большой теоретический и практический вклад в развитие отраслевой терминологии, способствует созданию параллельного корпуса казахского языка, расширению функций казахского языка в глобальном мире. Все это обусловлено, во-первых, междисциплинарным интересом к изучению отраслевой терминологии. Во-вторых, прикладным характером результатов исследования: практического апробирования интероперабельности полиязычных тезаурусов. В-третьих, расширением функций казахского языка в научной коммуникации.

Работа поддержана грантом финансирования научных, научно-технических проектов на 2022-2024 гг. Комитетом науки Министерства образования и науки Республики Казахстан (AP14972834).

#### Список использованной литературы

1. А. М. Федотов, О. А. Федотова, М. А. Самбетбаева. Информационные системы: модели и технологии : Учеб.пособие // ОТВ. РЕД. В. Б. Барахнин; Новосиб. гос. ун-т. — Новосибирск : ИПЦ НГУ, 2019. — 265 с.

2. Самбетбаева М.А. Ғылыми-білім беру қызметін қолдауға арналған ақпараттық жүйе үшін ақпараттық технологиялар бойынша қазақ тілі морфологиясын ескере отырып көптілді тезаурус жасау [Мәтін]: Философия докторы (PhD) дәрежесін алу үшін дайындалған диссертация: 6D070300 – Ақпараттық жүйелер. – Астана, 2016 – 148 б.

3. Федотов А.М., Идрисова И.А., Самбетбаева М.А., Федотова О.А. Использование тезауруса в научно-образовательной информационной системе // Вестник Новосибирского государственного университета. Серия: Информационные технологии. - 2015. - Т.13, № 2. - С.86-102.

4. Шрейдер Ю.А. О количественных характеристиках семантической информации // НТИ.Сер.2.-1963.-№10. — с.35-39.
5. ISO 25964-1:2011 Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval, 2011
6. ISO 25964-2:2013 Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies, 2013

## АНАЛИЗ ТЕХНОЛОГИЧЕСКИХ, ЭКОНОМИЧЕСКИХ ФАКТОРОВ, КОТОРЫЕ ПОВЛИЯЮТ НА БУДУЩЕЕ SIEM

*Қуатова Жұлдыз Қуатқызы, Ерманов Әлқожа Маратұлы, Азанов Николай Прокопьевич  
Казахский национальный университет имени Аль-Фараби, г. Алматы  
E-mail: [zhuldyz.kuatova@mail.ru](mailto:zhuldyz.kuatova@mail.ru)*

**Аннотация:** В настоящее время SIEM-системы и связанные с ними решения постепенно приближаются к инструментам анализа больших данных. Мы рассматриваем наиболее широко используемые SIEM с точки зрения их критической функциональности и предоставляем анализ внешних факторов, влияющих на ландшафт SIEM.

**Ключевые слова:** Эволюция Siem; Совершенствование SIEM; Тенденции SIEM.

### Введение

Риски кибербезопасности, влияющие на промышленные системы управления (ИКТ), чрезвычайно возросли за последние пару лет, главным образом из-за возросшей активности национальных государств и киберпреступников. Злоумышленники стали более изощренными и опасными, и их надлежащее и своевременное обнаружение стало реальной проблемой. Примерами текущих инцидентов кибербезопасности, влияющих на ИТ и ИКТ, являются: Атаки программ-вымогателей; Вредоносное ПО, влияющее на способность утилиты вести бизнес и операции; Фишинговые кампании, направленные на руководителей, помощников руководителей, инженеров SCADA, ИТ-администраторов или других привилегированных пользователей; Случаи компрометации деловой электронной почты, включая захват учетных записей или выдачу себя за руководителей; Утечка данных и кражи; и Социальная инженерия для сбора конфиденциальной информации от персонала. Согласно недавнему отчету NIST [10], решения по кибербезопасности в промышленных системах управления должны обеспечивать обнаружение поведенческих аномалий в режиме реального времени, обеспечивать более быстрое управление инцидентами и обеспечение интеллектуальной визуализации сети и всех ее взаимосвязанных узлов. Системы информации о безопасности и управления событиями (SIEM) рассматривают вышеупомянутые возможности как встроенные функции.

### Обзор литературы

В этой статье мы рассмотрим наиболее широко используемые средства защиты информации и управления событиями (коммерческие и с открытым исходным кодом) с целью определения их основных характеристик, преимуществ и ограничений для обнаружения текущих сценариев атак и реагирования на них.

### Метод

Мы проводим углубленный анализ функций и возможностей существующих SIEM и фокусируемся на их ограничениях, чтобы предложить потенциальные усовершенствования для интеграции в существующие платформы SIEM. Анализ внешних факторов (например, политических, экономических, социальных), которые потенциально могут повлиять на будущее SIEM в средне- и долгосрочной перспективе, предоставляется как способ выявления факторов, способствующих внедрению SIEM-систем нового поколения, и барьеров на их пути.

### Обсуждение

Меняющийся характер угроз безопасности, распространение мобильных устройств, глобализация, бурный рост социальных сетей и быстрые изменения в регулировании ускоряют эволюцию информации о безопасности и управления событиями. Целью данной статьи является анализ внешних факторов, которые потенциально могут повлиять на будущее систем SIEM и связанных с ними технологий в среднесрочной и долгосрочной перспективе на основе технологических факторов [1].

Облачное хранилище. Эту технологию можно четко рассматривать как средство, способствующее технологии SIEM, поскольку анализ сетевых событий с использованием больших данных может выполняться более эффективным способом, не беспокоясь о количестве хранимых журналов, информации и т.д.

Интеграция с облачными сервисами. Это относится отдельно к облачному хранилищу, поскольку оно больше ориентировано на выполнение программного обеспечения на удаленном сервере, а не только на “статическое” хранение данных в облачной инфраструктуре. Эта технология позволяет обеспечить масштабируемость и высокую доступность программных приложений, поскольку они не ограничены аппаратным обеспечением локального сервера и могут быть запущены из любого места.

Мобильные технологии. Рост числа мобильных устройств порождает новые угрозы, которые должны быть проанализированы SIEM-системами. В этом отношении наблюдается тенденция к тому, что сотрудники используют устройства, принадлежащие компании, а также персональные устройства для офисной работы. Необходимо было бы обеспечить безопасность корпоративных данных. Работа дома, например, с персональным компьютером, то, что сейчас обычно называют BYOD (Bring Your Own Device), является тенденцией в области кибербезопасности [2]. Однако это приводит к нескольким потенциальным проблемам: устройства BYOD не управляются ИТ-командой, поэтому они не находятся под контролем политики компании; некоторые BYOD не имеют предустановленного решения для обеспечения безопасности; данные на этих устройствах не зашифрованы; приложения, установленные на этих устройствах, невозможно отслеживать.

Аналитика больших данных. Как уже упоминалось ранее, SIEMs эволюционируют в системы анализа данных. Данные в подключенной среде растут экспоненциально и делают необходимым наличие мощных инструментов анализа, способных анализировать события в реальном времени, поддерживать принятие решений и т.д.. Развитие методов анализа данных, несомненно, способствует развитию SIEM-систем.

Технологии машинного обучения. Новые высокопроизводительные компьютеры с мощным оборудованием и современными программными языками, а также аналитика данных, описанная выше, позволяют создавать модели данных, основанные на опыте причинно-следственного анализа. SIEMs может воспользоваться преимуществами этих технологий, чтобы сделать обнаружение событий и принятие решений более разумными [3].

Internet of Everything (IoE) [1] – это вездесущая коммуникационная сеть, которая эффективно собирает, управляет и использует данные из миллиардов реальных объектов и физических действий. Это расширяет концепцию Интернета вещей (IoT), также включая людей, процессы, местоположения и многое другое. Влияние этой технологии в SIEMs заключается в том, что они предоставляют большой объем данных и событий для анализа.

Сети 5G. 5G представляет собой следующее поколение коммуникационных сетей и сервисов, подход к выполнению требований будущих приложений и сценариев. Эта технология увеличит скорость передачи данных, а затем может повлиять на объем данных, анализируемых SIEM в сети в единицу времени. Это может создать трудности для SIEMs в обнаружении событий.

Среди экономических факторов, которые повлияют на будущее SIEMs, можно выделить следующие:

Краткосрочная/временная работа. В 2014 году основным типом трудовых отношений в ЕС были постоянные контракты на полный рабочий день, на долю которых приходилось 59% занятости, хотя этот показатель снижается, в то время как доля нестандартных форм работы увеличивается. Если эта тенденция сохранится, вполне может случиться так, что в течение следующего десятилетия стандартные контракты будут применяться только к меньшинству работников [9]. Из-за новых видов работы, тяготеющих к более коротким срокам, люди теперь не остаются надолго в одной и той же

компаниям, особенно в первый период своей карьеры. Следствием этого является то, что компаниям необходимо свести к минимуму время, затрачиваемое сотрудниками на освоение нового инструмента или нового способа работы. Поэтому этот фактор делает крайне важным, чтобы будущие SIEM имели улучшенные и более дружелюбные интерфейсы на уровне принятия решений, правил настройки, ссылок на новые источники и датчики.

Внештатный сотрудник. Самозанятость растет по сравнению с обычной оплачиваемой занятостью в компании [8]. Фрилансеры работают в компании не как сотрудники, а как поставщики услуг. Этот вид работы может представлять угрозу для компаний, поскольку устройства, используемые фрилансерами, не принадлежат ИТ-отделу и их нелегко контролировать.

Кроме того, у них нет прочных связей с компанией, которая нанимает их услуги. Тем не менее, фрилансер по кибербезопасности может быть хорошим выбором для поставщиков SIEM, поскольку они могут обладать более широкими знаниями о потенциальных угрозах, влияющих на организацию, поскольку они накапливают большой опыт работы в разных компаниях.

Число рабочих мест в сфере кибербезопасности постоянно растет. Предполагаемый рост числа рабочих мест в сфере кибербезопасности к 2020 году составит 35% [7]. Это отражает важность кибербезопасности для компаний, и это может стать возможностью для SIEMs расти на рынке.

Крупные компании, глобализация. Глобальный рынок облегчает выживание и дальнейший рост крупных технологических компаний [6]. Однако уровень критичности этой информации может быть выше. Будущие SIEMs должны быть рассчитаны на такие крупные компании и глобальные сети.

Малые и средние предприятия. В будущем малые и средние предприятия станут более крупными объектами кибератак [5]. Они должны стать новой целью для роста рынка SIEM, что сделает такие модели, как SIEM как услуга, более привлекательными для малого и среднего бизнеса.

Общество становится все более зависимым от информационно-коммуникационных технологий (ИКТ), что ведет к быстрому социальному, экономическому и государственному развитию. Ниже рассказывается о том, как изменения в привычках общества, связанные с технологиями, повлияют на будущее SIEMs.

Поколение Z. Современные поколения понимают мир как большую сеть, в которой все подключено к Интернету. Можно предположить, что люди будущего будут лучше осведомлены о кибербезопасности и принесут компаниям более четкое представление о рисках, связанных с угрозами в сети.

Рост социальных сетей. За последние несколько лет наблюдается огромный рост использования социальных сетей среди молодого поколения. Активность в социальных сетях – это источник данных, которым нельзя пренебрегать, и он может иметь очень большое значение при анализе событий безопасности.

Кибератаки. В новых взаимосвязанных обществах развитие Интернета привело к новому типу атак, то есть кибератакам. Атаки на критические инфраструктуры можно считать новым оружием, что делает SIEMs незаменимыми в любой инфраструктуре, в которой данные имеют значение или атака на которую может привести к сбоям в работе или даже к ущербу для населения, не только с точки зрения интересов отдельной компании, но и с точки зрения пользователей, граждан и (в более общем плане) людей. [4].

Глубокая паутина. Глубокая паутина – это часть Всемирной паутины, содержимое которой не индексируется стандартными поисковыми системами [4]. Это может рассматриваться SIEM-системами как препятствие, поскольку затрудняет извлечение данных из сети.

## Заключение

В заключение, роль SIEM также была изучена в ближайшем и долгосрочном будущем с учетом различных аспектов (например, политических, экономических, социальных, технологических, экологических и правовых факторов) в различных критически важных инфраструктурах. Из этого анализа мы можем сделать вывод, что условия хороши для стимулирования инвестиций в совершенствование и расширение этой технологии в качестве ключевого компонента не только для промышленных систем управления с Центрами управления безопасностью, но и для обеспечения управления кибербезопасностью для МСП с ограниченными знаниями и возможностями в области безопасности.

## Список литературы

1. Лонгботтом, С. Познакомьтесь с основами и преимуществами программного обеспечения SIEM. Технический документ TechTarget, 2018 год.
2. Дж. Брэдли, Дж. Лоукс, Дж. Маколей, Р. Медкалф, Л. Бакалю.: BYOD: Глобальная перспектива, Отчет об исследовании Cisco, (2012).
3. Суарес-Тангил, Г.; Паломар, Э.; Рибаргора, А.; Санц, И. Обеспечение систем SIEM самоадаптацией, Слияние информации, 2015, 21, 145-158.
4. П. Пассери.: Статистика кибератак за 2016 год. Хакмагеддон, Сроки и статистика информационной безопасности, 2017.
5. С. О'Брайен.: Малые и средние предприятия станут еще более крупными объектами кибератак в 2016 году. Непрерывная разведка безопасности, (2016).
6. The economist.: Почему гиганты процветают. Сила технологий, глобализации и регулирования. Специальный доклад (2016).
7. С. Морган.: Один Миллион Вакансий В Сфере Кибербезопасности В 2016 Году. Статья в журнале Forbes (2016).
8. П. Лейтон.: Работа В Будущем: Рост Числа Независимых Профессионалов В Европе. Европейский форум независимых профессионалов, (2016).
9. Европейский парламент.: Нестабильная занятость в Европе. Часть 1: Закономерности, тенденции и политическая стратегия, (2016).
10. CyberX.: Рекомендации NIST по безопасности Интернета вещей и микросхем, 2019.

# ФЭЙК АККАУНТ АНЫҚТАУ МӘСЕЛЕСІНДЕ МАШИНАЛЫҚ ОҚЫТУДЫ ҚОЛДАНУ

*Азаматова Д.Т., Мусиралиева Ш.Ж., Голлманн Д.  
Әл-Фараби атындағы Қазақ Ұлттық Университеті, Алматы қаласы,  
[dinarazamatova@gmail.com](mailto:dinarazamatova@gmail.com)*

## Кіріспе

Facebook және Instagram сияқты Онлайн әлеуметтік медиа (OSN) барған сайын танымал болып, қазіргі әлемнің маңызды бөлігіне айналууда. OSN-ді байланыс құралы ретінде пайдаланудан басқа, олар бизнесті танымал ету және алға жылжыту үшін де қолданылады. Бір қарағанда, аккаунттың танымалдығы ізбасарлар саны немесе ұнатулар, пікірлер немесе көріністер саны сияқты жалпы мазмұн қасиеттері сияқты кейбір өлшемдермен өлшенеді. [1]

Бұл жұмыста әлеуметтік желі ретінде Instagram әлеуметтік желісі алынды, себебі бұл желі айына 1 миллиардқа жуық белсенді қолданушыларға және айына 2 миллион жарнама берушілерге жетті, ал пайдаланушылар күніне 4,2 миллиард хабарламаға лайк басады. Сондықтан осындай маңызды әлеуметтік платформада салауатты ортаны сақтау өте маңызды.

### 1. Фейк аккаунтты анықтауға мақсатында деректер жиынтығы

Жобаның бірінші кезеңі – аккаунт туралы ақпаратты жинау. Ол үшін кез келген веб-сайтты API-ге айналдыра алатын веб-беттерді жинау және автоматтандыру үшін Apify платформасын қолданылды. Бұл зерттеуде Instagram профильдерінен пайдалы деректерді алынды. Instagram Profile Scraper көмегімен Instagram пайдаланушылары туралы жалпыға қолжетімді деректерді жинауға мүмкіндік берді. Бұл жеке профильдерден емес, жалпыға ортақ профильдерден деректерді шығаруға болатынын білдіреді.[2]

Жобада деректер жиынтығы 2-ге бөліп қарастырылды: бірі 11 функциядан, яғни: профиль суреті, пайдаланушы атының ұзындығы, толық профильдің аты, bio length, сыртқы URL, жеке аккаунт растығына тексеру(is private), тексерілген аккаунт(is verified), бизнес аккаунт растығына тексеру, #post, #followers, #following, екіншісі 14 функциядан: профиль суреті, пайдаланушы атының ұзындығы, толық профильдің аты, bio length, сыртқы URL, тексерілген аккаунт(is verified), бизнес аккаунт растығына тексеру, #post, #followers, #following, соңғы хабарлама(пост) жарияланған уақыт, бір күнде жарияланатын посттың көлемі(саны), белсенділік индексі, ұнатулардың орташа көрсеткіші(мәні) құралған.

### 2. Препроцессинг кезеңі

Препроцессор - бұл басқа бағдарламаны (мысалы, компилятор) енгізуге арналған кіріс деректерін қабылдайтын және деректерді шығаратын компьютерлік бағдарлама. Препроцессордың шығысындағы мәліметтер олар кейінгі бағдарламалармен (компилятор) өңдеуге жарамды препроцессорлық формада болады деп айтылады. Өңдеудің нәтижесі мен түрі процессордың түріне байланысты; сонымен, кейбір процессорлар тек Қарапайым мәтінді алмастыра алады, ал басқалары бағдарламалау тілдерімен салыстыруға қабілетті. Препроцессорды қолданудың ең көп таралған жағдайы — бастапқы кодты компиляцияның келесі сатысына жібермес бұрын өңдеу.

Үлгілердегі дәлдікті арттыру көбінесе деректерді түрлендірудің алғашқы қадамдары арқылы алынады. Мүмкіндіктерді масштабтау машиналық оқытудағы деректерді алдын ала өңдеудің ең маңызды қадамдарының бірі болып табылады. Екі мәліметтер жиынтығы алдын-ала өңдеу кезеңінен өткізіледі. Бұл кезең екі деректер жиынтығын стандарттау мен қалыпқа келтіруден тұрады.[3]

### 3. Машиналық оқытудың жіктеу алгоритмдерін қолдану.

Алдын-ала өңделген функциялар келесі машиналық оқыту классификаторының алгоритмдеріне қабылданады және беріледі:[4]

- Шешімдер ағашы(Decision Tree);
- Кездейсоқ орман(Random Forest);
- Қолдау векторлық әдісі (SVM)

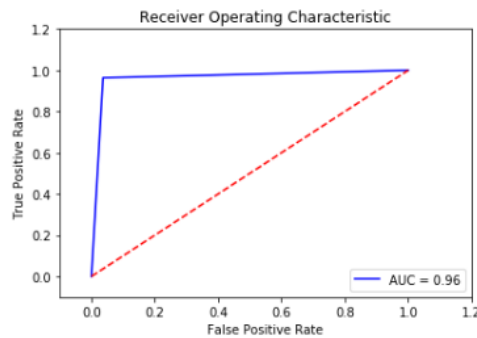
Әрбір алгоритм үшін оқыту мен тестілеу кезеңінен басқа ROC-қисығы және Жіктеу туралы есеп есептелді.

Модельді белгілі бір шекке байланбай, тұтастай бағалаудың бір жолы AUC-ROC (немесе ROC-қисығы) – қате қисығының астындағы аудан (Қисық астындағы аудан немесе Қабылдағыштың жұмыс сипаттамасы қисығы). Бұл қисық шынайы оң мөлшерлеме (TPR) және жалған оң мөлшерлеме (FPR) координаттарындағы (0,0) мен (1,1) аралығындағы сызық.

Айтылған классификация алгоритдерінің нәтижелері келесіндей:

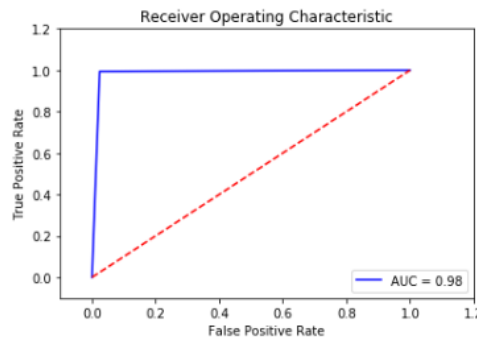
```
In [15]: plot_roc_curve(y_test, clf.predict(X_test))
```

```
False Positive rate: [0.          0.03636364  1.          ]
True Positive rate: [0.          0.96363636  1.          ]
```



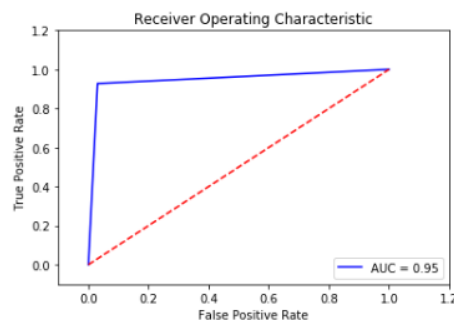
Сурет 1. Шешімдер ағашы алгоритмі бойынша ROC қисығы

```
False Positive rate: [0.          0.02424242  1.          ]
True Positive rate: [0.          0.99393939  1.          ]
```



Сурет 2. Кездейсоқ орман (Random Forest) алгоритмі бойынша ROC қисығы

```
False Positive rate: [0.          0.03030303  1.          ]
True Positive rate: [0.          0.92727273  1.          ]
```



Сурет 3. Қолдау векторлық әдісі (SVM) алгоритмі бойынша ROC қисығы



Көрсетілген нәтижелерге қарай қисықтар шынайы оң мөлшерлеме (TPR) және жалған оң мөлшерлеме (FPR) координаттарындағы аралық сызық (0,1) жақын. Және AUC-ауданының 0,95-тен жоғары.

Жіктеу туралы есеп

Машиналық оқытуда жіктеу үлгілерін пайдаланған кезде модельдің сапасын бағалау үшін қолданатын үш жалпы өлшем бар:

1. precision: жалпы оң болжамдарға қатысты дұрыс оң болжамдардың пайызы.
2. recall: жалпы нақты позитивтерге қатысты дұрыс оң болжамдардың пайызы.
3. F1-ұпайы: Дәлдік пен еске түсірудің өлшенген гармоникалық ортасы. 1-ге жақын болған сайын үлгі жақсырақ болады.
4. support: көрсетілген деректер жиынындағы сыныптың нақты оқиғаларының саны.

Осы төрт көрсеткішті пайдалана отырып, біз берілген жіктеу моделінің қандай да бір жауап айнымалысы үшін нәтижелерді болжай алатынын түсінуге болады.

```
In [16]: print(classification_report(y_test, clf.predict(X_test)))
```

	precision	recall	f1-score	support
0	0.96	0.96	0.96	165
1	0.96	0.96	0.96	165
accuracy			0.96	330
macro avg	0.96	0.96	0.96	330
weighted avg	0.96	0.96	0.96	330

Сурет 4. Шешімдер ағашы алгоритмі бойынша жіктеу туралы есеп

Шешім ағашының классификаторының деректерді жіктеудегі тиімділігі 96% құрады.

	precision	recall	f1-score	support
0	0.99	0.98	0.98	165
1	0.98	0.99	0.98	165
accuracy			0.98	330
macro avg	0.99	0.98	0.98	330
weighted avg	0.99	0.98	0.98	330

Сурет 5. Кездейсоқ орман(Random Forest) алгоритмі бойынша жіктеу туралы есеп

Кездейсоқ орман(Random Forest) классификаторының деректерді жіктеудегі тиімділігі 98% құрады.

	precision	recall	f1-score	support
0	0.93	0.97	0.95	165
1	0.97	0.93	0.95	165
accuracy			0.95	330
macro avg	0.95	0.95	0.95	330
weighted avg	0.95	0.95	0.95	330

Сурет 6. Қолдау векторлық әдісі (SVM) алгоритмі бойынша жіктеу туралы есеп

Қолдау векторлық әдісі (SVM) классификаторының деректерді жіктеудегі тиімділігі 95% құрады.

## Қорытынды

Желілер біздің өмірімізге өте терең еніп, танымал бола бастады, олардың адамдарға әсері әлі де бағаланбаған. Қазір жалпы типтегі желілер негізінен дамуда, бірақ жақын арада тақырып бойынша да, функционалдылығы бойынша да жаңа жобалар дамитын болады. Алайда, сонымен қатар жалған ақпарат таратушылар, жеке деректерді жинау мақсатындағы алайықтық істейтіндер де көбейе түседі. Сондықтан, әлеуметтік желі мен мессенджерлердегі фэйк аккаунтты анықтау мәселесінде машиналық оқытуды қолдану өте маңызды.

Бұл жұмыста фэйк аккаунтты анықтау мақсатында машиналық оқыту классификаторларын тиімді қолданылды. Және жобада таңдалған машиналық оқытудың классификатор алгоритмдерінен ең жоғары салыстырмалы дәлдік, яғни 98% көрсеткіш көрсеткен Random Forest алгоритмі болды.

## Әдебиеттер тізімі:

1. Estee Van Der Walt and Jan Eloff, "Using Machine Learning to Detect Fake Identities: Bots vs Humans," IEEE, 2018
2. Cao Qiang, Xiaowei Yang, Jieqi Yu, and Christopher Palow. "Uncovering large groups of active malicious accounts in online social networks." In Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, pp. 477-488. 2014.
3. Indira Sen et. al. "Worth its Weight in Likes: Towards Detecting Fake Likes on Instagram," ACM, 2018.
4. Cao Xiao, David Freeman and Theodore Hwa, "Detecting Clusters of Fake Accounts in Online Social Networks," ACM, 2015.
5. Boshmaf Y., Logothetis D., Siganos G., Leria J., Lorenzo J., Ripeanu M., Bezno-  
sov K., Halawa H. integro: Leveraging victim prediction for robust fake account detection in large scale OSNs. Comput. Secur., 2016
6. Egele M., Stringhini G., Kruegel C., Vigna G. Towards Detecting Compromised Accounts on Social Networks. IEEE Trans. Dependable Secure Comput PP, 2015
7. Avdoshin S., Lazarenko A. Deep Web Users Deanonymization System. Proceed-  
ings of the Institute for System Programming of the RAS, 2016
8. T. Information, "Instagram's Growing Bot Problem,"  
[www.theinformation.com/articles/instagrams-growing-bot-problem](http://www.theinformation.com/articles/instagrams-growing-bot-problem), accessed: 2019-06-10.
9. R. Raturi, "Machine learning implementation for identifying fake accounts in social network," International Journal of Pure and Applied Mathematics, vol. 118, no. 20, sf. 4785–4797, 2018.
10. R. Kaur and S. Singh, "A survey of data mining and social network analysis based anomaly detection techniques", Egyptian informatics journal, vol. 17, no. 2, pp. 199-216, 2016.

## ПРОБЛЕМЫ ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ХРАНЕНИЯ И ОБРАБОТКИ БОЛЬШИХ ДАННЫХ

*Кубигенова А.Т., Актаева А.У., Исмаилова А.А., Есмагамбетова Г.К.*  
*Казахский агротехнический университет им. С. Сейфуллина, г. Нур-Султан,*  
[akku\\_kubigenova@mail.ru](mailto:akku_kubigenova@mail.ru)

*Кокшетауский университет им.Ш.Уалиханова, г. Кокшетау, [aaktaewa@list.ru](mailto:aaktaewa@list.ru)*  
*Казахский агротехнический университет им. С. Сейфуллина, г. Нур-Султан,*  
[a.ismailova@mail.ru](mailto:a.ismailova@mail.ru)

*Монгольский университет науки и технологии, Улан-Батор, [gal.esm@mail.ru](mailto:gal.esm@mail.ru)*

**Аннотация.** Большие данные несут в себе несколько угроз безопасности, которые могут негативно сказаться на организациях. Отсутствие мер безопасности при хранении и обработке больших данных может привести к утечке данных. В то время как упрощение доступа к данным имеет важное значение для компаний, контроль над большими данными не менее важен для обеспечения доверия между их клиентами.

Статья начинается с краткого ознакомления с большими данными, затем освещаются проблемы, с которыми сталкивается безопасность больших данных.

**Ключевые слова:** большие данные, проблемы безопасности, конфиденциальность информации.

Сегодня большие данные имеют решающее значение для успеха любого бизнеса в мире, управляемом данными. Однако большие данные несут в себе несколько угроз безопасности, которые могут негативно сказаться на организациях. Благодаря передовым решениям инновационных технологий в области больших данных инфраструктуры организации могут оптимизировать поток данных информационных ресурсов для предоставления в режиме реального времени и более эффективного принятия решений.

Большие данные — это большие диверсифицированные наборы данных, поступающие из множества каналов: платформы социальных сетей, веб-сайты, электронные регистрации, датчики, покупка продуктов, журналы вызовов — выбор безграничен. Большие данные обладают тремя уникальными характеристиками:

- 1) volume – сверхбольшой объем данных,
- 2) velocity – очень высокая скорость передачи данных,
- 3) variety – слабая структурированность данных, которая понимается, прежде всего, как нерегулярность структуры данных и трудность извлечения однородных данных из потока и выявления корреляции.

С развитием технологий BIGDATA добавились и другие критерии, такие как, veracity (достоверность), variability (изменчивость), value (ценность), visibility (обозримость) (рис. 1).

Интеллектуальный анализ данных, машинное обучение и прогнозная аналитика — это лишь некоторые из недавно разработанных методов, используемых для получения новых сведений о неиспользованных областях источников данных для оптимизации бизнес-процессов. [1]

Постоянно растущие потоки данных представляют, как возможности, так и проблемы. Хотя перспектива более качественного анализа потоков данных информационных ресурсов позволяет принимать более обоснованные решения, есть определенные недостатки, такие как проблемы с безопасностью, из-за которых компании могут оказаться в затруднительном положении при работе с конфиденциальной информацией.

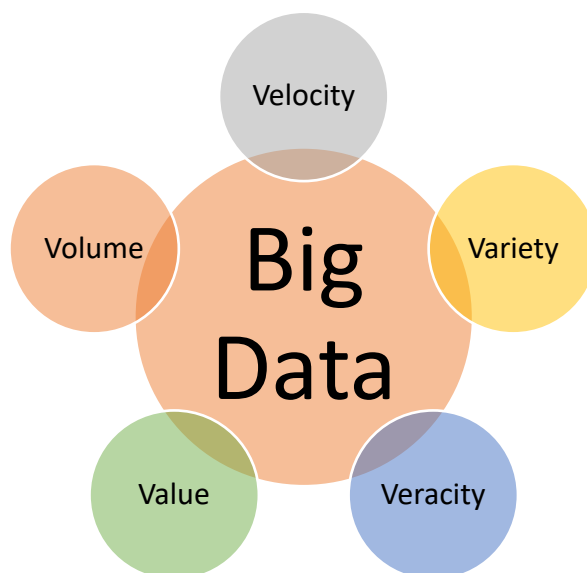


Рисунок 1. Критерии больших данных

Классификация основных проблем безопасности при использовании технологий больших данных (Рисунок 2).

<b>Классификация проблем безопасности Big DATA</b>	Хранение данных
	Поддельные данные
	Конфиденциальность данных
	Управление данными
	Сложность больших данных
	Соответствие требованиям безопасности в реальном времени
	Контроль доступа к данным
	Проблемы интеллектуального анализа данных
	Кража конфиденциальных данных сотрудников

Рисунок 2. Классификация основных проблем безопасности технологий больших данных

**Хранение данных.** Предприятия внедряют облачное хранилище данных, чтобы легко перемещать свои данные для ускорения бизнес-операций. Однако связанные с этим риски экспоненциальны с проблемами безопасности. Даже малейшая ошибка в управлении доступом к данным может привести к тому, что кто угодно получит множество конфиденциальных данных. В результате крупные технологические компании используют как локальные, так и облачные хранилища данных для обеспечения безопасности и гибкости.

В то время как критически важная информация может храниться в локальных базах данных, менее важные данные хранятся в облаке для простоты использования. Однако для реализации политик безопасности в локальных базах данных компаниям требуются специалисты по кибербезопасности. Хотя это увеличивает стоимость управления данными

в локальных базах данных, компании не должны воспринимать риски безопасности как должное, сохраняя все данные в облаке.

**Поддельные данные.** Генерация поддельных данных представляет серьезную угрозу для бизнеса, поскольку отнимает время, которое в противном случае можно было бы потратить на выявление или решение других насущных проблем. Существует больше возможностей для использования неточной информации в очень больших масштабах, поскольку оценка отдельных точек данных может быть сложной задачей для компаний.

Ложные флаги для поддельных данных также могут привести к ненужным действиям, которые потенциально могут снизить производительность или другие важные процессы, необходимые для ведения бизнеса. Один из способов избежать этого — обеспечить критическое отношение компаний к данным, над которыми они работают, для улучшения бизнес-процессов. Идеальным подходом является проверка источников данных путем периодической оценки и оценки моделей машинного обучения с использованием различных наборов тестовых данных для поиска аномалий.

**Конфиденциальность данных.** Конфиденциальность данных является большой проблемой в этом цифровом мире. Он направлен на защиту личной или конфиденциальной информации от кибератак, взломов и преднамеренной или непреднамеренной потери данных. Компании должны следовать более строгим принципам конфиденциальности данных с помощью служб управления доступом в облаке, включая очень строгое соблюдение конфиденциальности, чтобы усилить защиту данных. Лучше всего следовать нескольким правилам наряду с внедрением одной или нескольких технологий защиты данных. Общие правила включают в себя знание ваших данных, больший контроль над вашими хранилищами данных и резервным копированием, защиту вашей сети от несанкционированного доступа, проведение регулярных оценок рисков и регулярное обучение пользователей принципам конфиденциальности и безопасности данных.

**Управление данными.** Нарушение безопасности может иметь разрушительные последствия для бизнеса, включая уязвимость важной бизнес-информации для полностью скомпрометированной базы данных. Развертывание высокозащищенных баз данных жизненно важно для обеспечения безопасности данных на всех уровнях. Превосходная система управления базами данных поставляется с различными элементами управления доступом. Хотя рекомендуется следовать жестким и строгим методам физической безопасности, еще более важно соблюдать обширные программные меры безопасности для защиты хранения данных. Вот несколько методов эффективного достижения этой цели: шифрование данных, сегментация и разделение данных, защита на ходу и внедрение доверенного сервера. Кроме того, некоторые инструменты безопасности могут интегрироваться с базами данных для автоматического отслеживания обмена данными и уведомления предприятий о компрометации данных.

**Контроль доступа к данным.** Контроль того, какие данные пользователи могут просматривать или редактировать, позволяет компаниям обеспечивать не только целостность данных, но и сохранять их конфиденциальность. Но управлять контролем доступа непросто, особенно в крупных компаниях с несколькими тысячами сотрудников. Однако переход от локальных решений к облачным службам упростил процесс работы с Identity Access Management (IAM). IAM выполняет работу по управлению потоком данных посредством идентификации, аутентификации и авторизации. Соблюдение соответствующих стандартов ISO является хорошей отправной точкой для обеспечения того, чтобы организации соответствовали лучшим практикам IAM.

**Сложность больших данных.** Одной из самых больших угроз для больших данных является разнообразие наборов данных, используемых в любой момент времени. Эта информация может быть структурированной или неструктурированной, а ее источниками могут быть мобильные устройства, серверы, файлы электронной почты,

облачные приложения и многое другое. Данные становятся все труднее защитить, поскольку они становятся более сложными, поэтому важно использовать проверенную службу извлечения, преобразования и загрузки (ETL) для повышения совместимости данных.

**Соответствие требованиям безопасности в реальном времени.** В идеале хранилище больших данных должно включать инструменты, работающие в режиме реального времени, которые обеспечивают соответствие требованиям безопасности. Эти ресурсы постоянно контролируют процессы, чтобы гарантировать, что компания принимает надлежащие меры для защиты данных. Однако во время работы эти инструменты создают свои собственные большие объемы данных, которые затем необходимо обработать. Нарушение этих данных может быть столь же пагубным, как хакеры, получающие доступ к другим вашим базам данных.

Несмотря на то, что они создают множество дополнительных данных, которые необходимо защищать, эти инструменты обеспечения соответствия требованиям в режиме реального времени необходимы. Убедитесь, что вы изучаете свои варианты и выбираете с умом. Найдите вариант, в котором есть инструменты, предназначенные для минимизации потенциальных ложных срабатываний. Одно ложное предупреждение о нарушении может привести к напрасной трате ресурсов и возможности упустить из виду фактические нарушения. [6]

**Проблемы интеллектуального анализа данных.** Интеллектуальный анализ данных — это мощный инструмент, помогающий лучше понять и использовать данные, которыми владеет компания. Несмотря на то, что этим процессом занимаются профессионалы, он все равно может создавать проблемы с безопасностью больших данных, которые нельзя упускать из виду. Крайне важно, чтобы контролировать информацию, к которой могут получить доступ ИТ-специалисты при проведении интеллектуального анализа данных. Хотя доступ к возрастным и географическим демографическим данным может помочь выявить бесценные закономерности, например, обычно нет реальной причины предоставлять им доступ к информации о кредитных картах, номерам социального страхования или другим конфиденциальным данным.

**Кража конфиденциальных данных сотрудников.** Передовая культура данных позволила каждому сотруднику иметь определенный уровень критически важной бизнес-информации. Хотя это способствует демократизации данных, риск того, что сотрудник преднамеренно или непреднамеренно упустит конфиденциальную информацию, высок.

Кража конфиденциальных данных сотрудников распространена не только в крупных технологических компаниях, но и в стартапах. Чтобы избежать кражи сотрудников, компании должны внедрить правовую политику наряду с защитой сети с помощью виртуальной частной сети. Кроме того, компании могут использовать «Настольный компьютер как услугу» (DaaS), чтобы исключить функциональные возможности данных, хранящихся на локальных дисках. [2]

**Заключение.** Основные преимущества, которые бизнес может извлечь при использовании технологии больших данных:

Большие данные позволяют компаниям улучшать свои продукты и создавать индивидуальный маркетинг, получая всесторонний обзор поведения и мотивации своих клиентов.

Это позволяет предприятиям или поставщикам услуг отслеживать мошеннические действия в режиме реального времени, выявляя необычные модели и поведение с помощью прогнозной аналитики.

Он повышает эффективность цепочки поставок, собирая и анализируя данные, чтобы определить, доходят ли продукты до места назначения в желаемых условиях, чтобы привлечь интерес клиентов.

Прогнозный анализ позволяет компаниям сканировать и анализировать каналы социальных сетей, чтобы понять настроения клиентов.

Компании, которые собирают большие объемы данных, имеют больше шансов изучить неиспользованную область, а также провести более глубокий и подробный анализ в интересах всех заинтересованных сторон.

Чем быстрее и лучше бизнес понимает своего клиента, тем большую выгоду он извлекает. Большие данные используются для обучения моделей машинного обучения выявлению закономерностей и принятию обоснованных решений с минимальным вмешательством человека или без него.

Аналитики данных используют различные типы данных, прежде всего, для принятия более эффективных и эффективных бизнес-решений за счет понимания поведения клиентов и моделей их покупок. [3,4]

Но, сложность данных с точки зрения их структуры, источника, места хранения, формата, типа устройства и т. д. является одной из ключевых проблем безопасности, связанных при внедрении технологий больших данных. В сочетании с разнообразием процессов, происходящих с большими данными — хранением, очисткой, маскировкой и т. д. — управление данных информационных ресурсов является сложной задачей. Отсутствие мер безопасности при хранении и обработке больших данных может привести к утечке данных.

Обеспечение безопасности и конфиденциальности бизнес-данных и конфиденциальной информации клиентов являются основными вопросами безопасности при работе с большими данными. Поэтому необходимо повысить уровень устойчивости системы к кибератакам, настроить средства автоматической очистки данных, маскирования данных и санации документов, установить обязательную авторизацию для сотрудников и настроить постоянный мониторинг состояния системы.

#### Список использованной литературы

1. Статьев В.Ю. Информационная безопасность на пространстве «больших данных», DOI: 10.36724/2072-8735-2022-16-4-21-28
2. <https://www.integrate.io/blog/big-data-security-concerns/> (Дата обращения 24.09.2022)
3. <https://hevodata.com/learn/big-data-security/> (Дата обращения 24.09.2022)
4. Безопасность информационного пространства в контексте больших данных. Курбацкий В.А. В сборнике: Комплексная Защита Информации. материалы XXII научно-практической конференции. 2017. С. 250-253.
5. Информационная безопасность персональных данных граждан в условиях технологий "big data" ("большие данные"). Попова А.В., Воронова Ю.С. Аллея науки. 2020. Т. 1. № 10 (49). С. 683-686.
6. Проблемы обеспечения информационной безопасности в системах управления большими данными. Полтавцева М.А. В сборнике: XIII Всероссийское совещание по проблемам управления ВСПУ-2019. Сборник трудов XIII Всероссийского совещания по проблемам управления ВСПУ-2019. Институт проблем управления им. В.А. Трапезникова РАН. 2019. С. 2606-2611.
7. Технологии Больших Данных (Big Data) В Области Информационной Безопасности. Петренко А.С., Петренко С.А. В сборнике: The 2018 Symposium on Cybersecurity of the Digital Economy (CDE'18). Вторая международная научно-техническая конференция. 2018. С. 248-255.

## TEACHING 3D MODELING IN ART SCHOOLS

**Kystaubayeva Assel Samatkyzy**

student, faculty of Arts and Education

Speciality: Preparation of a teacher in artistic work and drawing

South Kazakhstan State Pedagogical University

Republic of Kazakhstan, Shymkent, e-mail: [assel.samatova.96@mail.ru](mailto:assel.samatova.96@mail.ru)

Supervisor: Nebessayeva Zhanar Orynbekovna

### **Annotation**

The main subjects in the state art schools are: painting, drawing, composition, sculpture and history of art. These subjects are necessary for study, as they give creative development and development of individual handwriting to each student. Nowadays, 3D projects in such areas as architecture, design and animation are very popular. Every year, 3D modeling becomes more in demand in certain areas. Students who study in areas such as architecture and architectural design take private courses in 3D visualization, as not all skills are taught in higher education institutions. Most often, children who are fond of creativity and take lessons in art schools, after graduating from grade 11, go to enter creative specialties. If you teach children from school age the basic skills of 3D visualization, this will simplify further education in higher educational institutions. Being further students can develop their talents and create projects.

**Keywords:** 3D modeling, teaching, education, art schools, computer graphics.

### **Purpose**

The purpose of this article is to create additional subjects in art schools, to develop new skills in computer modeling, teaching 3D modeling to children from 10 years old, mastering basic skills for further development in the chosen field, motivation for young artists to create desired works in digital format.

### **1. 3D modeling in art school: methodology, practice and education.**

3D visualization can be used in architecture design, graphic design and animation. Young artists have their own preferences in which directions they would like to develop. For each direction, subjects are created where they are taught from scratch, starting with easy programs for mastering 3D visualization.

Modern information technologies make it possible to display various visual information in three-dimensional space. Most information about the world around us is perceived with the help of the organs of vision. In order to achieve maximum accuracy in the perception of information, it is necessary to provide visual perception that is as close to reality as possible. Modeling training is of great importance both for achieving personal results, including the development of creative abilities. Children are exposed to 3D from a very young age through models and special effects used in movies and TV shows using computer graphics. And recreating your characters in 3D is an exciting creative process for kids, limited only by their imagination.

The use of 3D modeling in education provides teachers with new learning tools. These tools make learning easier for students to comprehend, increase their motivation, and speed up the assimilation of large amounts of knowledge.

In fact, 3D helps to gain experience in several disciplines, this makes it necessary for teachers to take into account emerging new technologies for the purpose of their subsequent competent implementation in the educational process.

In the process of assimilation of educational material, according to the operational concept of intelligence by J. Piaget, any information that perceived by a person, goes through four stages:

- sensorimotor (sensory perception);
- symbolic stage (figurative convolution of the sensory-logical information);



- logical stage (discursive-logical comprehension information);
- linguistic stage (accommodation of information in the mind through the word-image developed at the previous stages)

In the usual lecture form of the lesson, as in traditional printed textbooks, the physiologically necessary sensorimotor stage of information perception is practically absent, since the educational material is presented at the lexical level. This is one of the reasons for the complexity of the perception of information. In order to achieve maximum accuracy in the perception of information, it is necessary to provide visual perception that is as close to reality as possible. Compliance with the natural order of perception and processing of information leads to saving time in the learning process. When educational material is presented with the help of visual images, various channels are involved in the process of perception: hearing, vision, etc.

A figurative visual representation of an object can be obtained by studying the object itself or its physical model, as well as their display obtained by multimedia means (electronic posters, video clips, animation, etc.), including computer (virtual) 3D images. An important feature of three-dimensional models is the ability to change the properties of both the constituent elements of the model and the entire model as a whole, depending on the needs of the developer. Thanks to this, you can change the location of individual elements in space, change their appearance, use additional objects, and so on. You can show not only static graphics, but also complex spatial animation, as well as processes (including hidden ones) that occur both with the object and inside it. And this leads not only to a significant increase in the visibility of the material, but also to a significant saving of time during training. In addition, modern technologies make it possible to obtain photorealistic graphics based on such a model, which is not inferior in quality to photo and video materials.

The advantages of learning using information technology in the form of using 3D models are obvious. Unlike flat static images, these models are interactive: you can choose any point of view, make any transformations, with a minimum of effort. Computer 3D models in the process of studying can either be disassembled into separate elements, or assembled into a single product.

The interactivity of 3D computer models means that students and teachers are given the opportunity to actively interact with these tools. Interactivity means the presence of conditions for an educational dialogue-interaction, one of the participants of which is a computer model.

A 3D laboratory equipped with affordable, simple and inexpensive systems will allow both teachers and students to reproduce three-dimensional models using computer programs at the touch of a button. 3D laboratories are actively being introduced into the education system, for the following reasons:

- helps to involve children in research work, makes the learning process interesting and understandable;
- helps to understand the intricacies of the future profession at the training stage;
- stimulation of creative activity of each child;
- significantly increased the level of training of specialists.

## **2. An example of several art schools abroad, where they teach 3D modeling.**

### **School of Visual Arts** in New York City, USA.

Founded in 1947 as the School of Cartoonists and Illustrators. Silas H. Rhodes and illustrator Burn Hogarth co-founded the School of Cartoonists and Illustrators, where New York-based professionals work in the arts as educators, a practice that continues to this day. In 1956, C&I renamed the School of Visual Arts. The school teaches 3D animation, architectural visualization, game design, product design, visual effects, etc. Students are provided with all the necessary equipment for studying programs and mastering skills in the field in which they wish to develop.

### **AnimSchool** in Provo, Utah, USA.

AnimSchool is an accredited online school for teaching 3D animation and modeling. The school has over 30 instructors currently employed by studios such as Blue Sky Studios, Dreamworks, Disney, Sony and Pixar.

AnimSchool was founded in 2010 by David Gallagher, an animator, rigger and former character development supervisor at Blue Sky Studios. Students learn to specialize in a particular field and also learn support skills so that students can prepare well for the 3D workforce.

**Creative Career Academy** in Roswell, Georgia, USA.

Creative Career Academy is a Christian school that offers 21st century education in art, technology and design. Students can receive a professional education that branches into many aspects of life skills that can be applied to college and future careers. The school accepts students from 8 years old. Children can learn 3D animation and 2D animation.

In the art schools of the Republic of Kazakhstan, they mainly teach the skills of drawing and using paints, in some schools there are lessons in 2D computer graphics. There are private courses in 3D modeling, but mostly these courses are for adults who want to change careers or university students who want to learn skills in their specialty.

### **The conclusion**

The inclusion of studying the basics of 3D modeling in art schools opens up great opportunities for students to create fundamentally new products of labor, master new heights in the study of modern technologies. Students gain practical knowledge of modeling and parametric design, create their own engineering projects, develop engineering skills and ensure their competitiveness in professional fields of technical orientation. The development of a program for such a course is an urgent need due to rapidly developing, rapidly changing technologies, existing programs quickly become obsolete.

### References

1. Пиаже, Ж. Психология интеллекта — СПб.: Питер, 2004. — 192 с. — (Серия «Психология-классика»).
2. Возможности 3D-технологий в образовании [Электронный ресурс]. – Режим доступа: URL: <https://cyberleninka.ru> свободный.
3. School of Visual Arts | SVA | New York City | Fine Arts and Graphic Design School in New York City
4. Online 3D Animation School | Accredited Animation Programs | Learn at AnimSchool
5. About Us | Creative Career Academy ([u-create.org](http://u-create.org))

## ЭНЕРГИЯ ВЫЧИСЛЕНИЙ: ОПТИМИЗАЦИЯ ЭНЕРГИИ И ПРОИЗВОДИТЕЛЬНОСТИ ПРОГРАММ

Алексей Ластовецкий

Университетский Колледж Дублина, Дублин, Ирландия, [alexey.lastovetsky@ucd.ie](mailto:alexey.lastovetsky@ucd.ie)

В 2020 году на информационные и коммуникационные технологии приходилось 7% мирового потребления электрической энергии. Ожидается, что эта цифра вырастет до 21% к 2030 году [1]. Такое развитие событий делает энергетическую эффективность цифровых платформ одним из главных технологических вызовов современности.

Различают два подхода к решению проблемы энергетической эффективности вычислений – аппаратный и программный. Первый подход имеет дело с аппаратными решениями на уровне интегральных схем, направленными на разработку вычислительных устройств, потребляющих как можно меньше электроэнергии.

Второй подход имеет дело с разработкой энергетически эффективного программного обеспечения. В зависимости от уровня принимаемых решений его можно в свою очередь разбить на подходы системного уровня и прикладного уровня.

Подход системного уровня пытается оптимизировать окружение выполнения, а не саму выполняемую прикладную программу. Этот подход в настоящее время является основным и использует технологии динамического масштабирования напряжения и частоты (Dynamic Voltage and Frequency Scaling, DVFS), динамического управления мощностью (Dynamic Power Management, DPM) и планирования с учетом энергии для оптимизации энергетической эффективности выполнения прикладной программы.

Методы оптимизации энергии прикладного уровня используют целевые переменные уровня прикладной программы и направлены на оптимизацию прикладной программы, а не окружения выполнения [2].

В идеальном мире высокопроизводительных вычислений нет места для оптимизации энергии на уровне прикладной программы. Дело в том, что идеальный мир линеен и однороден. Линейность означает, что как энергия, потребленная процессором во время выполнения программы, так и время ее выполнения являются линейными функциями объема вычислений. Однородность означает, что все процессоры в вычислительной системе идентичны. С математической точки зрения, это означает, что при любом распределении заданного объема вычислений между идентичными линейными параллельными процессорами потребленная динамическая энергия будет той же самой, а распределение, балансирующее вычислительную нагрузку процессоров, всегда будет оптимальным как по времени выполнения, так и с точки зрения потребленной энергии.

Допущение *однородности*, приемлемое во времена, когда высокопроизводительные платформы строились из идентичных процессоров, неприемлемо в современном мире, который становится все более *неоднородным*. Свыше 30% систем из списка Top500 самых мощных машин в мире являются неоднородными, интегрируя различные процессоры общего назначения (CPU) и графические ускорители (GPU). Последствия неоднородности для оптимизации энергии и производительности прикладных программ поистине фундаментальны. Глубокое и всестороннее исследование этих последствий было опубликовано совсем недавно [3]. Оно обнаружило, что в случае  $p$  неоднородных линейных процессоров, существует бесконечное число оптимальных по Парето решений, распределяющих заданный объем вычислений  $n$  между процессорами, а множество Парето в системе координат “время выполнения – энергия” представляет собой непрерывную кусочно-линейную функцию, состоящую из  $p-1$  линейных звеньев. В исследовании были предложены эффективные (сложности  $O(p^3 \times \log_2 n)$ ) точные

алгоритмы построения множеств Парето как в пространстве “время выполнения – динамическая энергия”, так и в пространстве “время выполнения – полная энергия”, а также точный алгоритм, находящий за линейное время  $O(p)$  оптимальное распределение вычислений между процессорами для каждой точки найденного множества Парето. Обнаружено, что за исключением одного, все остальные оптимальные по Парето решения распределяют вычислительную нагрузку несбалансированно. Единственное сбалансированное решение минимизирует время выполнения. Решение, использующее для вычислений лишь один, самый эффективный с точки зрения потребления энергии процессор, также будет оптимальным по Парето и будет минимизировать динамическую энергию вычислений. Хотя здесь мы говорили только о линейных процессорах, в исследовании [3] решена более общая проблема, допускающая не только линейные, но и любые непрерывные, монотонно возрастающие функции времени выполнения.

Допущение *линейности*, приемлемое для одноядерных процессоров, перестает быть приемлемым для многоядерных процессоров и ускорителей. Их *нелинейность* была показана экспериментально на примере тщательно оптимизированных для многоядерных устройств программ [2]. Построенные дискретные нелинейные функции времени выполнения и потребленной энергии в зависимости от объема вычислений оказались крайне нерегулярными, не поддающимися аппроксимации гладкими функциями. Степень их нелинейности заметно возрастала с увеличением числа ядер, используемых в вычислениях. Вариации времени выполнения и энергии оказались очень существенными (средние вариации для широко используемой программы двумерного дискретного быстрого преобразования Фурье на 24-ядерном процессоре превысили 70%). При этом относительная величина вариаций не уменьшалась с увеличением размера решаемой задачи. Воспроизводимость вариаций и их неумещающаяся с увеличением объема вычислений величина доказывают, что они вызваны не случайными факторами и флуктуациями в среде выполнения, а системной сложностью интеграции ресурсов в многоядерных вычислительных устройствах. Существенно, что вариации энергии и времени выполнения в зависимости от объема вычислений очень часто не коррелируют [4]. Оказалось, что при увеличении объема вычислений время выполнения может уменьшаться, а энергия расти, или наоборот, время расти, а энергия уменьшаться.

Последствия нелинейности *однородных* многопроцессорных систем для минимизации времени выполнения программы [2], энергии ее выполнения [2], а также двухкритериальной минимизации как времени, так и энергии [4], были недавно изучены. Оказалось, что в случае  $p$  однородных процессоров как оптимальные по времени, так и оптимальные по энергии решения не обязательно являются сбалансированными [2].

Были предложены точные эффективные алгоритмы сложности  $O(m^2 \times p^2)$  ( $m$  - мощность дискретного множества точек, в которых заданы функции энергии и времени выполнения), находящие распределения вычислений, минимизирующие их время и энергию [2]. Был также предложен точный эффективный алгоритм сложности  $O(m^2 \times p^2)$ , решающий двухкритериальную проблему минимизации времени и энергии [4]. Число оптимальных по Парето решений, найденных этим алгоритмом для гибридных параллельных программ, использующих высокоэффективный библиотечный код, оказалось значительным. В то время, как сбалансированность вычислительной нагрузки официально считается главным критерием оптимальности программы, сбалансированные решения оказались очень далеки от множества решений, оптимальных по Парето [4].

Также были разработаны алгоритмы оптимизации энергетической эффективности параллельных программ, выполняющихся на  $p$  *нелинейных неоднородных* процессорах. Были предложены точные эффективные алгоритмы сложности  $O(m^3 \times p^3)$ , минимизирующие время [5], динамическую энергию [6], полную энергию [6], а также решающие задачи двухкритериальной минимизации времени и динамической энергии, и времени и полной энергии [7].

Методам [2-7] для оптимизации энергетической эффективности программ нужны профили энергии и времени выполнения всех компонент гибридной программы в качестве исходных данных. Для построения таких профилей требуются методы точного измерения энергии и времени отдельных компонент программы, выполняющихся на разных устройствах гибридного сервера. Часы, встроенные в процессоры и ускорители, можно использовать для измерения времени. Измерение же энергии выполнения отдельных компонент гибридной параллельной программы является нерешенной проблемой. Без ее решения, для построения полного энергетического профиля гибридной программы на сервере с  $p$  неоднородными процессорами потребуется  $m^p$  измерений на уровне сервера (для всех возможных комбинаций распределения вычислительной нагрузки) вместо  $m \times p$  при наличии решения (в этом случае,  $p$  профилей отдельных компонент можно использовать для расчета энергии всех возможных распределений нагрузки).

Существуют три метода измерения энергии выполнения отдельных компонент гибридной программы [8]. Первый использует внешний измеритель мощности всей системы. Это наиболее точный, но очень затратный метод. Получение одного измерения с достаточной точностью может потребовать часов и даже дней [8].

Второй метод использует сенсоры мощности, встроенные в интегральные схемы процессоров и ускорителей, такие как Intel RAPL (Running Average Power Limit), Intel Xeon Phi SMC (System Management Controller), AMD APM (Application Power Management), Nvidia NVML (Nvidia Management Library). Быстрые и недорогие, сенсоры, однако, плохо документированы и очень неточны [8]. Эксперименты с популярными библиотечными программами и основными процессорами и ускорителями показали, что построенные с помощью сенсоров профили энергии качественно неточны. Это означает, что даже их калибровка не сможет помочь, так как сенсоры могут регистрировать увеличение потребления энергии в случаях, когда оно уменьшается, и наоборот. Эксперименты показали, что использование полученных с помощью RAPL профилей энергии для оптимизации энергетической эффективности параллельной программы умножения матриц привело к существенному увеличению потребления энергии (от 37% до 84% в зависимости от размера матриц) по сравнению с использованием точных профилей, построенный с помощью измерителей мощности [8].

Третий метод использует аналитические модели энергии с различными измеряемыми индикаторами активности вычислительных устройств в качестве переменных. Их точность пока недостаточно высока, но постепенно увеличивается. Этот метод – единственная практическая альтернатива использованию измерителей мощности. Наиболее популярные и изученные модели используют в качестве переменных счетчики слежения за производительностью (Performance Monitoring Counters, PMC). PMC – это специальный регистр для хранения числа определенной программной или аппаратной активности, изначально предназначенный для низкоуровневого анализа эффективности программ. Количество разных событий, за которыми можно следить, очень велико (164 для Intel Haswell, 396 для Intel Skylake), а количество регистров для их хранения – очень мало (3–4 для одного выполнения программы). Кроме того, события специфичны для каждой архитектуры и поэтому непереносимы.

Наиболее популярные модели динамической энергии – линейные модели, использующие PMC-переменные. Основные проблемы с этими моделями связаны с большим количеством событий – кандидатов в переменные, огромными временными затратами на кодирование и сбор событий, и недостаточной переносимостью событий. Существующие методы выбора событий для модели либо используют все события, чтобы учесть все потенциальные источники потребления энергии, либо основаны на статистических инструментах, таких как корреляция и анализ главных компонент, либо используют советы экспертов и интуицию. Проблема оптимального выбора небольшого числа событий для энергетических моделей далека от решения. Было предложено множество моделей, но все они оказались недостаточно точными [9]. Хотя авторы этих

моделей и сообщают об их высокой точности, они обычно имеют в виду точность предсказания не динамической, а полной энергии, да еще и в системах с очень высокой статической мощностью. Кроме того, большинство опубликованных результатов, проанализированных в обзоре [9], оказались невоспроизводимыми. Для воспроизводимых же моделей лучшая подтвержденная средняя ошибка предсказания динамической энергии в экспериментах на Intel Haswell оказалась в пределах 90–100% [9].

Одним из источников неточности этих моделей, открытым в 2017, является то, что многие события, часто выбираемые в качестве переменных модели, оказались неаддитивными в современных многоядерных процессорах [10]. Энергия аддитивна. Действительно, энергия последовательного выполнения программ **A** и **B** будет равна сумме энергий выполнения каждой из программ,  $E_{AB} = E_A + E_B$ . Поэтому любая переменная **x** в линейной энергетической модели должны быть аддитивна,  $x_{AB} = x_A + x_B$ . Хотя, согласно их описанию, все события аддитивны, на практике многие из них неаддитивны, причем существенно, с отклонениями от аддитивности достигающими 200% [10], а число неаддитивных событий быстро растет с увеличением числа ядер [10].

Другим источником неточности моделей является нарушение ими основных законов сохранения энергии [10-11], такие как неоднородные модели динамической энергии, отрицательные коэффициенты в аддитивных членах, а также нелинейность наиболее сложных моделей (включая модели машинного и глубокого обучения). Недавно было доказано, что для того, чтобы не нарушать основные законы сохранения энергии, модели динамической энергии, использующие РМС-переменные, должны быть линейными однородными функциями с положительными коэффициентами, [11].

Таким образом, точность моделей динамической энергии, использующих РМС-переменные, может быть увеличена путем исключения неаддитивных переменных и соблюдения закона сохранения энергии. Применение этого подхода уже позволило снизить ошибку предсказания этих моделей до 25–30% [12]. Модели, комбинирующие РМС-переменные с другими измеряемыми переменными, такими как утилизация процессора, снизили ошибку предсказания до 10–20% для некоторых популярных программ [13]. Хотя еще многое предстоит сделать, эти результаты обнадеживают.

Модели энергии для ускорителей находятся в зачаточном состоянии из-за недостаточной аппаратной поддержки отслеживания событий, потребляющих энергию.

Таким образом, энергетическую эффективность современных цифровых платформ можно значительно увеличить путем оптимизации программного обеспечения. Наиболее фундаментальные алгоритмы оптимизации уже разработаны и поэтому не являются технологическим вызовом. Настоящий вызов – это точное и эффективное измерение энергии компонент программы, особенно компонент, выполняющихся ускорителями.

## References.

- [1] A. S. G. Andrae, "New perspectives on Internet electricity use in 2030," Eng. Appl. Sci. Lett., vol. 3, pp. 19–31, Jun. 2020.
- [2] A. Lastovetsky and R. Reddy, "New Model-based Methods and Algorithms for Performance and Energy Optimization of Data Parallel Applications on Homogeneous Multicore Clusters", IEEE Transactions on Parallel and Distributed Systems 28(4):1119-1133, 2017.
- [3] H. Khaleghzadeh, R. Manumachu, and A. Lastovetsky, "Efficient exact algorithms for continuous bi-objective performance-energy optimization of applications with linear energy and monotonically increasing performance profiles on heterogeneous high performance computing platforms", Concurrency and Computation: Practice and Experience, 2022.
- [4] R. Reddy and A. Lastovetsky, "Bi-Objective Optimization of Data-Parallel Applications on Homogeneous Multicore Clusters for Performance and Energy", IEEE Transactions on Computers 67(2):160-177, 2018.

- [5] H. Khaleghzadeh, R. Manumachu, and A. Lastovetsky, "A novel data-partitioning algorithm for performance optimization of data-parallel applications on heterogeneous HPC platforms", *IEEE Transactions on Parallel and Distributed Systems* 29(10):2176-90, 2018.
- [6] H. Khaleghzadeh, M. Fahad, R. Manumachu, and A. Lastovetsky, "A novel data partitioning algorithm for dynamic energy optimization on heterogeneous high-performance computing platforms", *Concurrency and Computation: Practice and Experience* 32(21), 2021.
- [7] H. Khaleghzadeh, M. Fahad, A. Shahid, R. Manumachu, and A. Lastovetsky, "Bi-objective Optimization of Data-Parallel Applications on Heterogeneous HPC Platforms for Performance and Energy through Workload Distribution", *IEEE Transactions on Parallel and Distributed Systems* 32(3):543-560, 2021.
- [8] M. Fahad, A. Shahid, R. R. Manumachu, and A. Lastovetsky, "A Comparative Study of Methods for Measurement of Energy of Computing", *Energies* 12(11), pp. 42, 2019.
- [9] K. O'Brien, I. Petri, R. Reddy, A. Lastovetsky, and R. Sakellariou, "A survey of power and energy predictive models in HPC systems and applications", *ACM Computing Surveys*, vol. 50, issue 3, 2017.
- [10] A. Shahid, M. Fahad, R. R. Manumachu, and A. Lastovetsky, "Additivity: A selection criterion for performance events for reliable energy predictive modeling", *Supercomputing Frontiers and Innovations*, vol. 4, pp. 50–65, 2017.
- [11] A. Shahid, M. Fahad, R. Manumachu, and A. Lastovetsky, "Energy Predictive Models of Computing: Theory, Practical Implications and Experimental Analysis on Multicore Processors", *IEEE Access* 9:63149-63172, 2021.
- [12]. A. Shahid, M. Fahad, R. R. Manumachu, and A. Lastovetsky, "Improving the Accuracy of Energy Predictive Models for Multicore CPUs Using Additivity of Performance Monitoring Counters", *HerteroPar* 2019.
- [13]. A. Shahid, M. Fahad, R. R. Manumachu, A. Lastovetsky, "Improving the accuracy of energy predictive models for multicore CPUs by combining utilization and performance events model variables", *Journal of Parallel and Distributed Computing* 151:38–51, 2021.

# СОВЕРШЕНСТВОВАНИЕ АЛГОРИТМОВ ТОПОЛОГИЧЕСКОЙ МОДЕЛИ УСТАНОВИВШЕГОСЯ РЕЖИМА ЭЛЕКТРОЭНЕРГЕТИЧЕСКИХ СИСТЕМ

*Ахметбаев Д.С., Джандигулов А.Р., Быстрова С.В., Бежежанова Ш.С.*  
Казахский агротехнический университет им. С. Сейфуллина, г. Астана,  
[axmetbaev46@mail.ru](mailto:axmetbaev46@mail.ru)

*Евразийский национальный университет им. Л. Гумилева, г. Астана, [abeked@mail.ru](mailto:abeked@mail.ru)*  
*Екибастузский инженерно-технический институт им. К. Сатпаева, e-mail:*  
[bystrova.75@mail.ru](mailto:bystrova.75@mail.ru)

*Евразийский национальный университет им. Л. Гумилева, г. Астана, :*  
[begezhanova.sh@mail.ru](mailto:begezhanova.sh@mail.ru)

## Аннотация

В работе совершенствуются алгоритмы формирования установившегося режима электроэнергетических систем, разработанные на основе топологической модели. Топологическая модель разработана на основе теории направленных графов с применением матрицы коэффициентов токораспределения. Алгоритмы формирования установившегося режима предусматривают ведение расчетов относительно всех независимых узлов сложной сети энергосистемы, что приводит к увеличению количества операции итерационных процессов. В работе предложено преобразование топологической модели системы, при котором в процессе решения преобразованной модели итерация производится только по узлам, в которых имеются ненулевые мощности нагрузки или генерации. Остальные узловые напряжения просто вычисляются через напряжения вышеуказанных узлов. Полученное преобразование позволяет сократить не только количество операций в каждой итерации, но и уменьшить само число итераций.

**Ключевые слова:** электроэнергетическая система, топологическая модель, установившийся режим, итерационные методы.

## Введение

Методологические проблемы анализа электроэнергетических систем (ЭЭС) охватывают широкий круг вопросов и задач, связанных с управлением ее режимами, прежде всего это исследования по формированию уравнений установившихся режимов ЭЭС. Сложность решения задачи управления ЭЭС с учетом ограничений, учитывающих экономические и надежность показатели, объясняется большой размерностью нелинейных уравнений, описывающих ее режим и состав телеизмерений и связанных с этим многообразием математических моделей режимов [1]. Применение матричного метода расчета в теории электрических сетей, разработанного в [2] и получившего дальнейшее развитие в работах [3, 4] однозначно определило направление исследования по формированию уравнений установившегося режима ЭЭС на основе обобщенных параметров ее схемы замещения. Топологическое содержание  $\underline{Z}$  – узловых обобщенных параметров позволило принять его в основу математических моделей и методов расчета сложных схем ЭЭС.

Основы топологии схемы замещения электрических сетей были заложены в классических работах Кирхгофа и Максвелла. Ими впервые введены понятия деревьев и получены топологические выражения для определителей матрицы проводимостей узлов  $\underline{Y}_y$  и матрицы сопротивлений контуров  $\underline{Z}_k$ .

В современной литературе эти топологические выражения для определителей  $\Delta_y$  и  $\Delta_k$  записанные через величины ненаправленных деревьев и дополнений, известны как производные законов Кирхгофа [5].

Практическое применение и развитие топологических методов в анализе электрических цепей стало возможным после опубликования работ Персиваля, Сешу, Мэзона, Коутса и других. В работах [6,7] Персиваля В.С. выведены понятия общей ветви, отображающей группу деревьев и пары обобщенных ветвей, отображающей группу 2 –



деревьев графа и даны теоремы для определения групп деревьев графа схемы, составленной из подсхем, связанных деревом параллельно и в виде контура. Эта теорема послужила основой при разработке, Ионкиным П.А., Соколовым А.А. [8,9], способов отыскания деревьев путем разложения исходного графа схемы на узловые пары, по ветвям и по узлу. Дальнейшее развитие топологического метода, применительно к задачам анализа установившихся режимов ЭЭС, получило в работах О.Т. Гераскина [10,11].

Для нахождения коэффициентов токораспределения электрических сетей, рассчитываются отношения суммы весов специфических деревьев к сумме значений весов всевозможных деревьев графа [11]. Сложность формирования матрицы коэффициентов токораспределения заключается в определении числителей топологических выражений путем деления сети на две части, с целью нахождения двух деревьев графа. Ахметбаевым Д.С. предложен аналитический подход к определению топологического содержания матрицы коэффициентов токораспределения на основе свойств возможных деревьев графа, без деления сети на две части [12,13]. В работе [14] реализован эффективный алгоритм направленного поиска и определения весов возможных деревьев графа без привлечения ранее определенных деревьев.

Исследования, проведенные с целью анализа методов и алгоритмов поиска возможных деревьев сложного графа, привело к разработке оптимальных алгоритмов на основе принципов диакоптики [15]. Основой идеей оптимизации является построение специальных классов деревьев. При этом в процессе группировки выделяются части будущих графов, являющиеся «родительскими» для групп графов и строятся соответствующие графы существенно меньшей размерности.

На основе топологического алгоритма коэффициентов токораспределения формируется матричное уравнение установившегося режима.

#### **Методы**

Узловые напряжения определяются по выражению, полученного путем аналитического преобразования известных уравнений электрического состояния сетей [16,17]:

$$U = U_0 + C^T Z_B C U_D^{-1} S(U) \quad (1)$$

где  $C$  –прямоугольная комплексная матрица коэффициентов токораспределения;  $Z_B$  – диагональная матрица сопротивлений ветвей;  $U_D$  – диагональная матрица узловых сопряженных напряжений;  $S(U)$  – вектор-столбец, зависящий от мощностей узловых нагрузок и генераторов и проводимостей линий передач;  $T$  - знак транспонирования матрицы.

Полученная система нелинейных уравнений установившегося режима электроэнергетических систем, разработана на основе топологической модели матрицы коэффициентов токораспределения. Данная система была предложена и применена авторами и представляет с собой  $(n - 1)$  комплексных нелинейных уравнений относительно  $(n - 1)$  комплексных значений узловых напряжений. Где  $n$  – общее число узлов в сети. Полученная система решается итерационными методами. По мере усложнения расчетной схемы ЭЭС количество операции итерационного процесса существенно возрастает, что снижает оперативность получения результатов расчета. Ниже предлагается усовершенствованный алгоритм организации итерационного процесса.

#### **Обсуждения**

Система (1) решается итерационными методами. В данной работе предложено преобразование рассматриваемой системы, при котором в процессе решения преобразованной системы итерация производится только по узлам, в которых имеются ненулевые мощности нагрузки или генерации. Остальные узловые напряжения просто вычисляются через напряжения вышеуказанных узлов.

В системе (1) вектор-столбец  $S(U)$  выражается формулой

$$S(U) = S_0 - j \cdot A \cdot |U|^2 \quad (2)$$

где  $S_0$  – вектор столбец заданных (постоянных) мощностей нагрузки и генерации,  $A$  – диагональная матрица с действительными элементами, вычисляемая через заданные коэффициенты проводимостей  $b$ .

Обозначим через  $M = C^T Z_B C$ . Подставим выражение (2) в (1).

$$U = U_0 + M\bar{U}_D^{-1}(S_0 - j \cdot A \cdot |U|^2)$$

Раскроем скобки и проведем преобразования

$$U = U_0 + M\bar{U}_D^{-1}S_0 - j \cdot M \cdot A \cdot U; \Rightarrow (E + j \cdot M \cdot A)U = U_0 + M\bar{U}_D^{-1}S_0; \Rightarrow U = (E + j \cdot M \cdot A)^{-1}(U_0 + M\bar{U}_D^{-1}S_0) \quad (3)$$

Систему (3) можно переписать в виде

$$U = V_0 + G\bar{U}_D^{-1}\bar{S}_0 \quad (4)$$

где  $V_0 = (E + j \cdot M \cdot A)^{-1}U_0$ ,  $G = (E + j \cdot M \cdot A)^{-1}M$ .

Покомпонентно можно записать в виде:

$$U_k = (V_0)_k + \sum_l G_{kl} \cdot (\bar{S}_0)_l / \bar{U}_l \quad (5)$$

Из системы уравнений (4-5) мы видим, что при решении итерационными методами, в правой части системы используются только те значения узловых напряжений, в которых есть ненулевые мощности генерации или нагрузки. То есть систему (4-5) можно разделить на две подсистемы: уравнения с ненулевыми мощностями генерации и/или нагрузки и уравнения без них. Итерационными методами решается только первая часть. А вторая часть является только формулами для вычисления оставшихся узловых напряжений.

Полученное преобразование позволяет сократить не только количество операций в каждой итерации, но и уменьшить само число итераций. В частности, при исследовании 5-и узловой тестовой схемы (см. рис.1), исходные данные для которой приведены в таблицах 1 и 2, вместо четырех комплексных уравнений исходной системы достаточно провести итерацию только для одного уравнения (сокращение на 75%).

Таблица 1. Исходные данные по узлам

№уз	Unom	Напряжение		Мощность нагрузки		Мощность генерации	
	кВ	фаза,град	модуль,кВ	P,МВт	Q,МВар	P,МВт	Q,МВар
1	500.0	68.734	484.529	0.00	0.00	5400.00	0.00
2	500.0	45.051	482.049	0.00	0.00	0.00	0.00
3	500.0	21.974	489.348	0.00	0.00	0.00	0.00
4	500.0	0.118	499.904	0.00	0.00	0.00	0.00
5	500.0	0.000	500.000	0.00	0.00	0.00	457.00

Таблица 2. Данные по ветвям

№ ветви	нач.	кон.	R, Ом	X, Ом	b, мСм (емк+,инд-)	Ктр
1	1	2	0.7000	17.5000	8.3200	
2	2	3	0.7000	17.5000	8.3200	
3	3	4	0.7000	17.5000	8.3200	
4	4	5	0.0000	0.1000	0.0000	

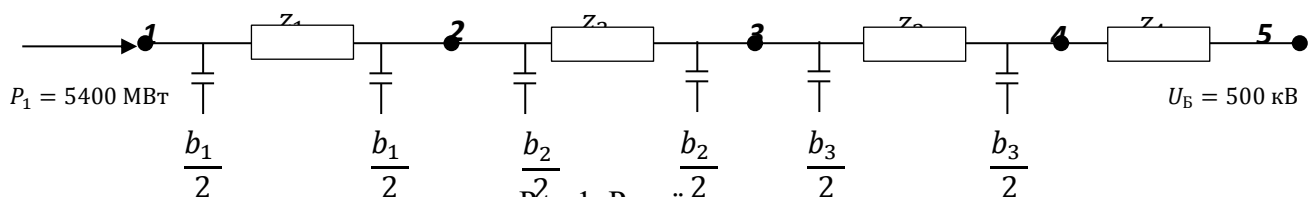


Рис.1. Расчётная схема

Матрицы  $M$  и вектор-столбец  $S_0$  в данном случае имеют вид:

$$M = \begin{pmatrix} 2.1 + 52.6j & 1.4 + 35.1j & 0.7 + 17.6j & 0.1j \\ 1.4 + 35.1j & 1.4 + 35.1j & 0.7 + 17.6j & 0.1j \\ 0.7 + 17.6j & 0.7 + 17.6j & 0.7 + 17.6j & 0.1j \\ 0.1j & 0.1j & 0.1j & 0.1j \end{pmatrix}$$

$$A = \begin{pmatrix} 4.16 & 0 & 0 & 0 \\ 0 & 8.32 & 0 & 0 \\ 0 & 0 & 8.32 & 0 \\ 0 & 0 & 0 & 4.16 \end{pmatrix} \cdot 10^{-3}, \quad S_0 = \begin{pmatrix} 5400 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$(E + j \cdot M \cdot A)^{-1} U_0 = \begin{pmatrix} 1232 - 64.767j \\ 01142 - 56.465j \\ 886.8 - 33.289j \\ 502.4 - 0.102j \end{pmatrix}$$

$$(E + j \cdot M \cdot A)^{-1} M \cdot \text{diag}(\bar{S}_0) = \begin{pmatrix} (0.4754 + 5.676j) \times 10^5 & 0 & 0 & 0 \\ (0.3865 + 4.319j) \times 10^5 & 0 & 0 & 0 \\ (0.2161 + 2.335j) \times 10^5 & 0 & 0 & 0 \\ 69.948 + 1.33j \times 10^3 & 0 & 0 & 0 \end{pmatrix}$$

Таким образом, первая часть системы (5) в этом случае представляет с собой одно комплексное уравнение относительно напряжения первого узла. Которое запишем в виде

$$U_1 = (V_0)_1 + G_{11} \cdot (\bar{S}_0)_1 / \bar{U}_1 \quad (6)$$

Если записать уравнение (6) покомпонентно, приняв следующие обозначения

$$U_1 = x + j \cdot y, \quad (V_0)_1 = x_0 + j \cdot y_0 = 1232 - 64.767j, \\ G_{11} \cdot (\bar{S}_0)_1 = \alpha + j \cdot \beta = (0.4754 + 5.676j) \times 10^5$$

то получим систему уравнений

$$x = x_0 + \frac{\alpha x - \beta y}{x^2 + y^2}; \quad y = y_0 + \frac{\alpha y + \beta x}{x^2 + y^2}$$

которая сводится к квадратному уравнению относительно, например, переменной  $x$ , решая которую находим два решения

$$x_1 = 175.637, \quad x_2 = 1.105 \times 10^3.$$

соответствующие мнимая части равны

$$y_1 = 451.481, \quad y_2 = 402.641$$

Второе решение отсеивается из соображения ограничений  $|U_1| \leq 500$ . Соответствующие фаза и модуль будут равны:  $\arg(U_1) = 68.743$ ,  $|U_1| = 484.441$ . Общие результаты итераций по всем узлам приведены в таблице 3.

Таблица 3. Результаты итераций по всем узлам

№ уз	Мощность генерации		Напряжение				Относительная разность, %	
			Тестовые результаты		Результаты по упрощенному алгоритму			
	P, МВт	Q, МВар	фаза, град	модуль, кВ	фаза, град	модуль, кВ	фаза, град	модуль, кВ
1	5400.00	0.00	68.734	484.529	68.734	484.441	0,00%	-0,02%
2	0.00	0.00	45.051	482.049	45.734	481.865	1,52%	-0,04%
3	0.00	0.00	21.974	489.348	21.976	489.233	0,01%	-0,02%

4	0.00	0.00	0.118	499.904	0.118	499.904	0,00%	0,00%
5	0.00	457.00	0.000	500.000	0.000	500.000	0	0

Таким образом, в данном случае, нет необходимости применять итерационные методы решения, а решение находится точно.

#### **Заключение**

Указанное преобразование существенно сокращает объем выполняемых вычислений при расчетах установившегося режима.

Изучение вопроса об эффективности преобразования в более общем случае, представляет весьма хороший интерес и является предметом дальнейших исследований.

#### **Список использованной литературы**

1. Мелентьев Л.А. Системные исследования в энергетике. - М.: Наука, 1974. - 416с.
2. Мельников Н.А. Матричный метод анализа электрических цепей. -М.: Энергия. 1972. – 232с.
3. Хачатрян В.С. Метод и алгоритм расчета установившихся режимов больших энергосистем // Известия АН СССР. Энергетика и транспорт. 1973.- №4. -С. 45-57.
4. Краюшкин А.Д. Программная реализация метода Ньютона с переменным шагом на ЕС ЭВМ для расчета нормальных электрических режимов сети большой размерности // Системы автоматизированного управления в энергетике. – Москва, 1982. -56-64с.
5. Гераскин О.Т. Топологическое содержание узлового и контурного определителей электрической сети и расчет их величин при помощи ЦВМ // Известия АН СССР. Энергетика и транспорт. 1966.- №2. - С. 59-70.
6. Percival W.S. Solution of passive electrical networks by means of mathematical trees // Proceedings IEE (London.- 1953, v.100, pt. III, №65. - P. 143-150.
7. Percival W.S. Implored matrix and determinant methods of solving networks // Proceedings IEE (London).- 1954.- v.101, pt.IV, №7. - P. 258-265.
8. Ионкин П.А., Соколов А.А. Основы построения и преобразования графов для расчета электрических цепей // Электричество. 1964.- №5. - С. 67-73.
9. Ионкин П.А. Общие уравнения для расчета электрических цепей с помощью графов // Электричество. 1964.- №8. - С. 27-31.
10. Гераскин О.Т. Графы электрической сети и топологические формулы // Известия АН СССР. Энергетика и транспорт. 1971.- №4. - С. 66-75.
11. Гераскин О.Т. Топологический анализ коэффициентов токораспределения в электрических сетях // Известия АН СССР. Энергетика и транспорт. 1967.- №5. - С. 20-28.
12. Ахметбаев Д.С. Топологический метод анализа стационарных режимов сложных электрических сетей // Электрооборудование: эксплуатация и ремонт. Москва. 2010.- С.150.
13. Ахметбаев Д.С. Метод расчета установившихся режимов электрических сетей на основе коэффициентов токораспределения // Электричество. Москва. 2010., №11, С.23-27
14. Нахождение всех остовных графов заданного графа. Программа для ЭВМ. Свидетельство о внесении сведений в государственный реестр прав на объекты, охраняемые авторским правом. №1551 от 31 января 2019г.
15. Ахметбаев Д. С., Джандигулов А.Р., Ахметбаев А.Д., Ахметова С.О. Оптимизация алгоритма нахождения остовных деревьев графа. Методические вопросы исследования надежности больших систем: Вып. 70 Методические и практические проблемы надежности систем энергетики. В 2хКнигах. / Книга2 / Отв. ред. Н.И. Воропай – ИСЭМ СО РАН, 2019 г.- С.361-366.
16. Ахметбаев Д.С., Ахметбаев А.Д., Бердыгожин А.С. Топологический метод формирования узловых уравнений в обращенной форме для электроэнергетических систем// Электричество. Москва. 2018, №5, С.18-27.

17. Akhmetbaev D.S., Dzhandigulov A.R. Development of algorithms for a new topological method for calculating current distribution coefficients in complex electrical networks//Eurasian journal of mathematical and computer applications. ISSN 2306–6172,Volume 7, Issue 3 (2019 ) 4 – 12

## 3DS MAX БАҒДАРЛАМАСЫНЫҢ КӨМЕГІМЕН ҮШ ӨЛШЕМДІ САХНАЛАРДЫ МОДЕЛЬДЕУ

*Өмірзақова Фариза Нүрділдақызы*  
*Қорқыт Ата атындағы университет, Қызылорда қаласы*  
*Fari\_30\_10\_93@mail.ru*

**Аннотация.** Қазіргі цифрландырылған кезеңде үшөлшемді компьютерлік графиканың алатын орны ерекше. Бұл оның әсерлі жарнамалық роликтерді, арнайы әсерлерді өндіру үшін, сондай - ақ бүгінгі күні пәндік саладағы дизайнерлер арасында интерьерлер мен архитектуралық жобаларды визуализациялау үшін жаппай қолдануымен байланысты. 3D-графика пакеттері нарығында танылған көшбасшылардың бірі - 3DS Max бағдарламасы.

**Кіріспе.** Соңғы он жылдағы технологияның қарқынды дамуы компьютерлік технологияның және оны бағдарламалық қамтамасыз ету саласының тез дамуына ықпал етті. Қазіргі өлшемдер бойынша маңызды емес болып көрінген арнайы эффектілер көмегімен құрылған фильмдер эпизоды, кезінде таңданыс толқыны мен пікірталас тудырғаны белгілі. Қазір кино мен теледидардағы арнайы эффектілермен ешкімді таңқалдыра алмайсың. Олар компьютерлік графиканы және соның ішінде үш өлшемді модельдеуді құру бағдарламасының жалпылай тарауының арқасында күнделікті құбылыс бола бастады. Үш өлшемді графиканың бағдарламалары өз мүмкіндіктері бойынша ең қызықтылары және меңгерілуі бойынша ең күрделілері. Осындай бағдарламалардың ішінде 3 ds Max бағдарламасы көшбасшы орындардың бірін алады. Өзінің тамаша мүмкіндіктерінің және меңгеруге қолжетімділігінің арқасында бұл бағдарлама қызығушылардың арасында да, кәсіпқойлардың арасында да ең танымалы. Шынында да, 3ds Max пайдаланылмайтын үш өлшемді графикамен байланысты адам қызметінің саласы өте аз қалды. Оны ойындармен фильмдерді құру үшін, сәулет және құрылыс, медицина және физика, сондай - ақ басқа да көптеген салаларда белсенді түрде қолданады.

Сондай ақ, жалғыз қолданушылар, бағдарламаны және студияны толығымен жақсы меңгергендер, компьютерлік графиканың аумағында кәсіпқой жұмыс істеушілер, ешқандай қиындықсыз, айрықша еңбексіз, өздерінің жеке сценарийлерін арнайы осы бағдарламаға ендірілген MAXScript макрокоманд тілімен құруына болады.

**Талқылау.** Бұл бағдарлама аты әлемге әйгілі Autodesk фирмасының өнімі. Тез дамуының нәтижесінде 3 ds max салалық стандартқа айналды, оның қолдану аясы үлкен және көп қырлы. Негізінде, бұл үш өлшемді модельдеу және анимация бағдарламасы дүние жүзіндегі үйдегі қолданушылар мен кәсіпқой киноиндустрия мамандарының арасында үлкен сұранысқа ие болды. 3 ds max авторларының идеялары, практикада керемет іске асырылады, қазіргі кезде бұл ең күшті ғана емес, сонымен қатар үш өлшемді әлемдегі ең сатылымды пакет. Сәулетшілік ішкі көріністерді және фасадтарды модельдеу, кейіпкерді анимациялау, Internet-ке арналған шындыққа негізделген фотолық 3D сахналар, физикалық процесстерді визуализациялау - бұлар бағдарлама мүмкіндіктерінің бірі ғана. Бүкіл интернет-сервердің коммерциялық орындауларын немесе ірі компанияның жарнамалық видеобейнесін жасауға болады.

Бағдарламаның жұмыс істеуі нәтижесінде статикалық сахналар құрылады, олар үшөлшемді геометриялық объектілердің жиынтығынан құралған (жалпақ және көлемді), яғни үш координаталармен сипатталады. Қарапайым бұл координаталарды ұзындығы, ені және биіктігі деп атауға болады. Төртінші өлшеуіш - уақыт, анимацияны (жан бітіруді) қолданатын сахналарда немесе динамикалық сахналарда қатысады. Статикалық сахнаны сипаттаудың негізгі үлгісі - объект сәулетінің үшөлшемді моделі, автомобиль двигателінің динамикалық жұмысын демонстрациялау. Кез-келген сахна стандартты алгоритмнен қалыптасады және ол келесі бейнеде сипатталуы мүмкін:

- Геометрия құру.

- Жарық көздерін, түсіру камераларын және материалдарды жөндеу.
- Анимацияны баптау.
- Визуализация.

Статикалық үш өлшемді сахнаның жұмысын аяқтайтын ақырғы нәтиже «сурет»-бейнелеу графикалық файлы болып табылады. Сахна объектілерінің әрбір кадрының өзгерістерін көрсететін динамикалық сахна «суреттердің» жиынтығын немесе анимациялық жүйелілікті береді. Визуализация нәтижелері қағазға, пленкаға, матаға аударылуы немесе бейнелентаға CD-дискке және т.б. жазылуы мүмкін.

**Әдістер.** 3ds max ортасының құралдары мен негізгі командаларын білу және маңызды дағдыларды білумен сипатталатын талаптар - бұлар жұмыстың негізгі сатысы болып табылады. Және де олардың физикалық қасиеттеріне немесе әрекеттестіктеріне емес, тек денелердің геометриясына негізделеді. Сахнаны модельдеу жұмыстарын меңгере келе алғашқы сабақтарды кез-келген үйренуші тез қабылдайды және ақырғы нәтижеге оңай қол жеткізуге болады.

#### *Анимация*

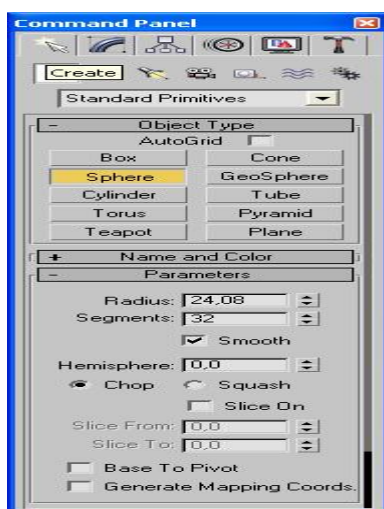
Жеке объекттерді, бүтін топтарды, динамикалық сахнаны модельдеу кезінде қозғалысты басқаратын күшті механизмдер нақты нәтижеге жетуге мүмкіндік береді. Бейнекамерамен түсірілген бейнелерге өте ұқсас етіп нақты кадрлар жасауға болады. Баяулату мен тездету, циклдер мен қайталаулар, уақыт аралықтарын масштабтау және басқа да параметрлер анимацияны басқарады және қолданушыға қолайлы жағдай береді.

#### *Визуализация*

Соңғы саты, параметрлерді баптау, алынған «суреттің» сапасын дұрыстау, генерацияланатын кадрлардың форматы мен типі, арнайы эффекттерді үстемелеу (нұр шашу, бейненің түсуі мен камера линзасындағы дақтар, ашықтықты өзгерту, тез болатын қимылдардың жақсы шығуы, тұман және т.с.с.). Әрбір кадрдың тез оқылуы сахнаның күрделілігіне, қолданылған материалдарға және жұмыс істеп отырған компьютерге тікелей байланысты.

*3ds Max бағдарламасының көмегімен стандартты примитивтерді құру*

*Sphere (Сфера)* объектісін құру үшін *Create (Құру)* панелінен *Geometry (Геометрия)* батырмасын басып, негізгі объект түрлері тізімінен *Standard Primitives (Стандартты примитивтер)* тобын таңдаймыз.



Сурет 1- Құрылатын сфераның параметрлері

*Object Type (Объект түрі)* тізімінен *Sphere (Сфера)* батырмасын таңдаймыз және параметрлер орамынан келесі мәндерді береміз:

- *Creation Method (Құру тәсілі)* орамында: *Center (Центрден)* ауыстырып қосқышы.

- *Parameters (Параметрлер)* орамында: *Segments (Сегменттер)* санағышы-32, *Smooth (Тегістеу)* жалаушасы-қосылған.

2. *Top (Жоғарыдан көрініс)* терезесін белсенді етеміз, көрсеткішті терезенің центріндегі тор координаттардың қиылысқан нүктесіне апарып тышқан батырмасын басып ұстап тұрып жылжытамыз. Көзбен өлшеп отырып сфераның *Radius (Радиус)* санағышының мәнін 65 бірлікке жеткіземіз.

3. *Name and Color (Атау және түс)* орамында атау тұсына «*Тегіс сфера*» деп енгіземіз.

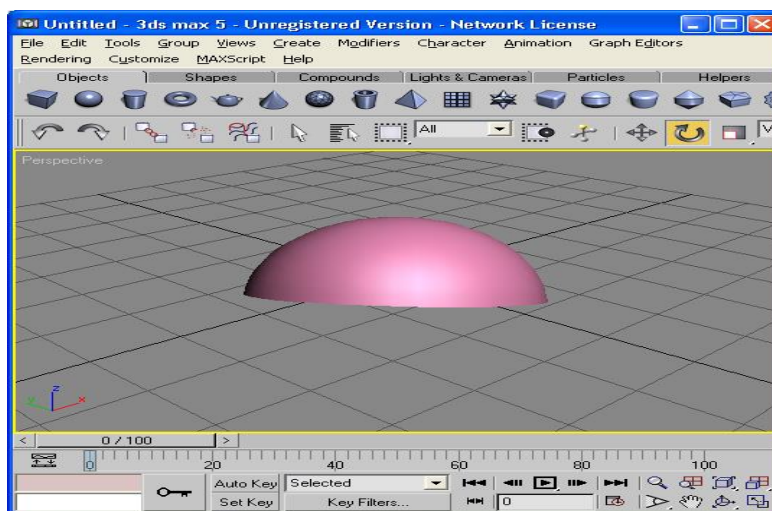
4. Көріністі терезелерді басқару батырмалар терімінің ішінен *Pan (Панорама (Көрініс))* батырмасын таңдап алып, сфераны оңға қарай жылжытамыз. *Sphere (Сфера)* батырмасын таңдап, жаңа объектіге басқа параметрлер енгіземіз.

- *Creation Method (Құру тәсілі)* орамынан: *Edge (Шетінен бастау)* ауыстырып қосқышын қосамыз.

- *Parameters (Параметрлер)* тобынан : *Segments (Сегменттер)* санағышы - 16, *Smooth (Тегістеу)* жалаушасы - алынған, және жоғарыда сипатталған әрекеттерді қайталап, сфера атауын «*Қырланған сфера*» деп енгіземіз.

Нәтижесінде бір-бірімен радиустары бірдей, бірақ сегменттер саны, қабырғаларының тегістігі және тірек нүктесінің орналасуы әртүрлі салынып болған соң 16-сегменттік сфера белгіленіп тұрады.

5. *Modify (Редакторлау)* панеліне өтіп, *Hemisphere (Жартысфера)* санағышының мәнін 0.5 етіп өгертеміз, сонда жарты сфералық сегмент аламыз (2-сурет).



Сурет 2- Squash (Қысу) тәсілімен алынған жарты сфералық сегмент

**Қорытынды.** Үшөлшемді сахнаның шындыққа жақындауы тағы бір маңызды фактор объектерге меншіктеген материалдарға байланысты. Жақсы нәтижеге жету үшін бапталатын қасиеттері мен параметрлерін өзгерту арқылы сфераны жарты сфераға айналдыра алдық.

#### Пайдаланылған әдебиеттер тізімі

1. Абдыкеримова Э.А. 3D модельдеу және анимация [Мәтін] : оқу құралы / Э.А. Абдыкеримова. - Алматы : "Отан" ЖК, 2021. - 178 б

2. Сулу Нагашыбаевна. Компьютерлік дизайн курсына арналған практикум [Мәтін] : оқу құралы / С.Н. Исабаева, Н.Б. Дүйсенова. - Алматы : "Отан" ЖК, 2020. - 96 б.

3. Түнғатаров Нұрмат Нұрғазыұлы. 3ds max-та компьютерлік модельдеу және анимация негіздері [Мәтін] : оқу құралы / Н.Н. Түнғатаров. - Алматы : Қазақ университеті, 2020. - 241 б.



4. Миловская О.С. 3ds Max-та сәулет және интерьерлер дизайны [Мәтін] = Дизайн архитектуры и интерьеров 3 ds Max : Оқулық / Қазақ тіліне ауд. А.О.Жакупова. - Алматы : Print-S, 2012. - 326 б
5. Аббасов, И.Б. Двухмерное и трехмерное моделирование в 3ds MAX / И.Б. Аббасов. - М.: ДМК, 2012. –
6. Ганеев, Р.М. 3D-моделирование персонажей в Maya: Учебное пособие для вузов / Р.М. Ганеев. - М.: ГЛТ, 2012.
7. Зеньковский, В. 3D-моделирование на базе Vue xStream: Учебное пособие / В. Зеньковский. - М.: Форум, 2011.
8. Петелин, А.Ю. 3D-моделирование в Google Sketch Up - от простого к сложному. Самоучитель / А.Ю. Петелин. - М.: ДМК Пресс, 2012. – 344
9. Сазонов, А.А. 3D-моделирование в AutoCAD: Самоучитель / А.А. Сазонов. - М.: ДМК, 2012. - 376 с.
10. Миловская О.С. 3ds max 2014. Питер.400 стр.

## ПРИМЕНЕНИЕ RFM АНАЛИЗА ДАННЫХ ДЛЯ СЕГМЕНТИРОВАНИЯ КЛИЕНТОВ И УДЕРЖАНИЯ ИХ ЛОЯЛЬНОСТИ

*Алимжанова Лаура Муратбековна,  
Кандидат технических наук, профессор  
КазНУ имени Аль-Фараби, Казахстан  
Нуриева Динара Яковлевна,  
Магистрант 2 курса,  
КазНУ имени Аль-Фараби, Казахстан  
[Dinaranuriyeva@gmail.com](mailto:Dinaranuriyeva@gmail.com)*

**Аннотация:** Актуальность выбранной темы обусловлена необходимостью множества компаний увеличить продажи и довести пользователя до совершения покупки. То есть ретаргетинг является стратегически важным компонентом рекламной кампании, направленный на укрепление бренда, повышение лояльности и увеличение продаж. Принцип работы это — «найти и вернуть»: при умелых настройках ретаргетинга можно обратиться к пользователям, уже побывавшим на сайте, но не совершивших на нем целевых действий.

**Ключевые слова:** ретаргетинг, RFM, сегментация, метод, потребительская лояльность

**Введение.** Дабы увеличить продажи, обширно известно, что значима индивидуальная работа с каждым клиентов. Для этого пользователей разделяют на категории и дальше формируют предложения для этих категорий исходя из их особенностей. Существует некоторое количество инструментов для сегментации клиентов. Один из них – RFM-метод. Суть способа – анализ поведенческих факторов групп покупателей. Его применение помогает наладить положительную коммуникацию между бизнесом и заказчиком.

**RFM-анализ как метод сегментации потребителей.** В большинстве случаев рекламные кампании, ориентированные на «широкую аудиторию», непроизводительны: они потребляют огромную часть маркетингового бюджета и не рассматривают надобности заказчиков. Необходимо тщательно анализировать различные маркетинговые стратегии, которые соответствуют личностным качествам и характеристикам заказчиков, для того чтобы обнаружить наилучший способ получить повторные покупки от целевой аудитории [1]. Впрочем, если число заказчиков либо пользователей служб в базе данных организации исчисляется тысячами, обнаружить предложение для всех будет не вполне вероятно. RFM-анализ дает готовую формулу распределения заказчиков на разные поведенческие группы. Финальная цель этого типа обзора — всподвинуть особенно прибыльных заказчиков к повторным заказам и покупкам товаров.

Рекламные стратегии многих компаний основаны на привлечении новых заказчиков. Но для бизнеса не менее значимо сберечь свою аудиторию и поощрять повторные заказы. В качестве основного способа сегментации RFM-анализ подходит для компаний со штатом больше 10 000 работников, т.е. компаний, работающих в секторе B2C [2]. Фирмы B2B с меньшим числом заказчиков также могут применять RFM-сегментацию, но это будет менее эффективно и результативно. Итоги обзора применяются для подготовки электронных писем заказчику в виде СМС либо уведомлений. Вполне возможно предлагать специальные предложения непрерывным заказчикам, напоминать людям, которые уже большое количество времени не совершали покупки, предлагать скидки и акции. Примеры предприятий, где RFM-анализ очень эффективен:

- a. гостиничный и ресторанный бизнес;
- b. салон красоты услуги бьюти-процедур;
- c. интернет-магазины, онлайн площадки;
- d. социальные сети [3].

**Методы и постановка задачи.** Как и все методы сегментации, RFM-анализ имеет свои плюсы и минусы. Среди его превосходств следует выделить: многообразие способов и широкая аналитика по множествам параметров, это в свою очередь дает высокие результаты в большинстве сфер бизнеса. Следующий плюс — это простота и точность. Благодаря автоматизации процесса исключается воздействие человеческого фактора и возможность ошибки. Аналитика разрешает определить маловероятных заказчиков и заблокировать их от отправки почты. Возможно непосредственно увеличение конверсию. Сегментация разрешает, верно, нацеливать рекламу, что приводит к увеличению продаж. Недочеты метода: Связанность итогов от объема базы данных [4]. Если у организации 200–300 заказчиков, данный метод, скорее всего, не сработает. Его ограниченное применение не подходит для компаний, работающих в области разовых продаж. Что следственно делает невозможным совершить объективную оценку параметров пользователей. От того что клиентская база непрерывно меняется, необходимо периодически проводить RFM-анализ.

**Применение MRFM-анализа на практике.** В качестве примера. Какие данные нужны для RFM-анализа? Для реализации RFM – анализа вам понадобятся следующие данные: (см. Таблица №1).

Таблица 1.

Вид данных	Тип
ФИО	Номер, e-mail, ФИО (текст)
Стоимость покупок	Сумма
Дата покупки	Дата
Число покупок	Кол-во

Таблица 2.

/ ФИО) клиента	дата последней покупки	количество покупок	общая стоимость покупки
1	01.07.2018	5	3 245,00
2	17.03.2021	2	721,00
3	21.02.2021	1	2578,00
4	05.05.2020	3	5476,00
5	10.01.2021	2	785, 00
6	31.07.2022	3	235,00

Эти данные можно получить из CRM-системы либо всякого аналитического программного обеспечения, используемого для ведения базы данных. К примеру, сквозная аналитику Calltouch для мониторинга данных [5]. Данный сервис анализирует и подробно определяет не только бюджет, но и производительность рекламных кампаний. Для сегментации заказчиков применяются три критерия: возраст покупки, регулярность и средний чек. Они устанавливаются в расчете на одного клиента. В большинстве случаев - от 1 до 3.

$f_x$  =ЕСЛИ (количество пройденных дней <=0,2\*МАКС (диапазон столбца количества пройденных дней);5;ЕСЛИ (количество пройденных дней <=0,4\*МАКС диапазон столбца количества пройденных дней);4;ЕСЛИ (количество пройденных дней <=0,6\*МАКС диапазон столбца количества пройденных дней);3;ЕСЛИ (количество пройденных дней <=0,8\*МАКС диапазон столбца количества пройденных дней);2;1)))

Figure 1 - Формула подсчета R, F, M

Таблица 3.

количество пройденных дней	R	F	M
363	2	4	3
467	1	3	1
133	4	3	3
59	5	5	5
174	4	2	1
334	2	2	1

Ранг последней покупки (Recency). В 1-ю очередь необходимо определить, какой период из последнего заказа считается отличным, средним либо плохим. Если речь идет о доставке еды, то заказы, сделанные месячной давности, могут быть оценены с наименьшим баллом.

Частота покупок (Frequency) Магазин, торгующий одеждой и обувью, видит в людях, которые приходят любой сезон, четыре раза в год, отменных заказчиков. Заказчику нехорошо делать заказ раз в год.

Сумма чека (Monetary) Данный критерий также зависит от направления организации. Для некоторых видов бизнеса денежных эквивалентов может не быть [6].

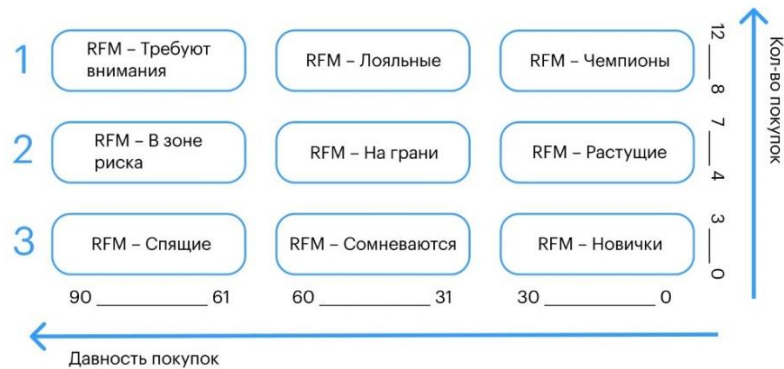


Figure 2 – Классификация клиентов по методу RFM

Для проведения анализа пользуют 5 числовых значению 1 – 5

1 – это наихудший показатель

5 – это наилучший показатель

В итоге исходя и расчетов выше, самым ценным покупателем является может клиент №4 (у которого показателем RFM 555: то есть покупает часто, тратит много), менее ценными покупателями являются – клиенты №2 (131) и №6 (221), так как делают незначительные и не частые заказы.

К примеру, новостные сайты могут измерять время ожидания либо клики по внутренним ссылкам. После того, как заказчики будут оценены по всем параметрам, дайте всякому из них 111, 112, 113, 121 и т. д. Дается до трехсот тридцати трех кодов. В конце концов получается двадцать семь секций [7]. Количество заказчиков в них может сильно различаться. В некоторых случаях до восьмидесяти процентов заказчиков принадлежат к одной группе. После этого нужно добавить добавочный четвертый параметр, но в данном случае количество секций увеличится до шестидесяти четырех.

Как используется RFM-анализ на практике?

1. Аудит клиентов. Можно понять какие клиенты более лояльны, а какие на грани потери.
2. Сегментированные рассылки. Вероятно, увеличить эффективности реклам, методом направления сообщения на конкретный сегмент рассылки.

Инновации — один из важнейших критериев RFM-обзора [8]. Если заказчик долгое время не делает целевого функционирования (не приобретает), можно говорить о высоком риске перехода к сопернику. Предлагая заказчику превосходство, вполне вероятно увеличить частоту его покупок и среднее число чеков, но вернуть утраченных заказчиков — непростая задача. Чем больше напоминаний о вас отправит магазин, тем больше вероятность, что они вновь подадут жалобу [9]. Опытные предприниматели рекомендуют делать RFM аналитику некоторое число раз в квартал. Это разрешает увеличить число повторных запросов и контролировать поток заказчиков. Анализ полученных данных разрешить главную задачу. Таковую как - контроль клиента. Вполне вероятно, увидеть число реальных людей, которые приносят выручка организации. Ретаргетинг — показ рекламы в Интернет, дабы напомнить заказчикам, которые уже посещали ваш веб-сайт, о совершении покупки.

Программа определит особенно результативные ключевые слова, рассчитает цену кампании и предложит варианты рекламы. RFM — не исключительный инструмент обзора аудитории, который можно применять в маркетинговых мероприятиях [10]. Но он считается самым дешевым, экономит время экспертов и дает точные итоги, когда количество заказчиков в секторе B2C превышает некоторое количество тысяч.

**Выводы.** В данной статья была сделана сегментация клиентов в основе RFM анализа. Полученные данные были отражены в матрице лояльности по данному метод. На основе этих классификаций были вывалены наиболее ценные пользователи на которых следует организациям делать упор при повышении продаж, а так же выявлять те группы пользователей кто находится на грани потери. Что в свою очередь поможет компаниям предпринимать меры по повышению лояльности клиентов. Данная методика является универсальной и может быть использована в дальнейшем компаниями разных секторов.

#### Список литературы

1. RFM-анализ: как сегментировать аудиторию по лояльности [Электронный ресурс]. URL: [https://zg-brand.ru/statiy/marketingovye-issledovaniya/rfm\\_analiz\\_kak\\_segmentirovat\\_auditoriyu\\_po\\_loyal\\_nosti/](https://zg-brand.ru/statiy/marketingovye-issledovaniya/rfm_analiz_kak_segmentirovat_auditoriyu_po_loyal_nosti/) (Дата обращения: 20.09.2022).
2. RFM-анализ: что это такое и как его использовать для повторных продаж [Электронный ресурс]. URL: <https://www.calltouch.ru/blog/rfm-analiz-chto-eto-takoe-i-kak-ego-ispolzovat-dlya-povtornyh-prodazh/>
3. Чем полезен RFM-анализ и как использовать данные анализа [Электронный ресурс]. URL: <https://www.retailcrm.ru/blog/chem-polezen-rfm-analiz-i-kak-ispolzovat-dannye-analiza>
4. Дугар-Жабон Т.З., Симакина М. А. Таргетинг и ретаргетинг как инструменты маркетинга /
5. РЕМАРКЕТИНГ И РЕТАРГЕТИНГ: ПОЛНАЯ ИНСТРУКЦИЯ [Электронный ресурс]. URL: <https://www.promowebcom.by/analytics/articles/context-advert/remarketing-i-retargeting-polnaja-instrukcija/>
6. Сегментация клиентов при помощи RFM-анализа [Электронный ресурс] // Crosssellguide.com. — URL: <https://crosssellguide.com/rfm-segmentatsiya-poshagovaya-instruktsiya-how-to/>
7. Цой М.Е., Щеколдин В.Ю. RFM-analysis as a tool for high-tech products' consumers' segmentation // Актуальные проблемы электронного приборостроения

8. Герия И.А. Программы лояльности и оценка их эффективности // Управление и экономика в XXI веке. – 2015. – № 1
9. Голубков Е.П. Маркетинговые исследования: теория, методология и практика. - М.: Финпресс, 2005. – 464 с
10. Сысоева С., Нейман А. Насильно мил не будешь, или что такое лояльность покупателей // Маркетолог. – 2004. – № 2. – с. 32-35.

## ЖОО СТУДЕНТТЕРІНЕ PYTHON ТІЛІНДЕ БАҒДАРЛАМАЛАУДЫ ОҚИТУ

Қапбар Г.С.<sup>1</sup>, Карымсақова А.Е.<sup>2</sup>, Ғапбарова П.С.<sup>3</sup>

Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Астана қаласы, [gkapbar@mail.ru](mailto:gkapbar@mail.ru)<sup>1</sup>,

Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Астана қаласы, [hatae@mail.ru](mailto:hatae@mail.ru)<sup>2</sup>,

Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Астана қаласы, [pgapbarova@mail.ru](mailto:pgapbarova@mail.ru)<sup>3</sup>

**Андатпа.** Мақалада ЖОО – да "Алгоритмдеу және бағдарламалау негіздері" пәнінде бағдарламалау тілі ретінде Python тілін оқытуды ұсынады. Бағдарламалауды жаңадан үйренуші студенттерге Python тілінде оқытудың тиімді тұстарын әлемдік сарапшылардың тұжырымдамасы арқылы және де жоғарғы оқу орнындағы студенттерге жүргізілген іс – тәжірибе арқылы айқындап көрсетілді.

**Abstract.** The article suggests teaching Python as a programming language at a university in the discipline "Fundamentals of algorithmization and programming". The advantages of learning Python for students studying Programming were determined by the concept of world experts, as well as the practice conducted for students of higher educational institutions.

**Түйін сөздер.** Python, алгоритмдеу, ЖОО, бағдарламалау, C, C++, Java.

**Кіріспе.** Еліміздегі жоғарғы оқу орындарында "Алгоритмдеу және бағдарламалау негіздері" пәнін оқу процесінде ерекше назар аударуды қажет ететін ең қызықты мәселелердің бірі - студенттердің бағдарламалауды оқыту деңгейін арттыру. Қазіргі уақытта студенттерге жоғарғы оқу орындарында бағдарламалау негіздерін үйрету көбінесе C++, C#, C, Pascal тілдерінде жүреді. Сонымен қатар, соңғы жылдары Python бағдарламалау тілі кеңінен танымал бола бастады, ол бүкіл әлемде бағдарламалау негіздерін оқыту үшін де, мәтіндік диалогтардан бастап маңызды веб-қосымшалар мен 3D ойындарына дейін әртүрлі бағдарламалар жасау үшін кеңінен қолданылады.[1]

"Алгоритмдеу және бағдарламалау негіздері" пәнінің мақсаты - студенттерді келешекте алгоритмдеу әдістері мен программалауды қолдану және меңгеру үшін білім беру, қажетті іскерлік дағдыны игерту, бағдарламалау саласында маман ретінде пайдалануға оқыту. Негізгі міндеттері – қолданбалы есептерді шешу кезінде студенттерге алгоритм әдістері және берілгендер құрылымы туралы жүйелі түсінік беру, математикалық ақпараттар арқылы практикалық іс – әрекеттерді нақты әлемнің заңдылықтарын пайдалануда кездесетін әр – түрлі есептердің шешімін ЭЕМ – нің көмегімен есептелетін алгоритмдерді қолдануға және құруға студенттерді даярлау.[2-4]

Осылайша, мақаланың тақырыбы өзекті болып көрінеді, өйткені ол жоғарғы оқу орны студенттерін "Алгоритмдеу және бағдарламалау негіздері" пәнінде Python тілінде бағдарламалауды оқытуды ұсынады.

**Зерттеу мақсаты.** ЖОО – да өтетін "Алгоритмдеу және бағдарламалау негіздері" пәнінде Python тілінде бағдарламалауды үйретудің оңтайлы тұстарын ашып көрсету.

**Зерттеу міндеттері.** 1) Практикаға бағытталған тәсілді пайдалана отырып, Python-да бағдарламалауды оқыту оңтайлығының жай-күйіне талдау жүргізу;

2) ЖОО – ы студенттерінің Python-да бағдарламалауды оқытудың тиімділігіне практикаға бағытталған тәсілді қолданып, орындылығын негіздеу;

**Зерттеу материалдары мен әдістері.** Жоғарғы оқу орындарында әр түрлі авторлардың "Алгоритмдеу және бағдарламалау негіздері" пәні бойынша қолданыстағы силлабустарын талдау олардың C, C++, Pascal бағдарламалау тілдерін таңдайтындығын көрсетті. Сондай-ақ, оқыту жүйесін әзірлеудің қолданыстағы құралдарына талдау жүргізілді, олардың артықшылықтары мен кемшіліктері қарастырылды. Python тілінің бағдарламалауды бастаушы студенттер үшін қолайлы екендігін атап өтуге болады:

Python - бұл нәтижелерді шығару үшін интерпретаторды қолданатын жоғары деңгейлі ("адам оқи алатын") бағдарламалау тілі. Python-да тексерілген код

модульдерінің кең стандартты кітапханасы бар, оларды өз бағдарламаларыңызға оңай қосуға болады.

Гвидо ван Россум сексенінші жылдардың аяғы мен тоқсаныншы жылдардың басында Нидерландыдағы математика және информатика ұлттық ғылыми-зерттеу институтында жасаған. Python тілі көптеген басқа тілдердің, соның ішінде C, C ++ тілдерінің туындысы болып табылады.

Python тілінің ұтымды тұсы - бағдарламалауды жаңадан бастаушыларға қолайлы ететін кодтың оқылуы, Python философиясының принциптерінің бірі болып табылады.

Python коды оқуға бағытталғандықтан, ол ағылшын тіліндегі кілт сөздерді жиі қолданады, ал басқа бағдарламалау тілдері сол сөздер орнына әдетте тыныс белгілерін қолданады. Оның ерекше айырмашылығы - код блогындағы нұсқауларды топтастыру үшін Python кілт сөздерді немесе тыныс белгілерін емес, бос орындарды қолданады. Паскаль тілінде, мысалы, блоктардың басталуы begin кілт сөзімен белгіленеді және end кілт сөзімен аяқталады, ал C бағдарламашылары код блоктарын көрсету үшін жақшаларды қолданады. Көбінесе блоктарды бос орындармен топтастырудың бұл тәсілін басқа тілдермен таныс бағдарламашылар сынға алады, бірақ Python-да бос орындарды қолдану бағдарламалардың жинақы көрінуіне мүмкіндік береді.

Бағдарламалауды жаңадан үйренушілер үшін, Python тілінің маңызды ерекшеліктерін тізімдейміз.

- **Python тегін** - бұл еркін және ашық бастапқы кодты бағдарламалық жасақтама.

- **Python - ды үйрену оңай** - оның қарапайым синтаксисі бар

- **Python оқуға оңай кодты жасауға мүмкіндік береді** - ол тыныс белгілерімен артық жүктелмейді.

- **Python - ға қызмет көрсету оңай** – оның модульдік құрылымы бар

- **Python бай "арсеналға" ие** - ол сіздің бағдарламаларыңызға оңай енетін үлкен стандартты кітапхананы ұсынады.

- **Python портативті** - оны көптеген платформаларда іске қосуға болады және барлық жерде бірдей интерфейс болады.

- **Python түсіндірілген (“интерпретируемый”)** - компиляцияны қажет етпейді.

- **Python-бұл жоғары деңгейлі тіл** - оның статикалық жад үлестірімі бар.

- **Python кеңейтіледі** - төмен деңгейлі модульдерді қосуға мүмкіндік береді.

- **Python әмбебап** - процедуралық және объектіге бағытталған бағдарламалау әдістерін қолдайды.

- **Python қолдануға икемді** - оның көмегімен консоль бағдарламаларын, GUI қосымшасын, сонымен қатар сыртқы бағдарламалардың веб - серверлермен өзара әрекеттесуіне арналған сценарийлер жасауға болады.[5]

C және C++ тілдері көбінесе жылдамдық маңызды болған кезде қолданылады. Бірақ тілдердің синтаксисін түсіну қиынырақ және де аз мәліметті алу үшін үлкен көлемде код жазуды қажет етеді. C тілінде жазылған шағын бағдарлама:

```
#include <stdio.h>
int main(int argc, char *argv[]) {
int language = 1;
printf("Language %d: I am C! Behold me and tremble!\n", language);
return 0;
}
```

C және C++ тілдері синтаксисі жағынан бір – біріне ұқсас, бірақ ерекшеліктері бар. Жоғарыда C тілінде келтірілген шағын бағдарламаның C++ тілінде жазылуы:



```
#include <iostream>
using namespace std;
int main() {
int language = 2;
cout << "Language " << language << \
": I am C++! Pay no attention to that C behind the curtain!" << \
endl;
return(0);
}
```

Java және C# тілдері C++ тілінің синтаксисіне ұқсас. Бірақ олар артық және шектеулі. Жоғарыда C++ тілінде келтірілген шағын бағдарламаның Java тілінде жазылуы:

```
public class Overlord {
public static void main (String[] args) {
int language = 3;
System.out.format("Language %d: I am Java! Scarier than C!\n", language);
}
}
```

Кейбір тілдерге маңызды синтаксистік багаж жүктелген. Оларды кейде статикалық тілдер деп атайды, өйткені олар сізден компьютерге енгізілген мәліметтер типін көрсетуді талап етеді.

Бағдарламалау тілдерінде бағдарламалау барысында қолданылатын *айнымалы*, яғни мәліметтер есімін білдіретін ұғым бар. Статикалық тілдер әр айнымалының жадта қанша орын алатынын көрсететін мәліметтің типін көрсетуді талап етеді. Компьютер бұл ақпаратты бағдарламаны өте төмен деңгейлі машиналық тілге құрастыру үшін пайдаланады. Мәліметтер типін көрсету бағдарламаның тез жұмыс жасауына көмектеседі, бірақ ол код жазу барысын тежейді. C, C++, Java тілдерінде жазылған кодтың көп бөлігі мәліметтердің типін көрсетуден тұрады. Мысалы, жоғарыда келтірілген мысалдарда language айнымалысы бүтін мәнді қайтару үшін int мәліметтер типін жариялау қажет болды. C, C++, Java тілдерінде айнымалылар типтерін өзгерте алмайтындықтан оларды *статикалық тілдер* деп атайды.

Динамикалық тілдер – статикалық тілдерге қарама – қайшы болып келеді. Бұл тілдер сізге айнымалыларды қолдану үшін оның типін көрсетуді талап етпейді. Мысалы x = 5 болса, динамикалық тіл 5 санының бүтін сан екендігін анықтайды, сол себепті x айнымалысы int типін қабылдайды. Бұл тілдер ауқымды бағдарламаларда аз қатарлы код жазуға мүмкіндік береді. Олар компиляциядан өту орнына, интерпретатор қоладанады. Динамикалық тілдер әдетте статикалық тілдерге қарағанда баяу жүреді, бірақ интерпретаторлар оңтайландырылған сайын олардың жылдамдығы артады.

Көптеген жылдар бойы көп мақсатты динамикалық тіл Perl болды. Perl бағдарламалау тілі өте күшті және көптеген кітапханаларға ие. Алайда, оның синтаксисін түсіну қиын болуы мүмкін, ал Python және Ruby бағдарламалау тілдерінің пайда болуына байланысты Perl бағдарламалау тілі танымалдылығын жоғалтады. Жоғарыда келтірілген мысалды Python тілінде жазар болсақ:

```
language = 7
print("Language %s: I am Python. What's for supper?" % language)[6]
```

Жоғарыда келтірілген мысалдардан бір мағынадағы бағдарламалардың статикалық C, C++, Java тілдеріне қарағанда, Python бағдарламалау тілінде жазылған код

көлемнің аздығы жөнінен де және жеңіл оқылатындығымен ерекшеленетінін байқауға болады.

ЖОО студенттерін Python-да бағдарламалауды оқытудың тиімділігіне практикаға бағытталған тәсілді қолданудың орындылығын негіздеу барысында эксперимент жүргізілді. "Алгоритмдеу және бағдарламалау негіздері" пәнінде 1 – курс студенттеріне мәтіні бір есепті бірнешеде тілде шығарып көру арқылы талдау жасалынды. Есеп мәтіні: Берілген екі санның қосындысын экранға шығару. Мұнда  $a=5, b=8$  болса. Эксперимент нәтижесі 1 – кестеде көрсетілген.

Эксперимент нәтижесі

1 – кесте

Бағдарламалау тілдері	Есепті шығаруға кеткен уақыты	Есептің шығарылу коды	Бағдарламалау тілінің оң/теріс тұстары
C++	8 минут	<pre>#include &lt;iostream&gt; using namespace std; int main() { int a = 5; int b = 8; int sum = 0; sum = a + b; cout&lt;&lt;"sum ="&lt;&lt;sum; return 0; }</pre>	<ul style="list-style-type: none"> <li>• Тыныс белгілер көп</li> <li>• Айнымалының типін көрсету шартты</li> <li>• Бағдарлама денесін блок ішінде жазу қажет</li> <li>• “#include &lt;iostream&gt; using namespace std;”, “return 0” секілді тұрақты сөздерді қажет етеді</li> </ul>
Java	7 минут	<pre>public class Main { public static void main(String[] args) { int a = 5; int b = 8; int sum =0; sum = a + b; System.out.println("sum ="+sum); } }</pre>	<ul style="list-style-type: none"> <li>• Айнымалының типін көрсету шартты</li> <li>• Бағдарлама денесін блок ішінде жазу қажет</li> <li>• Бағдарлама барысында жазылу керек тұрақты сөздер көп</li> </ul>
C	8 минут	<pre>#include &lt;stdio.h&gt; int main() { int a = 5; int b = 8; int sum; sum = a + b; printf("sum =%d", sum); return 0; }</pre>	<ul style="list-style-type: none"> <li>• Тыныс белгілер көп</li> <li>• Айнымалының типін көрсету шартты</li> <li>• Бағдарлама денесін блок ішінде жазу қажет</li> <li>• “#include &lt;iostream&gt;” “return 0” секілді тұрақты сөздерді қажет етеді</li> </ul>
Python	4 минут	<pre>a = 5 b = 8 sum = a + b print ('sum =',sum)</pre>	<ul style="list-style-type: none"> <li>• Бағдарлама мәтіні қысқа</li> <li>• Бағдарлама мәтіні бір қараған адамға</li> </ul>

			түсінікті • Тыныс белгілерді қажет етпейді
--	--	--	--

Жүргізілген эксперимент арқылы C,C++,Java тілдеріне қарағанда мәтіні және де шығарылу алгоритмі бірдей есептің Python бағдарламалу тілінде код көлемінің аздығымен, есепті шығаруға салыстырмалы түрде аз уақыт жұмсалыуымен, тұрақты сөздер мен блоктары аз болуымен,яғни қарапайым синтаксисімен ерекшеленетінін байқадық.

### Қорытынды

ЖОО студенттері “Алгоритмдеу және бағдарламалау негіздері” пәнінде бағдарламалауды жеңіл үйрену үшін динамикалық Python тілін қолдану арқылы өту оңтайлы шешім болып табылады.Себебі, сарапшылар бағдарламалауды үйренуде ең жақсы тілдердің бірі Python деп айтады. Қарапайым синтаксис бұл тілді қабылдауға ыңғайлы және түсінікті етеді. Сондай – ақ Python білімі бар бағдарламалаушылар қазіргі таңда сұранысқа ие. Олар веб-әзірлеу, machine learning және Data Science салаларында жұмыс істейді.

Алгоритмдеуді және бағдарламалау тілдерін жаңадан үйренуші студенттер үшін Python тілін оқыту тиімді екенін жүргізілген эксперимент барысында байқадық. ЖОО – да “Алгоритмдеу және бағдарламалау негіздері” пәнінде бағдарламалау тілі ретінде Python тілін қарастырсақ, студенттер Python тілінің қарапайым әрі түсінікті синтаксисін тез меңгеріп қана қоймай келешекте нарықта сұранысқа ие маман болып шығары анық.

### Әдебиеттер тізімі

1. Кочеткова О.А., Долгополов И.В. «Программирование на языке Python» курсы бойынша электрондық оқу құралдары // Университеттік білім (МКУО - 2016): XX Халықаралық ғылыми-әдістемелік конференцияның мақалалар жинағы (Пенза, 7-8 сәуір 2016 ж.). Пенза: Пенза мемлекеттік университетінің баспасы, 2016. 104-105 Б.

2.Торайғыров университеті сайты.[Электрондық ресурс]:<https://www.tou.edu.kz/kz/> (Қаралған күні:19.09.22)

3. Халықаралық Ақпараттық Технологиялар университеті сайты.[Электрондық ресурс]: <https://iitu.edu.kz/kk/articles/ob-universitete-kk/o-muit-kk/>

4.Сатпаев университеті сайты.[Электрондық ресурс]: <https://satbayev.university.ru>

5.Mike McGrath Python in easy steps, - 2013. -192 с.

6.Билл Любанович Простой Python. Современный стиль программирования, - 2016. -480 с.

# ДЕТЕКТИРОВАНИЕ АНОМАЛЬНЫХ СИТУАЦИЙ ИНФОРМАЦИОННЫХ СИСТЕМ

*Б.Т. Рзаев<sup>1</sup>, Ж.Т. Бельдеубаева<sup>2</sup>, И.М. Увалиева<sup>3</sup>*

*<sup>1</sup>Казахский агротехнический университет им. С.Сейфуллина, г. Астана,  
pathinchaos@gmail.com*

*<sup>2</sup>Казахский агротехнический университет им. С.Сейфуллина, г. Астана,  
zh.beldeubayeva@mail.ru*

*<sup>3</sup>Восточно-Казахстанский технический университет им. Д. Серикбаева,  
г. Усть-Каменогорск, iuvalieva@mail.ru*

## **Введение**

Обнаружение аномалий относится к проблеме поиска закономерностей в данных, которые не соответствуют ожидаемому поведению. Эти несоответствующие паттерны часто называют аномалиями, выбросами, несогласованными наблюдениями, исключениями, абберациями, неожиданностями в различных областях применения. Из них аномалии и выбросы – это два термина, наиболее часто используемые в контексте обнаружения аномалий, иногда взаимозаменяемые. Обнаружение аномалий находит широкое применение в самых разнообразных приложениях, например, обнаружение мошенничества с кредитными картами, страховкой или медицинским обслуживанием, обнаружение вторжений для обеспечения кибербезопасности, обнаружение неисправностей в критически важных системах безопасности и военное наблюдение за действиями противника.

Важность обнаружения аномалий обусловлена тем фактом, что аномалии в данных приводят к существенной и часто критической информации, пригодной для применения в самых разных областях. Например, аномальное МРТ-изображение в медицине может указывать на наличие злокачественных опухолей, аномалии в данных транзакций по кредитной карте могут указывать на кражу кредитной карты или личных данных, аномальные показания датчика космического корабля могут указывать на неисправность в каком-либо компоненте космического корабля, аномальная структура трафика в телекоммуникационной сети (далее - ТС) может означать, что взломанный компьютер отправляет конфиденциальные данные неавторизованному адресату [1]. Именно последний пример аномалий наиболее близко подходит к теме данного исследования и непосредственно отражает суть данной работы.

## **1. Обзор существующих методов**

Общие свойства сетевого пакетного трафика в информационных системах (далее - ИС) интенсивно изучались в течение многих лет [2-4]. Большинство этих исследований по анализу трафика были сосредоточены на типичном поведении на уровне пакетов. В частности, исследование [4] сосредоточено на выявлении аномалий путем удаления сначала из сигнала его предсказуемой, внешней части и только затем использования статистических методов. Также было проведено множество исследований методов обнаружения сетевых неисправностей, где использовались статистические отклонения от нормального поведения трафика для выявления неисправностей [5]. Эти исследования были сосредоточены на точном выявлении отклонений от нормального поведения. Для обнаружения DDoS атак предлагались новые информационные метрики [6], которые позволяют уменьшить ложные срабатывания и увеличивают чувствительность к обнаружению.

На данном этапе развития ИС, широкое применение получили методы машинного обучения (далее - МО) [7-8] и нейронных сетей (далее - НС) [9].

## **2. Подход к детектированию вредоносного трафика на основе голосования классификаторов МО**

Идентификация вредоносного трафика в сети является довольно непростой задачей, так как применяемые злоумышленниками технологии постоянно совершенствуются - растет качество инструментов и используются нестандартные решения. Но все же человеческий фактор дает о себе знать, и злоумышленники оставляют свои следы пребывания в сети. К тому же все нестандартные решения приводят к новым явлениям в сети - аномалиям, ранее не происходившим, и этим можно успешно воспользоваться специалистам по безопасности, с помощью применения интеллектуальных систем обработки данных.

Можно выделить и обобщить следующие признаки аномальных явлений в сети:

1) Большое количество одинаковых по типу пакетов. Это может сигнализировать о попытках исполнения атак отказа в обслуживании (DoS, DDoS), которое вследствие чрезмерной загрузки каналов передачи данных и процессоров серверов может привести к потере связи с критичными сервисами и выходу из строя сетевого оборудования.

2) Последовательность пакетов и их значения. Данный признак трафика может характеризовать о несанкционированном перемещении в сети. Злоумышленник, получив доступ к незнакомой ему сети, будет искать соседние подходящие для его задач устройства и пробовать к ним подключаться [10]. Если предугадать последовательность его действий, то можно с большой вероятностью обнаружить неправомерные действия в сети.

3) Закономерности появления сетевых пакетов. Данный признак содержит несколько подпризнаков, к примеру может появиться пользовательская активность в нерабочее время, периодичность возникновения одинакового по типу трафика, необъяснимо большая продолжительность трафика. Эти и другие виды закономерностей в сети требуют тщательной обработки и проверки на безопасность.

В рамках данного исследования предлагается подход по детектированию аномальных явлений в сетевом трафике информационных систем, на основе голосования, который производит сопоставление результатов нескольких классификаторов.

Формализованное описание предлагаемого подхода представлено в работах [11-12].

В общей постановке задача выглядит следующим образом. Имеется выборка, в котором получены значения анализируемых параметров по различным состояниям. Это позволяет сделать соответствие набора кортежей параметров из множества  $X$  каждому состоянию.

$X_i = (x_1, \dots, x_n)$  содержит кортеж значений исследуемых параметров длины  $n \geq 2$ .

Множество состояний определяется кортежами  $\{X_1, X_2, \dots, X_m\} \in X$ , где  $m$  - количество записей в выборке, которые отражают поведение процесса в разных состояниях. Каждому состоянию ставится в соответствие метка бинарного множества подмножества опасных  $C_1$  и безопасных  $C_2$  состояний.

В следствии чего получена обучающая выборка размеченная по классам опасности. Необходимо для входного кортежа значений  $X_i$  построить алгоритм классификации  $a = \{b_1, b_2, \dots, b_k\}$ , отображающий  $X \rightarrow C$ , где  $k$  - количество базовых алгоритмов классификации в ансамбле. В ходе функционирования отдельно классифицирующего алгоритма возникает ошибка, которая сглаживается последовательностью  $b_q, q = 1, \dots, k$ , обученных независимо друг от друга. Ответ алгоритма  $b_q(X_i)$ , который получен по итогам работы, определяет класс подмножеств  $C_j, j = 1, 2$ , относящийся к бинарному множеству классов состояний  $C$ .

Результат работы алгоритмов  $b_q, q = 1, \dots, k$  по отдельности - вероятность принадлежности поступившего на вход кортежа  $X_i$  классу состояний:

$$b_q(X_i) = \max_{j=1,2} P_q(C_j|X_i). \quad (1)$$

Результирующий класс, определяемый ансамблем алгоритмов классификации для кортежа  $X_i$ , можно определить с применением значений функций  $F_1$  и  $F_2$  для подмножеств  $C_1$  и  $C_2$ :

$$F_1(b_1(X_i), \dots, b_k(X_i)) = \frac{1}{k} \sum_{q=1}^k P_q(C_1|X_i), F_2(b_1(X_i), \dots, b_k(X_i)) = \frac{1}{k} \sum_{q=1}^k P_q(C_2|X_i). (2)$$

Ансамбль базовых алгоритмов классификации описывается следующим выражением:

$$b(X_i) = f(F_1(b_1(X_i), \dots, b_k(X_i)), F_2(b_1(X_i), \dots, b_k(X_i))), (3)$$

где  $f$  – это решающее правило, которое определяет вероятностную оценку и устанавливает номер класса.

Таким образом, предлагаемое решение применяет бинарную классификацию алгоритмов. Формализация обобщается решающим правилом, которое переводит оценку в номер класса [9].

### 3. Эксперимент

Оценка применяемого подхода произведена на публичном наборе данных NSL-KDD [13-14]. Эксперимент предполагает бинарную классификацию состояний телекоммуникационной системы (идентификация паразитного (аномального) и нормального трафика). На данном этапе нет четко регламентированных правил по формированию состава ансамбля и количества алгоритмов в них, что требует проведения дополнительных экспериментов. Выбор определенной модели осуществляется исходя из задачи. Состав ансамбля необходимо уточнять заранее, так как его изменение в дальнейшем будет затруднено. В эксперименте предложен случай голосования, где использовались сильные и слабые классификаторы. В оценке участвовали следующие алгоритмы классификации: Naïve Bayes (NB), Decision Stump (DS), JRip (JR). Результаты анализа получены с помощью программного обеспечения Weka.

Размеченная выборка была разделена на две части, одна из них – для обучения, другая – для тестирования. Структуры выборки включала более 40 значений различных атрибутов [14]. Полученный из набора данных сетевой трафик рассматривается как последовательность кортежей значений пакетов, около 50 % из которых отражает нормальный (безопасный) трафик и 50 % - аномальный.

Гистограмма полученных результатов представлена на рисунке 1.

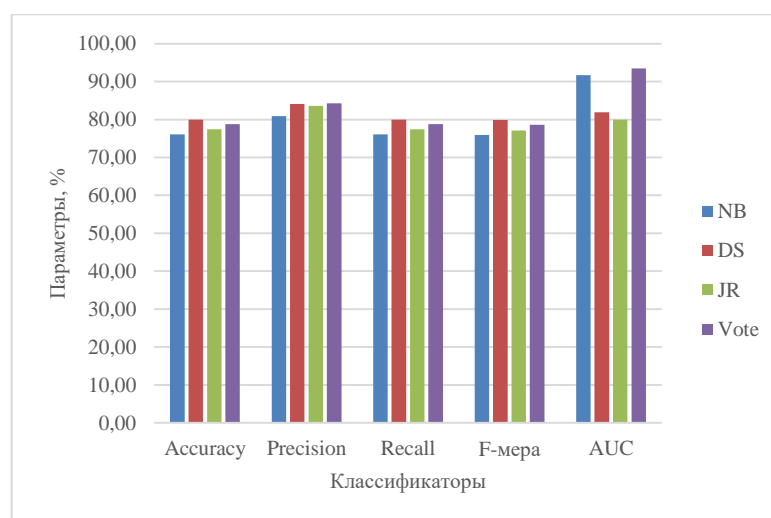


Рисунок 1. Гистограмма оценки классификаторов

Полученные результаты эксперимента на основе классификаторов в отдельности и с применением голосования представлены в таблице 1.

Таблица 1.

Классификаторы	Accuracy	Precision	Recall	F-мера	AUC
NB	76,12	80,90	76,10	75,90	91,71
DS	79,98	84,10	80,00	79,90	81,87
JR	77,42	83,60	77,40	77,10	79,99
Vote	78,76	84,30	78,80	78,60	<b>93,44</b>

Оценка алгоритмов классификации была произведена на основе площади под ROC-кривой (AUC) для тестового множества, рисунок 2.

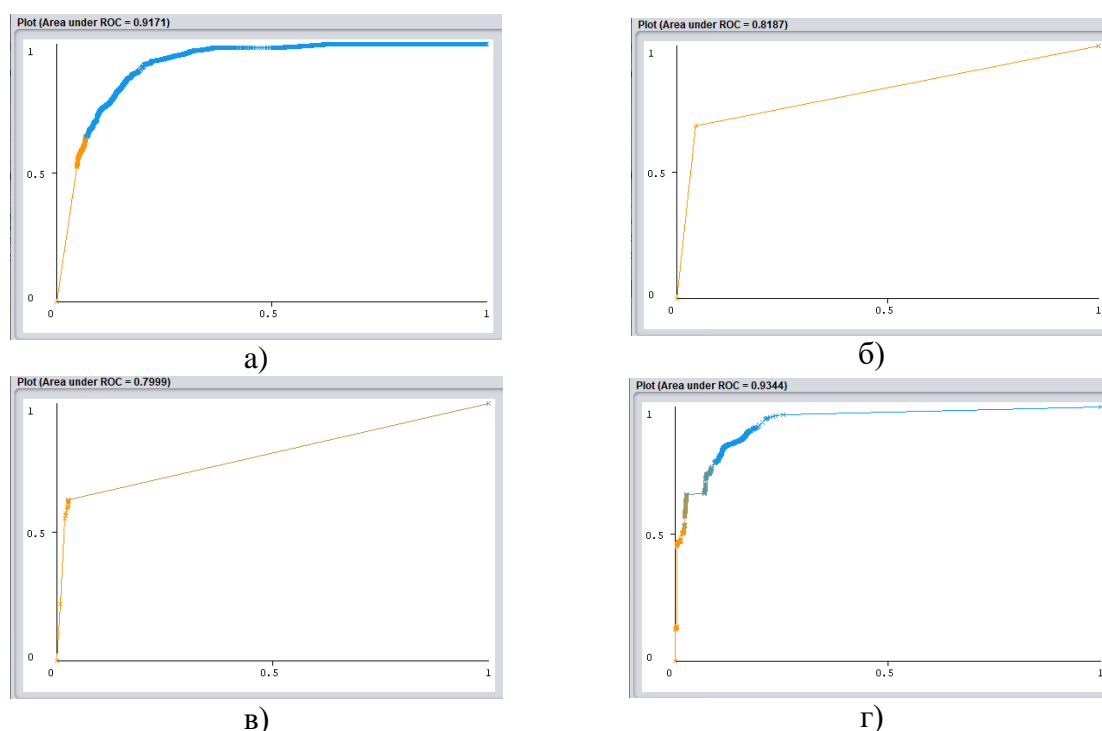


Рисунок 2. Площади под ROC-кривой (AUC) классификаторов NB (а), DS (б), JR (в), и голосования Vote (г)

Результаты тестирования на наборе данных NSL-KDD алгоритмами машинного обучения, с применением голосования, реализованными в Weka, показали точность более 93 %, даже на сравнительно небольшой по объему обучающей выборке.

### Заключение

Поиск аномальных ситуаций информационных систем является довольно сложной и ответственной задачей. От своевременного выявления возможных злонамеренных воздействий на ТС компаний зависит экономическое и репутационное благополучие компаний.

По полученным в пределах данной работы результатам, становится очевидным преимущество применения ансамблей разнородных классификаторов, чем их применение в отдельности, при том, что это требует относительно небольшой по объему выборки. Классифицирующие алгоритмы настраиваются для разных видов событий и аномалий, а полученный эффект разброса ответов сглаживается ансамблем. Каждый классификатор «специализируется» на определенной части событий, что позволяет адаптировать его к различным условиям функционирования сетевых сегментов.

Среди недостатков предложенного подхода необходимо отметить чувствительность классифицирующих алгоритмов к качеству данных в обучающей выборке. В случае его сильного влияния результаты ансамбля могут значительно снижаться.

#### Список использованных источников

- 1 Chandola V., Banerjee A., Kumar V. Anomaly detection. *ACM Computing Surveys*. (2009). 41(3). 1–58. doi:10.1145/1541880.1541882.
- 2 R. Càceres Measurements of wide-area Internet traffic. Tech.Rep. UCB/CSD 89/550. Computer Science Department. University of California. Berkeley, 1989.
- 3 W. Willinger, M. Taqqu, R. Sherman, D. Wilson Self-similarity through high-variability: Statistical analysis of Ethernet LAN traffic at the source level. *IEEE/ACM Transactions on Networking*. Vol. 5. No. 1. pp.71-86. 1997.
- 4 Barford P., Kline J., Plonka D., Ron, A. A signal analysis of network traffic anomalies. *Proceedings of the Second ACM SIGCOMM Workshop on Internet Measurement - IMW '02*. 2002. doi:10.1145/637201.637210.
- 5 F. Feather, D. Siewiorek, R. Maxion Fault detection in an ethernet network using anomaly signature matching. *Proceedings of ACM SIGCOMM '93*. San Francisco CA. 2000.
- 6 Xiang Y., Li K., Zhou W. Low-Rate DDoS Attacks Detection and Traceback by Using New Information Metrics. *IEEE Transactions on Information Forensics and Security*, 6(2), 426–437. 2011. doi:10.1109/tifs.2011.2107320.
- 7 Alrashdi I., Alqazzaz A., Aloufi E., Alharthi R., Zohdy M., Ming, H. AD-IoT: Anomaly Detection of IoT Cyberattacks in Smart City Using Machine Learning. 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC). doi:10.1109/ccwc.2019.8666450.
- 8 Рзаев Б.Т., Лебедев И.С. Применение бэггинга при поиске аномалий сетевого трафика. *Научно-технический вестник информационных технологий, механики и оптики*. 2021. Т. 21, № 2. С. 234–240. doi: 10.17586/2226-1494-2021-21-2-234-240.
- 9 Рзаев Б.Т., Лебедев И.С. Applying of Machine Learning for Analyzing Network Traffic in the Conditions of an Unbalanced Data Sample. [Studies in Computational Intelligence](#) 1026. 2022. pp. 69-78. doi: 10.1007/978-3-030-96627-0\_7.
- 10 Е. Килушева Как выявить атаку злоумышленников в сетевом трафике. <https://www.anti-malware.ru>. Дата обращения: 22.09.22.
- 11 Семенов В.В., Лебедев И.С., Сухопаров М.Е. Идентификация состояния отдельных элементов киберфизических систем на основе внешних поведенческих характеристик. *Прикладная информатика*. 2018. Т. 13. № 5 (77). С. 72–83.
- 12 Сухопаров М.Е., Лебедев И.С. Идентификация состояния информационной безопасности устройств интернета вещей в информационно-телекоммуникационных системах. *Системы управления, связи и безопасности*. 2020. № 3. С. 252–268. doi: 10.24411/2410-9916-2020-10310.
- 13 Ingre B., Yadav A. Performance Analysis of NSL-KDD dataset using ANN. *Proc. 4th International Conference on Signal Processing and Communication Engineering Systems (SPACES)*. 2015. P. 92–96. doi: 10.1109/SPACES.2015.7058223
- 14 Dhanabal L., Shantharajah Dr. S.P. A Study on NSL-KDD dataset for intrusion detection system based on classification algorithms. *International Journal of Advanced Research in Computer and Communication Engineering*. 2015. V. 4. N 6. P. 446–452. doi: 10.17148/IJARCCCE.2015.4696.



## ШЕШІМ ҚАБЫЛДАУДЫҢ КӨП КРИТЕРИАЛДЫ ЕСЕБІНІҢ МАТЕМАТИКАЛЫҚ МОДЕЛІ

*Мауина Г.М., магистр*

*Айтимова У.Ж., ф.-м.ғ.к., қауымд.проф. м.а.*

*С.Сейфуллин атындағы Қазақ агротехникалық университеті, Астана қ.,  
alema85@mail.ru*

### **Аннотация.**

Материалда Солтүстік Қазақстанның ауылшаруашылық кәсіпорындарының өндірістік сценарийлері бойынша басқарушылық шешімдер қабылдау мәселесі қарастырылады. Мәселе өндірістік және нарықтық конъюнктураның критерийлер (параметрлер) жиынтығының болуына байланысты, критерийлердің дұрыстығы ауылшаруашылық объектісінің жұмыс істеуінің қолайлы сценарийі туралы басқарушылық шешім қабылдау сапасын анықтайды. Зерттеудің мақсаты проблеманы құрылымдау және көп өлшемді әсер ету факторлары жағдайында Солтүстік Қазақстанның аграрлық секторының кәсіпорындарында өндірісті ұтымды ұйымдастыру сценарийін таңдау бойынша шешімдер қабылдау міндетін ресімдеу болып табылады. Зерттеу әдістемесі күрделі мәселені талдауға ережелер беретін жабық логикалық құрылым ретінде өзара тәуелді компоненттер жүйесін (ресурстар, өндірістік сценарийлер) қарастыруға мүмкіндік беретін жүйелік тәсілге негізделген.

### **Кіріспе.**

Қазіргі уақытта ауылшаруашылық салаларында басқарушылық шешімдер қабылдауды қолдау әдістерін әзірлеу өте өзекті болып табылады. Бұл ауыл шаруашылығы өндірісінің тиімділігін арттыру, ресурстық міндеттерді шешу және басқа да көптеген салаларға әсер етеді. Қазіргі заманғы ауылшаруашылық кәсіпорындарының жұмыс істеу ерекшелігі-ауылшаруашылық объектісінің жұмыс істеу сценарийлерін анықтайтын өндірістік және нарықтық жағдайлардың көптеген параметрлерінің (өлшемдерінің) болуы, мысалы, шығындар, пайда, тәуекелдер және т.б.. Бұл болжау мен жоспарлаудың белгілі бір проблемасы бойынша шешім қабылдаушы тұлға (ШҚТ) үшін ауылшаруашылық кәсіпорнын өндірудің таңдаулы сценарийін таңдауды қиындатады.

Ауылшаруашылық кәсіпорнының оңтайлы сценарийін іздеудің көп критерилі мәселесін шешу әдісін таңдағанда, осы пәндік аймақтың объектісі үшін ситуациялық модельдің ерекшеліктерін ескеру қажет. Бұл мәселені шешу әдісін таңдаудың алғышарттары ауылшаруашылық өндірісінің жағдайлық моделінің келесі факторлары болып табылады.

Баламаларды бағалау және салыстыру үшін, яғни қалаған сценарийді таңдау туралы шешім қабылдау үшін ақпараттың едәуір көлемін өңдеу қажет. Шешім қабылдаушы тұлға мінез-құлқы шешім таңдау процесіне де, осы процестің нәтижелеріне де айтарлықтай әсер етуі мүмкін.

### **Материалдар мен әдістер.**

Өндіріс сценарийін таңдау үшін көп өлшемді тапсырманың дұрыс ресімделген моделінің күрделілігі эвристикалық процедураны қолданудың орындылығын анықтайды. Басқару тәжірибесі ШҚТ үшін бұл тәсіл қолайлы және уақытты қажет ететіндігін көрсетеді. Агрокәсіпорындарға әсер ететін критерилер бойынша бастапқы ақпаратты сарапшылар сапалы форматта бере алады. Ауылшаруашылық кәсіпорнының ситуациялық моделі үшін сарапшылардың өздері басымдық беру процедурасы орынсыз.

Көптеген әсер ету критерийлері жағдайында оңтайлы сценарийді таңдау бойынша менеджерлердің шешімдерін қабылдауды қолдау міндеті көп өлшемді міндеттер класына жатады. Алгоритмдік деңгейде шешім қабылдаудың көп өлшемді әдістері келесі әдістерді қамтуы мүмкін: негізгі критерий, сызықтық жинақтау, максималды жинақтау, лексиграфиялық оңтайландыру, Нелдер – Мида, бейімделу және т. б.

Жалпы алғанда, шешім қабылдаудың көп критерилі есептеріне арналған математикалық модель келесі жолмен ұсынылуы мүмкін [1].

$$\langle S; E_1, \dots, E_m; M \rangle \quad (1)$$

мұнда  $S$  – көптеген шешім нұсқалары (біздің міндетіміз үшін – ауылшаруашылық өндірісінің балама сценарийлері),  $E_1, \dots, E_m$  – тапсырма критерийлері (біздің міндетіміз үшін – бұл кәсіпорын жұмысының сценарийінің нұсқасын таңдау критерийлері),  $m$  – критерийлер саны ( $m \geq 2$ ),  $M$  – өлшемдерді салыстыру үшін сарапшылардың көптеген артықшылықтары (қатаң емес артықшылық қатынастары).  $E_i(s)$  мәндерінің жиынтығы көптеген  $S$  нұсқаларынан  $s$  шешімінің векторлық бағасын құрайды.

Солтүстік Қазақстанның аграрлық кәсіпорындарында өндіріс үшін басқару шешімін таңдауды оңтайландыру мәселесінің мәнін анықтайтын элементтерді иерархиялық бейнелеудің жүйелі рәсімі ретінде иерархияларды талдаудың эвристикалық әдісі таңдалды. Бұл әдіс қарастырылып отырған мәселенің элементтерін жұптасып салыстыру және нәтижелердің кейінгі иерархиялық синтезі нәтижесінде алынған сараптамалық ақпаратты өңдеуге негізделген.

Бұл әдіс мәселені қарапайым компоненттерге ыдыратудан және шешім қабылдаушылардың пікірлерінің дәйектілігін жұптық салыстыруларға сәйкес одан әрі өңдеуден тұрады. Нәтижесінде иерархиядағы барлық критерийлер үшін зерттелген баламалардың салыстырмалы маңыздылығы анықталады. Салыстырмалы маңыздылық басымдықтардың векторлары түрінде сандық түрде көрсетіледі. Осы формада алынған векторлардың мәндері қатынастар шкаласындағы бағалар болып табылады және қатаң бағалау деп аталады.

Көптеген баламалардан көп өлшемді шешім қабылдау мәселесін талдау үшін біз (1) математикалық модельді (2) шешім нұсқаларын, иерархия деңгейлерінің санын, әр деңгейдегі сарапшылардың артықшылық қатынастарының жиынтығын, әр деңгейдегі критерийлер санын ұсынамыз [2,3,4]:

$$\langle S^1, S^2, \dots, S^z; E_1^1, E_2^1, \dots, E_i^1; E_1^2, E_2^2, \dots, E_j^2; \dots; E_1^z, E_2^z, \dots, E_f^z; M^1, M^2, \dots, M^z \rangle \quad (2)$$

$S^z$  мұнда-көптеген шешімдер (өндіріс сценарийлерінің көптеген балама түрлері),  $z$  – иерархия деңгейлерінің саны ( $z = 1, 2, \dots, Z$ ), – тапсырма критерийлері (сценарий тиімділігінің критерийлері),  $E_1^z, \dots, E_m^z$  – әр деңгейдегі сарапшылардың қалауының көптеген қатынастары, – әр деңгейдегі критерийлер саны.

Көптеген  $S^z$  нұсқалардың ішінен  $S$  шешімінің әр нұсқасы осы опцияның  $P(s)$  векторлық бағасын құрайтын мәндермен сипатталады:

$$p(s) = (E_1(s), \dots, E_m(s)) \quad (3)$$

Бұл преференциялар қатаң емес артықшылық қатынасын қолдана отырып  $M$  -ды  $P: P'Mp''$  модельденеді: яғни  $P'$  векторлық бағалау  $P''$  кем емес артықшылық балама шешімдер бағаланатын критерийлер жиынтығы негізінде жасалады және т.б.. Әрбір критерий үшін  $E_1, \dots, E_m$  оның үлкен мәндері кішіге қарағанда жақсырақ екенін қабылдаймыз. Содан кейін варианттардың векторлық бағалауларында Парето қатынасын анықтауға болады.

### **Нәтижелер және талқылау.**

Аграрлық кәсіпорын өндірісінің әрбір баламалы сценарийін кешенді бағалау барлық басқару өлшемдерінің әсерін ескере отырып жүргізілуге тиіс. Солтүстік Қазақстан Агроөнеркәсіптік кешені өндірісінің қолданбалы ғылыми зерттеулері агрокәсіпорынның жұмыс істеуіне әсер ететін өндірістік-нарықтық жағдайлар өлшемдерінің (параметрлерінің) болуын анықтады [5]. Басқару критерийлері төрт топта ұсынылды:

- 1 топ. Егіс құрылымы және ауыспалы егіс бойынша шарттар
- 2 топ. Кәсіпорын ресурстары
- 3 топ. Нарық сыйымдылығы және келісімшарттық міндеттемелер бойынша талаптар
- 4 топ. Тәуекелдер

Біз ауылшаруашылық өндірісінің сценарийін таңдау бойынша зерттелген шешімдерді кейіннен бағалау үшін барлық топтар үшін басқару критерийлерін анықтаймыз. Басқару критерийлерінің иерархиясы өндіріс сценарийін таңдау міндеттері 1-суретте көрсетілген.

Сарапшылар ұсынған басымдықты таңдау мәселесінің критерийлері сапалы компонентке ие болуы мүмкін екенін ескеріңіз. Сондықтан критерийлерді бағалау бір шкала негізінде сапалыдан сандық деңгейге ауысады.

Иерархиялық талдау әдісін (ИТӘ) әзірлеуші американдық математик Томас Т. Саати салыстырмалы маңыздылық шкаласында 1-ден 9-ға дейінгі шаманы таңдаудың орындылығы туралы математикалық негіздеме береді [6,7,8]. Ол айтарлықтай айырмашылықтары бар ынталандыру (маңыздылығын бағалау ретінде) экспоненциалды түрде орналастырылғанын көрсетті. Сонымен қатар, тиісті қабылдау айтарлықтай айырмашылықтар байқалатын дискретті нүктелерде арифметикалық прогрессияны құрайды. Шамамен бес дискретті нүкте бар, олар көрші бағалар арасындағы саудасаттықты білдіреді.

Т. Саатидің пікірінше, шкаланың жоғарғы шегін белгілеудің бірнеше себептері бар, олар – 9 маңыздылық дәрежесі бойынша сипатталады.

Сапалық айырмашылықтар іс жүзінде маңызды және бір ретті немесе заттарды салыстыратын заттардың санын салыстыру үшін қолданылатын мәнге жақын болған кезде дәлдік элементі болады [8,9,10].



1 сурет – Агроөндіріс сценарийін таңдау критерийлерінің иерархиясы

Айта кету керек, әр топтағы критерийлердің саны ауылшаруашылық кәсіпорындарының әлеуметтік-экономикалық жағдайларының өзгеруіне байланысты өзгеруі мүмкін.

### Қорытынды

Солтүстік Қазақстанның аграрлық кәсіпорындары үшін өндіріс сценарийінің оңтайлы нұсқасын таңдау проблемасы тән. Мәселе өндірістік және нарықтық жағдайлардың көптеген өлшемдерінің (параметрлерінің) болуына және шешім қабылдаушының процеске қатысу ерекшеліктеріне байланысты. Өлшемдердің маңыздылығы туралы ақпаратты өндірістік сарапшылардан алу көзделеді.

Ауылшаруашылық кәсіпорнының жұмыс істеуінің қолайлы сценарийі бойынша басқарушылық шешім қабылдау сапасы сараптамалық бағалауды есепке алудың дұрыстығына және жергілікті атқарушы органдарға қатысу дәрежесіне байланысты анықталады. Дамудың бастапқы позициясы келесідей. Шешім қабылдаушыға кез-келген "дұрыс" шешімді тағайындамау үшін, оған интерактивті режимде мәселенің мәні мен оны шешуге қойылатын талаптарды түсінуіне сәйкес келетін өндіріс сценарийінің нұсқасын табуға мүмкіндік беру.

Зерттелетін проблема үшін иерархиялық талдау әдісін қолданудың негізделген тәсілі шешім қабылдау міндетін бастапқы шарттарға сәйкес ұтымды түрде құрылымдауға мүмкіндік берді: жеке құрамның қатысу сипаты және сараптамалық бағалауды қолдану.

Солтүстік Қазақстанның агрокәсіпорны мысалында өндіріс сценарийін таңдау міндеттің доминантты иерархиясымен әзірленген модель параметрлердің кең диапазонында есептеулерді жүзеге асыруға мүмкіндік береді. Модельдің бұл сапасы Солтүстік Қазақстанның осындай агрокәсіпорындарында өндірістің оңтайлы сценарийлерін болжау бойынша ауқымды есептеулерді орындау үшін оның әмбебаптығын айқындады.

#### Қолданылған әдебиеттер тізімі

- 1 Подиновский В.В. Теоретические основы выработки решений в сложных ситуациях. М.: МО СССР, 1988.
- 2 Т.Саати. Принятие решений. Метод анализа иерархий // Thomas L. Saaty «The Analytic Hierarchy Process». (In Russian). - 1993. - 320 с.
- 3 Халин В.Г. и др. Теория принятия решений в 2 т. – М.: Изд. ЮРАЙТ, 2020.
- 4 Figueira J., Greco S., Ehrgott M. Multiple Criteria Decision Analysis: State of the Art Surveys. – New York: Springer Science & Business Media, 2005. – 1045 p.
- 5 AHP Software for Decision Making and Risk Assessment / <https://www.expertchoice.com/2021>.
- 6 Transparent Choice / <https://www.transparentchoice.com/>
- 7 Super Decisions / <https://www.superdecisions.com/>
- 8 Mauina G.M., Chertkova E.A., Nukusheva S.A., Aitimova U.Zh., Ismailova A.A. Expert-statistical method of management decision support for agricultural enterprises of Northern Kazakhstan // Journal of Theoretical and Applied Information Technology, – 2021. – Vol 99, Issue 12. P. 3071-3083.
- 9 Трансферт и адаптация технологий по точечному земледелию при производстве продукции растениеводства по принципу «демонстрационных хозяйств» (полигонов) в Северо-Казахстанской области [Текст]: отчет о НИР: / КАТУ им. С. Сейфуллина; рук. Куришбаев А.К. – Нур-Султан., 2019. – 349 с. № ГР 0118РК01393.
- 10 Саати Т., Керис К. Аналитическое планирование. Организация систем: Пер. с англ. – М.: Радио и связь, 1991. – 224 с

## АДАМ ДЕНЕСІНДЕГІ ТЕМПЕРАТУРАНЫ ЦИЛИНДРЛІК БӨЛУДІҢ МАТЕМАТИКАЛЫҚ МОДЕЛІ

*Қабдешов Толағай, Хажсайв Кумис*

*Дәулет Серікбаев атындағы Шығыс Қазақстан техникалық университеті*

*Өскемен қаласы*

*[tolagai27@gmail.com](mailto:tolagai27@gmail.com), [khajaiv.kumys@mail.ru](mailto:khajaiv.kumys@mail.ru)*

Аңдатпа

Бұл мақаланың мақсаты-адам ағзасындағы температураның таралуының математикалық моделін құру. Бұл зерттеуде денені цилиндр ретінде қарастыра және деректерді радиалды бағытта талданған. Бессельдің модификацияланған дифференциалдық теңдеуі және модификацияланған функциясы биологиялық жылу үшін Пеннс теңдеуінің аналитикалық шешімін алу үшін қолданылады. Дене температурасының таралу әсерін қоршаған орта температурасы, тіндердің жылу өткізгіштігі, метаболизм жылдамдығы, қанның перфузия жылдамдығы, жылу беру коэффициенттері сияқты әртүрлі параметрлерден өлшенді. Сандық нәтижелер әртүрлі жағдайларда графиктерді құру үшін қолданылады.

Кілттік сөздер: Бессель функциясы, биологиялық жылу үшін Пеннс теңдеуі, жылу балансы математикалық пәндік жіктеу.

Кіріспе

Адам ағзасы-бұл тінге енетін қан тамырларының күрделі желісі, ал қан ағымы биологиялық жүйеде жылу берудің маңызды механизмі. Метаболизм ішкі энергияның маңызды көзі. Өмірдің осы екі жылу аспектісі өмір сүруді жалғастырады. Қан ағымы тірі организмнің терморегуляциясының негізгі элементі өте күрделі болып табылады.

Дене ядросының орташа температурасы 37°C градус. Бұл дене температурасы жылу өндірісі мен жылу жоғалту арасындағы тепе-теңдіктің нәтижесі. 27°C-тан төмен температура және 42°C-тан жоғары температура өте маңызды, бұл жағдайда өлім себебі болуы мүмкін. Сондықтан дене температурасын 37°C деңгейінде ұстау керек. Қан тамырларындағы жылу алмасу қоршаған орта температурасының өзгеруіне қарамастан дене ядросының біркелкі температурасын сақтауға көмектеседі[2]. Жылу алмасу ядродан, ұлпалар арқылы бетіне жылу ағынын реттейтін ішкі өткізгіштікке байланысты, перифериялық қан ағымымен, дене мен тері арасындағы температура градиентімен және тіндік дененің өткізгіштігімен реттелетін жылу берудің құрамдас бөлігі. Қан ағымы перифериялық өткізгіштіктің көп бөлігін қамтамасыз етеді, онда қан мен тіндер арасында конвекция және артериялар мен тамырлар арасында кері жылу алмасу болады. Қан ағымы дененің метаболикалық қажеттіліктеріне, сондай-ақ тиісті дене температурасын ұстап тұру қажеттілігіне сәйкес реттеледі[7].

Метаболикалық жылу дененің әр бөлігіндегі метаболикалық және химиялық реакциялар нәтижесінде біркелкі шығарылады, бірақ жылдамдық әрқашан бірдей бола бермейді. Бұл реакция болмаса, өмірді жалғастыру мүмкін емес еді.

Ғалымдар физиологиялық жүйені зерттеуге және оны медициналық зерттеулердің басынан бастап қолдануға мүдделі болды[3]. Олар қоршаған ортаның әртүрлі жағдайларына реакцияларды модельдеу үшін адамның жылу жүйесінің ең дәл моделін табуға тырысты. 1876 жылы Бернардтың эксперименттік зерттеулерінен бастап, қан тамырлары жүйесі мен тамырлар арасындағы күрделі жылу әрекеттесуінің математикалық модельдеуі басталды. Тамыр жүйесі мен тіндердің өзара әрекеттесуі көптеген физиологтар, дәрігерлер мен инженерлер үшін қызығушылық тудыра бастады.

1948 жылы Пеннс қан перфузиясы мен жылу тасымалындағы жаппай тасымалдау арасындағы математикалық байланысты сипаттайтын іргелі жұмысты жариялады. Оның жұмысы тоғыз адамның білектеріндегі радиалды позицияға байланысты температураның таралуын өлшеу бойынша бірқатар тәжірибелерден тұрды[5]. Қарапайым және икемді

болғандықтан, Пеннс моделі өтпелі температураны болжау үшін әлі де қолайлы нәтижеге ие, дегенмен көптеген зерттеушілер перфузия жылдамдығын және жүрек қан температурасы мен жергілікті тіндердің температурасы арасындағы айырмашылықты сипаттайтын балама модельдер жасады[1].

Әдістер

Пеннс метаболизм мен қанның перфузиясының әсерін, сондай-ақ атриальды қан температурасының интракраниальды температураның таралу әсерін ескеретін моделін жасады[4]. Бұл екі әсер жылу диффузиясының стандартты теңдеуіне енгізілді

$$\rho c \frac{\partial T}{\partial t} = \nabla(K\nabla T) + M(T_a - T) + S \quad (1)$$

мұндағы,  $\rho$  =терінің тығыздығы,  $c$  =терінің нақты жылу сыйымдылығы,  $K$  = жылу өткізгіштік,  $M = \rho_b w_b c_b S = qm$  =метаболикалық жылу  $\rho_b$  = қанның тығыздығы,  $w_b$ =қанның бірдік көлемдегі перфузия жылдамдығы,  $c_b$  =қанның нақты жылу жылдамдығы,  $T_a$  =артерия арқылы тінге түскен кездегі қан температурасы.

Адам денесінің бөліктері цилиндрге ұқсайды, сондықтан модельде цилиндрлік координаттарды  $(r, T, z)$  температура профильдерін зерттеу үшін пайдалану ыңғайлы.

$$\rho c = \frac{\partial T}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left( K \frac{\partial T}{\partial r} \right) + \frac{1}{r^2} \frac{\partial}{\partial \theta} \left( K \frac{\partial T}{\partial \theta} \right) + \frac{\partial}{\partial z} \left( K \frac{\partial T}{\partial z} \right) + M(T_a - T) + S \quad (2)$$

Аналитикалық тәсілді қолдана отырып, адам ағзасындағы бір өлшемді стационарлық жағдайда температураны бөлу моделін құра отырып, теңдеудің күрделілігін жою үшін  $\gamma$  және  $z$  тұрақты болып саналады. Тек 'r' радиалды бағыты алынады[6]. Осылайша, бір өлшемді зерттеуде - күй жағдайы - (2) теңдеу келесі шекаралық жағдайды қолдана отырып, тұрақты күй жағдайы үшін

$$\frac{d}{dr} \left( r \frac{dT}{dr} \right) + \frac{M}{K} (T_a - T) + \frac{S}{K} = 0 \quad (3)$$

цилиндрлік пішінге радиалды бағытта беріледі.

$$r = 0, \quad \frac{dT}{dr} = 0 \quad (4)$$

$$r = R, \quad -K \frac{dT}{dr} = h_A (T - T_\infty) \quad (5)$$

мұндағы,  $R$ -терінің радиусы;  $h_A$  = терінің бетіндегі жылу беру коэффициенті,  $T_\infty$  = атмосфераның температурасы.

Теңдеу (3) шекаралық шартпен, сипаттамалық шамаларды енгізе отырып, өлшемсіз өңдеуді орындайды.

$$r^* = \frac{r}{R}, \quad \text{және} \quad T^* = \frac{T - T_\infty}{T_a - T_\infty} \quad (6)$$

Дифференциядан кейін (6) 'r' және теңдеуді ауыстыру (3) келесіге дейін азаяды

$$\frac{1}{r^*} \frac{d}{dr^*} \left( r^* \frac{dT^*}{dr^*} \right) + \frac{R^2 M (1 - T^*)}{K} + \frac{SR^2}{K(T_a - T_\infty)} = 0 \quad (7)$$

Өлшемсіз параметрлерді қайтадан енгізіледі,

$$M^* = \frac{R^2}{K}, \quad S^* = \frac{SR^2}{K(T_a - T_\infty)} \quad (8a)$$

$$h_A = \frac{h_A R}{K} \quad (8b)$$

(8a) және (8b) көмегімен келесідей теңдеуге дейін азаяды

$$\frac{1}{r^*} \frac{d}{dr^*} \left( r^* \frac{dT^*}{dr^*} \right) - M^* T^* + (M^* + S^*) = 0 \quad (9)$$

Есептеудің қарапайымдылығы үшін,  $M^* + S^* = U$  және  $M^* = V$  және  $\emptyset = U - VT$  қойып, (10) теңдеуді келесі түрде алынады

$$r^{*2} \frac{d^2 \varphi}{dr^{*2}} + r^* \frac{d\varphi}{dr^*} - Vr^* \varphi = 0 \quad (10)$$

бұл Бессельдің модификацияланған теңдеуі

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} - (x^2 + p^2)y = 0$$

$$[x = r^*, \beta = \sqrt{v}, p = 0]$$

Оның шешімі куат сериясы түрінде берілген,  $y = AI_p(\beta x) + BK_p(\beta x)$

Сонымен,  $\varphi$  үшін теңдеудің шешімі келесідей

$$\varphi = C_1 I_0(\sqrt{V} r^*) + C_2 K_0(\sqrt{V} r^*) \quad (11)$$

$$\frac{d\varphi}{dr^*} = C_1 I_1(\sqrt{V} r^*) - C_2 K_1(\sqrt{V} r^*) \quad (12)$$

мұндағы,  $C_1$  және  $C_2$ -шекаралық шарттардан анықталған ерікті тұрақтылар

$$\frac{dT^*}{dr^*} = 0 \quad r^* = 0$$

$$I_1(0) = 0$$

$$I_1(\sqrt{V} r^*) = 0$$

$$K_1(\sqrt{V} r^*) \neq 0, C_2 = 0$$

$$\varphi = U - VT$$

$$\frac{dT^*}{dr^*} = -\frac{1}{V} [C_1 \sqrt{V} I_1(\sqrt{V} r^*)] \quad (13)$$

Шекаралық шарттармен есеп айырысу

$$\frac{dT^*}{dr^*} = -h_A^* T^* \quad r = R \quad \frac{r}{R} = 1 \quad R^* = 1 \quad (14)$$

алынады,

$$C_1 = \frac{h_A^*}{h_A^* I_0(\sqrt{M^*} r^*) + \sqrt{M^*} I_1(\sqrt{M^*} r^*)} \quad (15)$$

Тағы да, есептеулер бойынша  $T^*$  және  $T$  үшін шешім жазылуы мүмкін.

Осылайша теңдеу (12)

$$T^* = \frac{M^* + S^*}{M^*} - \frac{1}{M^*} \left[ \frac{h_A^* I_0(\sqrt{M^*} r^*)}{h_A^* I_0(\sqrt{M^*} r^*) + \sqrt{M^*} I_1(\sqrt{M^*} r^*)} \right] \quad (16)$$

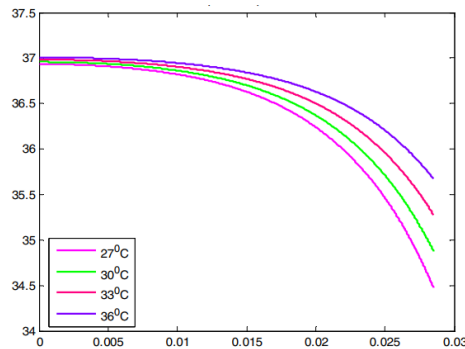
$$T = T_\infty + (T_a - T_\infty) \left[ \frac{U}{V} - \frac{1}{V} \left\{ \frac{h_A^* I_0(\sqrt{V} r^*)}{h_A^* I_0(\sqrt{V} r^*) + \sqrt{V} I_1(\sqrt{V} r^*)} \right\} \right] \quad (17)$$

Тірі ұлпаның цилиндрлік пішіні үшін (17) теңдеудегі аналитикалық шешімдер және физикалық қасиеттердің сандық нәтижелері көптеген факторларға байланысты.

Нәтиже

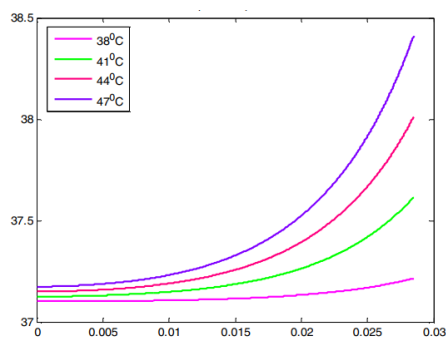
I. Атмосфералық температураның әсері

Әр түрлі атмосфералық температураларда температураны бөлу профильдері 1(a). және 1(b). суреттерінде көрсетілген.



Сурет 1(a). Атмосфералық температура 37°C төмен



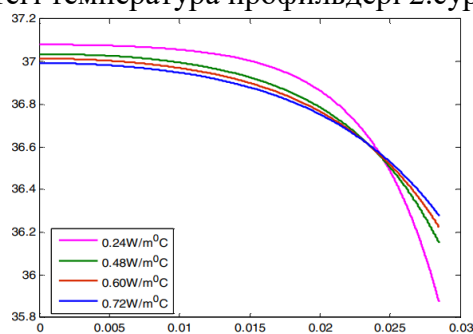


Сурет 1(b). Атмосфералық температура  $37^{\circ}\text{C}$  жоғары

1(a) суретте, егер атмосфералық температура дене температурасынан төмен болса, онда дене температурасы дене ядросынан тері бетіне қарай төмендейтінін көрсетеді. Себебі, дененің негізгі температурасы атмосфераның температурасынан жоғары, ал 1(b) суретте, егер атмосфералық температура температурадан жоғары болса, онда дене температурасы дененің өзегінен терінің бетіне қарай жоғарылайды.

## II. Жылу өткізгіштіктің әсері

Әр түрлі жылу өткізгіштегі температура профильдері 2.суретте көрсетілген.

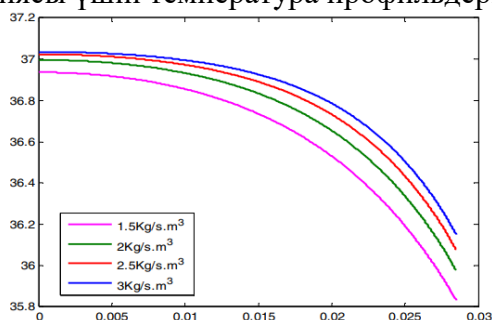


Сурет 2. Жылу өткізгіштіктің әсері

2.суретте Дене тіндерінің жылу өткізгіштігінің жоғарылауымен дененің ішкі бөлігіндегі дене температурасының жоғарылауы көрсетілген, бірақ дененің сыртқы бөлігінде жылу өткізгіштіктің жоғарылауымен температура төмендейді. Жылу өткізгіштік жоғарылаған сайын температура төмендейді. Бұл дененің бетіндегі конвективті шекарамен байланысты.

## III. Қан перфузиясының әсері

Қанның әртүрлі перфузиясы үшін температура профильдері 3.суретте көрсетілген.

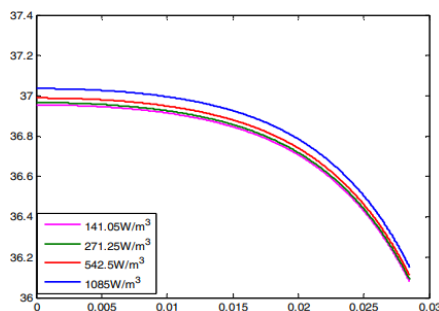


Сурет 3. Қанның перфузиясының әсері

Қанның перфузиясының жоғарылауы адам денесінің беткі температурасына айтарлықтай әсер етеді. 3.суретте Радиалды бағыттағы температура градиенті қанның перфузиясының жоғарылауымен төмендейді, бұл қанның перфузиясынан туындаған жылу бөлу жылдамдығының жоғарылауының нәтижесі.

## IV. Метаболикалық жылудың әсері

Әр түрлі метаболикалық жылу шығару кезінде температура профильдері 4.суретте көрсетілген.

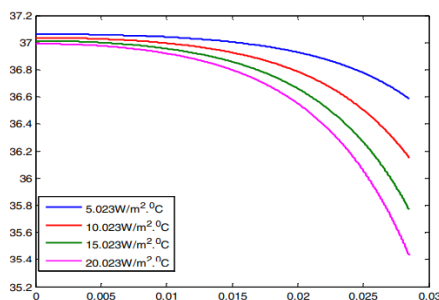


Сурет 4. Метаболикалық жылудың әсері

Метаболикалық жылудың бөлінуі адам ағзасындағы температураның таралуына өте аз әсер етеді. Метаболикалық жылудың  $542 \text{ Вт/м}^3$ -ден  $1085 \text{ Вт/м}^3$ -ге дейін өзгеруі дене температурасының шамамен  $0,1^\circ\text{C}$  өзгеруіне әкелді, ал терінің бетіне оның әсері шамалы. Бұл тері бетіндегі қан тамырларының аздап шоғырлануына байланысты.

#### V. Жылу беру коэффициенттерінің әсері

Жылу беру коэффициентінің әртүрлі мәндеріндегі температура профильдері 5.суретте көрсетілген.



Сурет 5. Жылу беру коэффициентінің әсері

5.суретте Жылу беру коэффициенттерінің тері бетіне айтарлықтай әсері көрсетілген. Бұл тері бетінің қоршаған ортамен тікелей байланысты.

#### Қорытынды

Аналитикалық модель жасалынып, Бессельдің модификацияланған дифференциалдық теңдеуі мен модификацияланған функциясын қолдана отырып, зерттелетін күйдің бір өлшемді жағдайында био-жылу үшін Пеннс теңдеуінің шешімі алынды.

Бұл модельдегі шешім атмосфералық температура, жылу өткізгіштік, метаболикалық жылу шығару жылдамдығы, қанның перфузия жылдамдығы және жылу беру коэффициенті сияқты әртүрлі температуралық профильдердің өзгеруі адам денесінің температурасының таралуына айтарлықтай әсер ететінін көрсету үшін қолданылады. Аналитикалық шешім жылу диагностикасы мен емдеуге қолданылуы мүмкін, бұл жағымсыз жанама әсерлерді барынша азайту кезінде терапевтік әсерің барынша арттырады.

#### Әдебиеттер тізімі

1. Arkin X., Sue Luxe және Holmes K. R., IEEE Transactions on Biomedical Engineering 10141 Nov - Feb (1994).

[\(PDF\) Recent developments in modeling heat transfer in blood perfused tissues \(researchgate.net\)](#)

2. Gurung D. B, Saxena V. P және Adhikari P. R, J. Appl. Math and information (2009)27, P.301-313.

- [Numerical-Solutions-of-One-Dimensional-Bioheat-Transfer-Equation-in-Cylindrical-Living-Tissues.pdf \(yandex.kz\)](#),
3. He Ying, Liu Hao, Himeno Ryutaro, Mass Transfer 47 (2004) 2735-2745.  
[A One-dimensional Thermo-fluid Model of Blood Circulation in Upper limb of Man | Request PDF \(researchgate.net\)](#)
  4. He Ying, жапон журналы Computers in Biology and Medicine, (2006), 36, P.1336-1350.  
[An advanced computational bioheat transfer model for a human body with an embedded systemic circulation | SpringerLink](#)
  5. Kai Yue Xinxin Zhang, Fan You, An Analytic Solution, 13(3), (2004), 52-54 б.  
[cbm06-heyang.pdf \(yandex.kz\)](#)
  6. Minkowycz W. J және Sparrow E. M, Vol. III, (2009).  
[Advances in Numerical Heat Transfer, Volume 3 - W. J. Minkowycz - Google Книги](#)
  7. Моис М.Ф. Press Maastricht, (1982) .  
[37591032685348.pdf \(yandex.kz\)](#)

# СИСТЕМА ФУНКЦИЙ ТИПА ХААРА

Орман Индира Маликовна

[Indira.malikovna@mail.ru](mailto:Indira.malikovna@mail.ru)

ЕНУ им.Л.Н.Гумилева, Нур-Султан, Казахстан  
Старший преподаватель

**Аннотация:** В работе рассмотрена проблема построения ортонормированной системы между функциями и доказано, что система типа Хаар, как и система Хаар, является ортонормированной системой. Кроме того, определены коэффициенты классификации данной степенной функции по ортонормированной системе типа Хаар.

**Введение.** В работе изучается построение ортонормированной системы функций на сегменте  $[0, 1]$ . Доказано, что введенная система функций является ортонормированной системой функций типа Хаара.

Основная часть. Основные определения и необходимые утверждения. Определения основных понятий

Определение 1. Пусть дана функция  $f$ , определенная на сегменте  $[a, b]$ . Если для функции  $f$  выполнено

$$\int_a^b |f| dx < \infty, \quad (1)$$

то говорят, что функция  $f$  интегрируема по Лебегу на  $[a, b]$ . Множество всех интегрируемых по Лебегу на сегменте  $[a, b]$  функции называется пространством Лебега и обозначается через  $L(a, b)$ .

Определение 2. Пусть даны  $f, g \in L(a, b)$ . Выражение вида

$$(f, g) = \int_a^b |f \cdot g| dx \quad (2)$$

называется скалярным произведением функций  $f, g \in L(a, b)$ .

Определение 3. Пусть дана последовательность функций  $\{\varphi_n\}_{n=1}^{\infty}$ , интегрируемых на сегменте  $[a, b]$ . Если выполнены

$$(\varphi_i, \varphi_k) = \begin{cases} 1, & i = k, \\ 0, & i \neq k \end{cases} \quad (3)$$

то последовательность  $\{\varphi_n\}_{n=1}^{\infty}$  называется ортонормированной.

Определение 4. Последовательность функций  $\chi_m^k(x)$ , определенных и ортонормированных на сегменте  $[0, 1]$  называется системой Хаара, если:

1. При  $m = 0, k = 0$

$$\chi_0^0(x) = 1, \quad \chi_0^{(1)}(x) = \begin{cases} 1, & x \in \left[0, \frac{1}{2}\right), \\ 0, & x = \frac{1}{2}, \\ -1, & x \in \left(\frac{1}{2}, 1\right]; \end{cases}$$

2. При  $m \geq 1, 1 \leq k \leq 2^m$

$$\chi_m^{(k)}(x) = \begin{cases} \sqrt{2^m}, & x \in \left( \frac{k-1}{2^m}, \frac{k-\frac{1}{2}}{2^m} \right), \\ -\sqrt{2^m}, & x \in \left( \frac{k-\frac{1}{2}}{2^m}, \frac{k}{2^m} \right), \\ 0, & x \in \left( \frac{l-1}{2^m}, \frac{l}{2^m} \right), l \neq k, 1 \leq l \leq 2^m. \end{cases}$$

Определение 5. Пусть дано число  $\alpha > -1$ . Последовательность функций  $\chi_m^k(x, \alpha)$ , определенных и ортонормированных на сегменте  $[0, 1]$  называется системой Хаара, если:

1. При  $m=1, k=1$

$$\chi_1^1(x, \alpha) = \begin{cases} \sqrt{\frac{2^{2(1+\alpha)}}{2}} & x \in \left( 0, \frac{1}{2^{2(1+\alpha)}} \right) \\ -\sqrt{\frac{2^{2(1+\alpha)}}{2(2^{1+\alpha}-1)}} & x \in \left( \frac{1}{2^{2(1+\alpha)}}, \frac{1}{2^{1+\alpha}} \right) \\ 0, & x=0, \quad x \in \left( \frac{(l-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{l^{1+\alpha}}{2^{m(1+\alpha)}} \right), l \neq 1 \end{cases}$$

2 При  $m=1, k=2$ ,

$$\chi_1^2(x, \alpha) = \begin{cases} \sqrt{\frac{2^{2(1+\alpha)}}{2}} & x \in \left( \frac{1}{2^{1+\alpha}}, \frac{1+2^{1+\alpha}}{2^{2(1+\alpha)}} \right) \\ -\sqrt{\frac{2^{2(1+\alpha)}}{2[2^{1+\alpha}(2^{1+\alpha}-1)]}} & x \in \left( \frac{1+2^{1+\alpha}}{2^{2(1+\alpha)}}, 1 \right) \\ 0, & x=0, \quad x \in \left( \frac{(l-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{l^{1+\alpha}}{2^{m(1+\alpha)}} \right), l \neq 1 \end{cases}$$

Для промежутков

$$(0, 1) = \bigcup_{k=1}^m \left( \frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{k^{1+\alpha}}{2^{m(1+\alpha)}} \right)$$

$$\chi_m^k(x, \alpha) = \begin{cases} \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2[(k-1)^{1+\alpha}(1-2^{1+\alpha})+k^{1+\alpha}]}} & x \in \left( \frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{(k-1)^{1+\alpha}+k^{1+\alpha}}{2^{(m+1)(1+\alpha)}} \right) \\ -\sqrt{\frac{2^{(m+1)(1+\alpha)}}{2[k^{1+\alpha}(2^{1+\alpha}-1)-(k-1)^{1+\alpha}]}} & x \in \left( \frac{(k-1)^{1+\alpha}+k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}, \frac{k^{1+\alpha}}{2^{m(1+\alpha)}} \right) \\ 0, & x=0, x \in \left( \frac{(l-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{l^{1+\alpha}}{2^{m(1+\alpha)}} \right), l \neq k \end{cases}$$

**Дискуссия.** Необходимые утверждения

Лемма 1. Для  $-1 < \alpha \leq 0$ ,  $k = 1, 2, \dots, 2^m$ ,  $n = m + 1$ ,  $n, m \in \mathbb{N}$  и  $k' \in \mathbb{N}$  такого, что

$$2k - 1 \leq k' \leq k \left[ \left( 1 - \frac{1}{k} \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}} \quad (4)$$

имеет место

$$\left( \frac{(k'-1)^{1+\alpha}}{2^{n(1+\alpha)}}, \frac{(k')^{(1+\alpha)}}{2^{n(1+\alpha)}} \right) \subset \left( \frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{k^{(1+\alpha)} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}} \right) \quad (5)$$

и для

$$1 + k \left[ \left( 1 - \frac{1}{k} \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}} \leq k' \leq 2k \quad (6)$$

справедливо

$$\left( \frac{(k'-1)^{1+\alpha}}{2^{n(1+\alpha)}}, \frac{(k')^{(1+\alpha)}}{2^{n(1+\alpha)}} \right) \subset \left( \frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}, \frac{k^{1+\alpha}}{2^{m(1+\alpha)}} \right). \quad (7)$$

Доказательство. При  $k = 1$  соотношения (5), (6) очевидны. Пусть  $k > 1$ . Если  $n = m + 1$ , то для натуральных чисел  $k, k'$  и  $-1 < \alpha \leq 0$  удовлетворяющих условию

$$2k - 1 \leq k' \leq k \left[ \left( 1 - \frac{1}{k} \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}},$$

из неравенства

$$(2k - 1)^{1+\alpha} \leq (k - 1)^{1+\alpha} + k^{1+\alpha}$$

следует

$$\frac{(k')^{(1+\alpha)}}{2^{n(1+\alpha)}} \leq \frac{(k - 1)^{(1+\alpha)} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}} \quad \text{и} \quad \frac{(k' - 1)^{(1+\alpha)}}{2^{n(1+\alpha)}} \geq \frac{(k - 1)^{(1+\alpha)}}{2^{m(1+\alpha)}}.$$

Из последних двух неравенств

$$\left( \frac{(k' - 1)^{1+\alpha}}{2^{n(1+\alpha)}}, \frac{(k')^{(1+\alpha)}}{2^{n(1+\alpha)}} \right) \subset \left( \frac{(k - 1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{(k - 1)^{(1+\alpha)} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}} \right).$$

Теперь, при условии

$$1 + k \left[ \left( 1 - \frac{1}{k} \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}} \leq k' \leq 2k \quad \text{верны} \quad \frac{(k')^{(1+\alpha)}}{2^{n(1+\alpha)}} = \frac{(k')^{(1+\alpha)}}{2^{(m+1)(1+\alpha)}} \leq \frac{(2k)^{(1+\alpha)}}{2^{(m+1)(1+\alpha)}} = \frac{k^{(1+\alpha)}}{2^{m(1+\alpha)}}$$

и  $\frac{(k' - 1)^{(1+\alpha)}}{2^{n(1+\alpha)}} \geq \frac{(k - 1)^{(1+\alpha)} + k^{(1+\alpha)}}{2^{(m+1)(1+\alpha)}}$ , а из двух последних неравенств, имеем

$$\left( \frac{(k' - 1)^{1+\alpha}}{2^{n(1+\alpha)}}, \frac{(k')^{(1+\alpha)}}{2^{n(1+\alpha)}} \right) \subset \left( \frac{(k - 1)^{(1+\alpha)} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}, \frac{k^{1+\alpha}}{2^{m(1+\alpha)}} \right).$$

Лемма 1 доказана.

Лемма 2. Пусть  $-1 < \alpha \leq 0$ ,  $k = 1, 2, \dots, 2^m$  и  $n > m + 1$ ,  $n, m \in \mathbb{N}$ . Тогда для  $k' \in \mathbb{N}$  справедливы:

Если

$$2^{n-m}(k - 1) + 1 \leq k' \leq 2^{n-m-1}k \left[ \left( 1 - \frac{1}{k} \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}}, \quad (8)$$

То при  $k = 1, 2, \dots, 2^m$

$$\left( \frac{(k'-1)^{1+\alpha}}{2^{n(1+\alpha)}}, \frac{(k')^{(1+\alpha)}}{2^{n(1+\alpha)}} \right) \subset \left( \frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{(k-1)^{(1+\alpha)} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}} \right). \quad (9)$$

Если натуральное число  $k'$  удовлетворяет условию

$$2^{n-m-1} k \left[ \left( 1 - \frac{1}{k} \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}} + 1 \leq k' \leq 2^{n-m} k, \quad (10)$$

то

$$\left( \frac{(k'-1)^{1+\alpha}}{2^{n(1+\alpha)}}, \frac{(k')^{(1+\alpha)}}{2^{n(1+\alpha)}} \right) \subset \left( \frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}, \frac{k^{1+\alpha}}{2^{m(1+\alpha)}} \right). \quad (11)$$

Доказательство. Если  $k=1$ , то  $k'=1$  и верны (5), (6). Пусть  $k>1$ . Рассмотрим следующие случаи.

Пусть  $n > m+1$  и выберем число  $k'$  так, чтобы удовлетворялось (8), т.е.

$$2^{n-m}(k-1)+1 \leq k' \leq 2^{n-m-1} k \left[ \left( 1 - \frac{1}{k} \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}}.$$

Тогда верно

$$\frac{(k'-1)^{1+\alpha}}{2^{n(1+\alpha)}} \geq \frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}.$$

Действительно, при  $k' \geq 2^{n-m}(k-1)+1$  выполнены

$$k'-1 \geq 2^{n-m}(k-1) \quad \text{и} \quad (k'-1)^{1+\alpha} \geq 2^{(n-m)(1+\alpha)}(k-1), \quad \text{отсюда}$$

$$\frac{(k'-1)^{1+\alpha}}{2^{n(1+\alpha)}} \geq \frac{2^{(n-m)(1+\alpha)}(k-1)^{1+\alpha}}{2^{n(1+\alpha)}} = \frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}.$$

Пусть имеет место  $k' \leq 2^{n-m-1} k \left[ \left( 1 - \frac{1}{k} \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}}$ . Тогда  $k>1$ ,  $n > m+1$  и для числа

$$-1 < \alpha \leq 0 \quad 2^{n-m} k \left[ \left[ \left( 1 - \frac{1}{k} \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}} - 1 \right] \geq 1 - 2^{n-m}.$$

Следовательно  $2^{n-m-1} k \left[ \left( 1 - \frac{1}{k} \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}} \geq 2^{n-m}(k-1)+1$ , Отсюда для чисел  $k' \in N$ ,

удовлетворяющих неравенствам  $2^{n-m}(k-1)+1 \leq k' \leq 2^{n-m-1} k \left[ \left( 1 - \frac{1}{k} \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}}$ ,

справедливы неравенства (5) и (7), поэтому

$$\left( \frac{(k'-1)^{1+\alpha}}{2^{n(1+\alpha)}}, \frac{(k')^{(1+\alpha)}}{2^{n(1+\alpha)}} \right) \subset \left( \frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{(k-1)^{(1+\alpha)} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}} \right).$$

2. Для чисел  $k \geq 2$  и  $n > m+1$ , если число  $k' \in N$  удовлетворяет неравенствам

$$2^{n-m-1} k \left[ \left( \frac{1}{k} - 1 \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}} + 1 \leq k' \leq 2^{n-m} k, \text{ то}$$

$$\left( \frac{(k'-1)^{1+\alpha}}{2^{n(1+\alpha)}}, \frac{(k')^{(1+\alpha)}}{2^{n(1+\alpha)}} \right) \subset \left( \frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}, \frac{k^{1+\alpha}}{2^{m(1+\alpha)}} \right).$$

Лемма 2 доказана.

**Методы.** Система функций типа Хаара. Ортонормированность системы типа Хаара

**Теорема 1.** Пусть даны числа  $-1 < \alpha \leq 0$  и  $m \geq 1$  целое неотрицательное. Тогда для  $k = 1, 2, \dots, 2^m$  система функций  $\chi_m^{(k)}(x, \alpha)$  определенных на  $[0, 1]$  является системой типа Хаара, т.е. ортонормированной системой:

1. При  $m = 1, k = 1$

$$\chi_1^1(x, \alpha) = \begin{cases} \sqrt{\frac{2^{2(1+\alpha)}}{2}} & x \in \left(0, \frac{1}{2^{2(1+\alpha)}}\right) \\ -\sqrt{\frac{2^{2(1+\alpha)}}{2(2^{1+\alpha}-1)}} & x \in \left(\frac{1}{2^{2(1+\alpha)}}, \frac{1}{2^{(1+\alpha)}}\right) \\ 0, & x = 0, \quad x \in \left(\frac{(l-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{l^{1+\alpha}}{2^{m(1+\alpha)}}\right), l \neq 1 \end{cases} \quad (12)$$

2. При  $m = 1, k = 2,$

$$\chi_1^2(x, \alpha) = \begin{cases} \sqrt{\frac{2^{2(1+\alpha)}}{2}} & x \in \left(\frac{1}{2^{(1+\alpha)}}, \frac{1+2^{(1+\alpha)}}{2^{2(1+\alpha)}}\right) \\ -\sqrt{\frac{2^{2(1+\alpha)}}{2[2^{1+\alpha}(2^{1+\alpha}-1)]}} & x \in \left(\frac{1+2^{(1+\alpha)}}{2^{2(1+\alpha)}}, 1\right) \\ 0, & x = 0, \quad x \in \left(\frac{(l-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{l^{1+\alpha}}{2^{m(1+\alpha)}}\right), l \neq 1 \end{cases} \quad (13)$$

3. Для  $m > 1, k = 1, 2, \dots, 2^m$  и

$$(0, 1) = \bigcup_{k=1}^m \left( \frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{k^{1+\alpha}}{2^{m(1+\alpha)}} \right)$$

$$\chi_m^k(x, \alpha) = \begin{cases} \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2[(k-1)^{1+\alpha}(1-2^{1+\alpha}) + k^{1+\alpha}]}} & x \in \left(\frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}\right) \\ -\sqrt{\frac{2^{(m+1)(1+\alpha)}}{2[k^{1+\alpha}(2^{1+\alpha}-1) - (k-1)^{1+\alpha}]}} & x \in \left(\frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}, \frac{k^{1+\alpha}}{2^{m(1+\alpha)}}\right) \\ 0, & x = 0, x \in \left(\frac{(l-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{l^{1+\alpha}}{2^{m(1+\alpha)}}\right), l \neq k \end{cases} \quad (14)$$

**Замечание.** Очевидно, что в теореме, вместо  $-1 < \alpha \leq 0$  можно ввести обозначение  $\tau = -\alpha$ , тогда имеет место  $0 \leq \tau < 1$ .

**Доказательство.** Покажем, что система нормирована

$$\|\chi_m^{(k)}(x, \alpha)\|_{L^2(0,1)} = 1. \quad (15)$$

Сначала, покажем, что при  $m = 1$  и  $k = 1, 2$  справедливо равенство



$$\|\chi_1^{(1)}(x, \alpha)\|_{L^2(0,1)} = 1. \quad (16)$$

Действительно, для  $m = 1$  и  $k = 1$  получим

$$\begin{aligned} \chi_1^{(1)}(x, \alpha) &= \int_0^{\frac{1}{2^{1+\alpha}}} [\chi_1^{(1)}(x, \alpha)]^2 dx = \int_0^{\frac{1}{2^{1+\alpha}}} [\chi_1^{(1)}(x, \alpha)]^2 dx + \int_{\frac{1}{2^{1+\alpha}}}^1 [\chi_1^{(1)}(x, \alpha)]^2 dx = \\ &= \frac{2^{2(1+\alpha)}}{2} \cdot \frac{1}{2^{2(1+\alpha)}} + \frac{2^{2(1+\alpha)}}{2(2^{1+\alpha} - 1)} \cdot \left( \frac{1}{2^{1+\alpha}} - \frac{1}{2^{2(1+\alpha)}} \right) = \frac{1}{2} + \frac{2^{2(1+\alpha)}}{2(2^{1+\alpha} - 1)} \cdot \frac{2^{1+\alpha} - 1}{2^{2(1+\alpha)}} = \frac{1}{2} + \frac{1}{2} = 1. \end{aligned}$$

Тогда, для  $m = 1$  и  $k = 2$  будем иметь

$$\begin{aligned} \chi_1^{(2)}(x, \alpha) &= \int_{\frac{1}{2^{1+\alpha}}}^1 [\chi_1^{(2)}(x, \alpha)]^2 dx = \int_{\frac{1}{2^{1+\alpha}}}^{\frac{1+2^{1+\alpha}}{2^{2(1+\alpha)}}} [\chi_1^{(2)}(x, \alpha)]^2 dx + \int_{\frac{1+2^{1+\alpha}}{2^{2(1+\alpha)}}}^1 [\chi_1^{(2)}(x, \alpha)]^2 dx = \\ &= \frac{2^{2(1+\alpha)}}{2} \cdot \left( \frac{1+2^{1+\alpha}}{2^{2(1+\alpha)}} - \frac{1}{2^{1+\alpha}} \right) + \frac{2^{2(1+\alpha)}}{2[2^{1+\alpha}(2^{1+\alpha} - 1)]} \cdot \left( 1 - \frac{1+2^{1+\alpha}}{2^{2(1+\alpha)}} \right) = \\ &= \frac{2^{2(1+\alpha)}}{2} \cdot \frac{1}{2^{2(1+\alpha)}} + \frac{2^{2(1+\alpha)}}{2(2^{1+\alpha} - 1)} \cdot \frac{2^{1+\alpha} - 1}{2^{2(1+\alpha)}} = \frac{1}{2} + \frac{1}{2} = 1. \end{aligned}$$

Таким образом, (16) доказано. Теперь, докажем справедливость(15) для любого  $m \geq 1$ . В самом деле для промежутка

$$(0,1) = \bigcup_{k=1}^m \left( \frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{k^{1+\alpha}}{2^{m(1+\alpha)}} \right)$$

справедливо

$$\begin{aligned} \int_0^1 [\chi_m^k(x, \alpha)]^2 dx &= \int_{\frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}}^{\frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}} [\chi_m^k(x, \alpha)]^2 dx + \int_{\frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}}^{\frac{k^{1+\alpha}}{2^{m(1+\alpha)}}} [\chi_m^k(x, \alpha)]^2 dx = \\ &= \int_{\frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}}^{\frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}} \left[ \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2[(k-1)^{1+\alpha}(1-2^{1+\alpha}) + k^{1+\alpha}]}} \right]^2 dx + \\ &+ \int_{\frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}}^{\frac{k^{1+\alpha}}{2^{m(1+\alpha)}}} \left[ \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2[k^{1+\alpha}(2^{1+\alpha} - 1) - (k-1)^{1+\alpha}]}} \right]^2 dx = \\ &= \frac{(k-1)^{1+\alpha}(1-2^{1+\alpha}) + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}} \cdot \frac{2^{(m+1)(1+\alpha)}}{2[(k-1)^{1+\alpha}(1-2^{1+\alpha}) + k^{1+\alpha}]} + \\ &+ \frac{k^{1+\alpha}(2^{1+\alpha} - 1) - (k-1)^{1+\alpha}}{2^{(m+1)(1+\alpha)}} \cdot \frac{2^{(m+1)(1+\alpha)}}{2[k^{1+\alpha}(2^{1+\alpha} - 1) - (k-1)^{1+\alpha}]} = \frac{1}{2} + \frac{1}{2} = 1, \end{aligned}$$

тем самым, равенство (15) доказано.

Докажем ортогональность введенной системы для  $m = 1, k = 1, 2$ . Для этого достаточно показать справедливость равенства

$$\int_0^1 \chi_1^1(x, \alpha) \cdot \chi_1^2(x, \alpha) dx = 0. \quad (17)$$

В самом деле, по определению функций, произведение подынтегральных функций равно 0, тогда и интеграл равен 0, т.е.

$$\begin{aligned} \chi_1^1(x, \alpha) \cdot \chi_1^2(x, \alpha) &= \left[ \sqrt{\frac{2^{2(1+\alpha)}}{2}} \cdot 0 + \left( -\sqrt{\frac{2^{2(1+\alpha)}}{2(2^{1+\alpha}-1)}} \right) \cdot 0 \right] + \\ &+ \left[ 0 \cdot \sqrt{\frac{2^{2(1+\alpha)}}{2}} + 0 \cdot \left( -\sqrt{\frac{2^{2(1+\alpha)}}{2[2^{1+\alpha}(2^{1+\alpha}-1)]}} \right) \right] = 0, \end{aligned}$$

отсюда,

$$\int_0^1 \chi_1^1(x, \alpha) \cdot \chi_1^2(x, \alpha) dx = \int_0^1 0 dx = 0.$$

Для случаев 1. и 2. при  $m=1, k=1$  теорема доказана. Рассмотрим следующий случай.

3. Пусть даны числа  $m > 1, k = 1, 2, \dots, 2^m$  и  $n > m$ . Сначала, рассмотрим случай  $n = m + 1$ . По лемме 1, т.е. в случае

$$2k - 1 \leq k' \leq k \left[ \left( 1 - \frac{1}{k} \right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}},$$

получим

$$\left( \frac{(k-1)^{1+\alpha}}{2^{n(1+\alpha)}}, \frac{(k')^{1+\alpha}}{2^{n(1+\alpha)}} \right) \subset \left( \frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}, \frac{k^{1+\alpha} + k'^{1+\alpha}}{2^{(m+1)(1+\alpha)}} \right). \quad (18)$$

Отсюда

$$\begin{aligned} &\chi_m^k(x, \alpha) \cdot \chi_{m+1}^{k'}(x, \alpha) = \\ &= \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2[(k-1)^{1+\alpha}(1-2^{1+\alpha}) + k^{1+\alpha}]}} \cdot \sqrt{\frac{2^{(m+2)(1+\alpha)}}{2[(k'-1)^{1+\alpha}(1-2^{1+\alpha}) + (k')^{1+\alpha}]}} + \\ &+ \frac{2^{(m+1)(1+\alpha)}}{2[(k-1)^{1+\alpha}(1-2^{1+\alpha}) + k^{1+\alpha}]} \cdot \left( -\sqrt{\frac{2^{(m+2)(1+\alpha)}}{2[k^{1+\alpha}(2^{1+\alpha}-1) - (k-1)^{1+\alpha}]} \right). \end{aligned}$$

Из последнего равенства, будем иметь

$$\begin{aligned} &\int_0^1 \left[ \chi_m^k(x, \alpha) \cdot \chi_{m+1}^{k'}(x, \alpha) \right] dx = \\ &= \int_{\frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}}^{\frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}} \left[ \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2[(k-1)^{1+\alpha}(1-2^{1+\alpha}) + k^{1+\alpha}]}} \cdot \sqrt{\frac{2^{(m+2)(1+\alpha)}}{2[(k'-1)^{1+\alpha}(1-2^{1+\alpha}) + (k')^{1+\alpha}]}} \right] dx + \\ &+ \int_{\frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}}^{\frac{k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}} \left[ \frac{2^{(m+1)(1+\alpha)}}{2[(k-1)^{1+\alpha}(1-2^{1+\alpha}) + k^{1+\alpha}]} \cdot \left( -\sqrt{\frac{2^{(m+2)(1+\alpha)}}{2[k^{1+\alpha}(2^{1+\alpha}-1) - (k-1)^{1+\alpha}]} \right) \right] dx = \end{aligned}$$

$$\begin{aligned}
&= \left[ \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2 \left[ (k-1)^{1+\alpha} (1-2^{1+\alpha}) + k^{1+\alpha} \right]}} \cdot \sqrt{\frac{2^{(m+2)(1+\alpha)}}{2 \left[ (k'-1)^{1+\alpha} (1-2^{1+\alpha}) + (k')^{1+\alpha} \right]}} \right] \cdot \left[ \frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}} - \right. \\
&\left. - \frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}} \right] + \left[ \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2 \left[ (k-1)^{1+\alpha} (1-2^{1+\alpha}) + k^{1+\alpha} \right]}} \cdot \sqrt{\frac{2^{(m+2)(1+\alpha)}}{2 \left[ (k'-1)^{1+\alpha} (1-2^{1+\alpha}) + (k')^{1+\alpha} \right]}} \right] \times \\
&\times \left[ \frac{k^{1+\alpha}}{2^{m(1+\alpha)}} - \frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}} \right] = 0.
\end{aligned}$$

Снова применяя лемму 1, при  $1 + k \left[ \left(1 - \frac{1}{k}\right)^{1+\alpha} + 1 \right]^{\frac{1}{1+\alpha}} \leq k' \leq 2k$ , получим

$$\left( \frac{(k'-1)^{1+\alpha}}{2^{n(1+\alpha)}}, \frac{(k')^{1+\alpha}}{2^{n(1+\alpha)}} \right) \subset \left( \frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}, \frac{k^{1+\alpha}}{2^{m(1+\alpha)}} \right). \quad (19)$$

Тогда  $\chi_m^k(x, \alpha) \cdot \chi_{m+1}^{k'}(x, \alpha) =$

$$\begin{aligned}
&= \left( -\sqrt{\frac{2^{(m+1)(1+\alpha)}}{2 \left[ k^{1+\alpha} (2^{1+\alpha} - 1) - (k-1)^{1+\alpha} \right]}} \right) \cdot \sqrt{\frac{2^{(m+2)(1+\alpha)}}{2 \left[ (k'-1)^{1+\alpha} (1-2^{1+\alpha}) + (k')^{1+\alpha} \right]}} + \\
&+ \left( -\sqrt{\frac{2^{(m+1)(1+\alpha)}}{2 \left[ k^{1+\alpha} (2^{1+\alpha} - 1) - (k-1)^{1+\alpha} \right]}} \right) \cdot \left( -\sqrt{\frac{2^{(m+2)(1+\alpha)}}{2 \left[ (k')^{1+\alpha} (2^{1+\alpha} - 1) - (k'-1)^{1+\alpha} \right]}} \right).
\end{aligned}$$

Отсюда  $\int_0^1 \left[ \chi_m^k(x, \alpha) \cdot \chi_{m+1}^{k'}(x, \alpha) \right] dx = \int_0^1 0 dx = 0$ . Если  $n > m+1$ ,  $k=1$  и  $k'=1$ , то

$$\left( 0, \frac{1}{2^{n(1+\alpha)}} \right) \subset \left( 0, \frac{1}{2^{(m+1)(1+\alpha)}} \right), \quad \text{А в случае } k'=2, \quad \text{то} \\
\left( \frac{1}{2^{n(1+\alpha)}}, \frac{1}{2^{(n-1)(1+\alpha)}} \right) \subset \left( \frac{1}{2^{(m+1)(1+\alpha)}}, \frac{1}{2^{m(1+\alpha)}} \right).$$

Таким образом, при  $k=1$ ,  $k'=1$  и  $k'=2$ , а также при  $k' > 2$ , имеем

$$\int_0^1 \chi_m^1(x, \alpha) \cdot \chi_n^{k'}(x, \alpha) dx = 0.$$

Теперь, пусть  $n > m+1$ ,  $k > 1$ . Тогда по лемме 2, можно провести аналогичные рассуждения.

Теорема 1 доказана полностью. На промежутке  $(0;1]$ , рассмотрим разложение функции  $f(x) \in L(0,1)$  по ортонормированной системе Хаара. Пусть дана функция  $f(x) \in L(0,1)$ . Разложения функции  $f(x) \in L(0,1)$  по ортонормированной системе Хаара будет иметь вид

$$f(x) = \sum_{m=1}^{\infty} \sum_{k=1}^m c_k^m \chi_m^{(k)}. \quad (20)$$

Определим, для некоторой функции коэффициенты Фурье данного разложения.

**Заключение.** Разложение функции  $f(x) = x^\beta$  по ортонормированной системе Хаара

Теорема 2. Пусть даны числа  $-1 < \alpha \leq 0$ ,  $\beta > -1$  и  $m \geq 1$  - целое число. Тогда для функции  $f(x) = x^\beta$ , определенной на  $(0;1]$  имеет место:

$$f(x) = \frac{1}{2^{(1+\alpha)(1+\beta)}(\beta+1)} \sum_{m=1}^{\infty} \sum_{k=1}^m [c_m^k(f, \beta, \alpha)_1 + c_m^k(f, \beta, \alpha)_2] \chi_m^{(k)}, \quad (21)$$

здесь,

$$c_m^k(f, \beta, \alpha)_1 = \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2[(k-1)^{1+\alpha}(1-2^{1+\alpha}) + k^{1+\alpha}]}} \cdot \frac{((k-1)^{1+\alpha} + k^{1+\alpha})^{\beta+1} - [2(k-1)]^{(1+\alpha)(\beta+1)}}{2^{m(1+\alpha)(\beta+1)}},$$

$$c_m^k(f, \beta, \alpha)_2 = \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2[k^{1+\alpha}(2^{1+\alpha}-1) - (k-1)^{1+\alpha}]}} \cdot \frac{(2k)^{(1+\alpha)(\beta+1)} - ((k-1)^{1+\alpha} + k^{1+\alpha})^{(\beta+1)}}{2^{m(\beta+1)(1+\alpha)}}.$$

Доказательство. пусть  $m = 1$ ,  $k = 1, 2$ . Тогда справедливо

$$\begin{aligned} c_1^1(f, \beta, \alpha) &= \int_0^{\frac{1}{2^{2(1+\alpha)}}} x^\beta \cdot \frac{2^{1+\alpha}}{\sqrt{2}} dx = \\ &= \frac{2^{1+\alpha}}{\sqrt{2}} \cdot \frac{1}{\beta+1} \cdot \left( \frac{1}{2^{2(1+\alpha)(1+\beta)}} \right) = \frac{2^{1+\alpha}}{\sqrt{2}(\beta+1)2^{2(1+\alpha)(1+\beta)}} \end{aligned}$$

$$\begin{aligned} c_1^2(f, \beta, \alpha) &= - \int_{\frac{1}{2^{2(1+\alpha)}}}^{\frac{1}{2^{1+\alpha}}} x^\beta \cdot \frac{2^{1+\alpha}}{\sqrt{2(2^{1+\alpha}-1)}} dx = \\ &= - \frac{2^{1+\alpha}}{\sqrt{2(2^{1+\alpha}-1)}} \cdot \frac{1}{\beta+1} \cdot \left( \frac{1}{2^{(1+\alpha)(1+\beta)}} - \frac{1}{2^{2(1+\alpha)(1+\beta)}} \right) = \\ &= - \frac{2^{1+\alpha}(2^{(1+\alpha)(1+\beta)} - 1)}{\sqrt{2(2^{(1+\alpha)(1+\beta)} - 1)}(\beta+1)2^{2(1+\alpha)(1+\beta)}}. \end{aligned}$$

Тогда, для  $m = 1$ ,  $k = 1, 2$

$$c_1^1(f, \beta, \alpha) = \frac{2^{1+\alpha}}{\sqrt{2}(\beta+1)2^{2(1+\alpha)(1+\beta)}}$$

и

$$c_1^2(f, \beta, \alpha) = - \frac{2^{1+\alpha}(2^{(1+\alpha)(1+\beta)} - 1)}{\sqrt{2(2^{(1+\alpha)(1+\beta)} - 1)}(\beta+1)2^{2(1+\alpha)(1+\beta)}}.$$

Пусть, теперь  $m > 1$ ,  $k = 1, 2, \dots, 2^m$ . Тогда

$$c_m^k(f, \beta, \alpha)_1 = \int_{\frac{(k-1)^{1+\alpha}}{2^{m(1+\alpha)}}}^{\frac{(k-1)^{1+\alpha} + k^{1+\alpha}}{2^{(m+1)(1+\alpha)}}} x^\beta \cdot \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2[(k-1)^{1+\alpha}(1-2^{1+\alpha}) + k^{1+\alpha}]}} dx =$$

$$\begin{aligned}
&= \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2\left[(k-1)^{1+\alpha}(1-2^{1+\alpha})+k^{1+\alpha}\right]}} \cdot \frac{1}{\beta+1} \cdot \left[ \frac{\left((k-1)^{1+\alpha}+k^{1+\alpha}\right)^{\beta+1}}{2^{(m+1)(1+\alpha)(\beta+1)}} - \frac{(k-1)^{(1+\alpha)(\beta+1)}}{2^{m(1+\alpha)(\beta+1)}} \right] \\
c_m^k(f, \beta, \alpha)_2 &= - \int_{\frac{(k-1)^{1+\alpha}+k^{1+\alpha}}{2^{m(1+\alpha)}}}^{\frac{k^{1+\alpha}}{2^{m(1+\alpha)}}} \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2\left[k^{1+\alpha}(2^{1+\alpha}-1)-(k-1)^{1+\alpha}\right]}} \cdot x^\beta dx = \\
&= - \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2\left[k^{1+\alpha}(2^{1+\alpha}-1)-(k-1)^{1+\alpha}\right]}} \cdot \frac{1}{\beta+1} \cdot \left[ \frac{k^{(1+\alpha)(\beta+1)}}{2^{m(\beta+1)(1+\alpha)}} - \frac{\left((k-1)^{1+\alpha}+k^{1+\alpha}\right)^{(\beta+1)}}{2^{(m+1)(1+\alpha)(\beta+1)}} \right].
\end{aligned}$$

Таким образом, при  $m > 1$ ,  $k = 1, 2, \dots, 2^m$

$$\begin{aligned}
c_m^k(f, \beta, \alpha)_1 &= \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2\left[(k-1)^{1+\alpha}(1-2^{1+\alpha})+k^{1+\alpha}\right]}} \times \\
&\times \frac{1}{\beta+1} \cdot \left[ \frac{\left((k-1)^{1+\alpha}+k^{1+\alpha}\right)^{\beta+1}}{2^{(m+1)(1+\alpha)(\beta+1)}} - \frac{(k-1)^{(1+\alpha)(\beta+1)}}{2^{m(1+\alpha)(\beta+1)}} \right]
\end{aligned}$$

и

$$\begin{aligned}
c_m^k(f, \beta, \alpha)_2 &= - \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2\left[k^{1+\alpha}(2^{1+\alpha}-1)-(k-1)^{1+\alpha}\right]}} \times \\
&\times \frac{1}{\beta+1} \cdot \left[ \frac{k^{(1+\alpha)(\beta+1)}}{2^{m(\beta+1)(1+\alpha)}} - \frac{\left((k-1)^{1+\alpha}+k^{1+\alpha}\right)^{(\beta+1)}}{2^{(m+1)(1+\alpha)(\beta+1)}} \right].
\end{aligned}$$

Отсюда, коэффициенты разложения

$$c(\alpha, \beta) = \frac{1}{2^{(1+\alpha)(1+\beta)}(\beta+1)}$$

определяются в виде

$$\begin{aligned}
c_m^k(f, \beta, \alpha) &= c_m^k(f, \beta, \alpha)_1 + c_m^k(f, \beta, \alpha)_2 = \\
&= \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2\left[(k-1)^{1+\alpha}(1-2^{1+\alpha})+k^{1+\alpha}\right]}} \cdot \frac{\left((k-1)^{1+\alpha}+k^{1+\alpha}\right)^{\beta+1} - \left[2(k-1)\right]^{(1+\alpha)(\beta+1)}}{2^{m(1+\alpha)(\beta+1)}} - \\
&- \sqrt{\frac{2^{(m+1)(1+\alpha)}}{2\left[k^{1+\alpha}(2^{1+\alpha}-1)-(k-1)^{1+\alpha}\right]}} \cdot \frac{(2k)^{(1+\alpha)(\beta+1)} - \left((k-1)^{1+\alpha}+k^{1+\alpha}\right)^{(\beta+1)}}{2^{m(\beta+1)(1+\alpha)}}.
\end{aligned}$$

Теорема 2 доказана полностью.

## Литература

1. Жәутіков О.А. Математикалық анализ курсы. Алматы, «Ғылым», 1958 ж.

2. Темірғалиев Н. Математикалық анализ, I т., Алматы, «Мектеп», 1987.
3. Темірғалиев Н. Математикалық анализ, II т., Алматы, «Анатілі», 1991.
4. Темірғалиев Н. Математикалық анализ, II т., Алматы, «Білім», 1997.
5. Кашин Б.С., Саакян А.А. Ортогональные ряды. М.: «АФЦ», 1999.
6. Никольский С.М. Приближение функций многих переменных и теоремы вложения, М., «Наука», 1977.
7. Бари Н.К. Тригонометрические ряды, М.: «Физматгиз», 1961.
8. Дзядык В.К. Введение в теорию равномерного приближения функций полиномами. М.: «Наука», 1977.
9. Төлегенов Б.Т. Математикалықанализденлекциялар курсы 1. Алматы, 1973.
10. Төлегенов Б.Т. Математикалықанализденлекциялар курсы 2. Алматы, 1973.

## РАЗРАБОТКА СКРИПТА ДЛЯ ВЫЯВЛЕНИЯ ВАРИАБЕЛЬНЫХ ЛОКУСОВ БАКТЕРИАЛЬНЫХ ШТАММОВ *FRANCISELLA TULARENSIS*

В.А. Шевцов, А.Д. Кауржанова, А.А. Исмаилова  
Казахский агротехнический университет им. С. Сейфуллина, Астана, Казахстан,  
shevtsovvladislav111@gmail.com

### Абстракт

Большой массив данных получаемый в настоящее время с использованием секвенаторов второго и третьего поколения требует привлечения специалистов из области информационных технологий для обеспечения процесса обработки данных и их интерпретации, а также разработки программного обеспечения. В настоящее время существует переходный период догеномной и постгеномной эры, в связи с этим исследователи сталкиваются с проблемой совмещения современных данных получаемых с использования прямого секвенирования и ранее получаемыми результатами методами капелярного электрофареза при анализе тандемных повторов. Целью нашей работы была разработка скрипта, позволяющего выявлять тандемные повторы в массивах геномных данных с последующим сопоставлением результатов с классическими видами ПЦР. Предлагаемый нами скрипт, позволил безошибочно установить размеры 7 используемых VNTR локусов при генотипировании *Francisella tularensis*, при этом было установлено полное соответствие с капелярным электрофарезом и выявлены неточности при использовании в *in silico* MLVA с использованием сборок полученных с использованием таких программ сборщиков как SKESA и SPAdes.

### Введение

Совмещение данных генотипирования микроорганизмов полученных в «догеномный» период с новыми геномными данными стало настоящей проблемой. Данная пропасть возникла в связи с различиями протоколов изучения ДНК микроорганизмов на протяжении развития технологий генотипирования. Генотипирование на основании фрагментации ДНК с разделением фрагментов в pulsed-field gel electrophoresis (PFGE), Random Amplified Polymorphic DNA (RAPD), Restriction Fragment Length Polymorphism (RFLP) сменились на более простые, точные и легко воспроизводимые методы как мультилокусный анализ переменных тандемных повторов (MLVA) и мультилокусное секвенс типирование (MLST). В настоящее время существует переходный период, когда для полноценной эпидемиологии необходимо использовать данные, полученные по всему миру различными методами генотипирования, а также при ретроспективном анализе ранее полученных результатов генотипирования циркулирующих штаммов. В переходный период многие лаборатории вынуждены параллельно использовать полногеномное секвенирование, PFGE and MLVA для ряда патогенов, чтобы проводить полноценный эпидемиологический контроль.

Определение нуклеотидной последовательности в современных биологических исследованиях является основной технологией. Всего несколько десятилетий назад стоимость секвенирования генома человека составляла 3 миллиарда [1] долларов и заняло 15 лет исследований, однако на сегодняшний день существуют так называемые секвенаторы нового поколения, благодаря которым цена за секвенирование существенно снизилась и составляет приблизительно 1000 долларов за секвенс генома человека [2].

С появлением секвенаторов нового поколения возрос и уровень сложности исследований. Так, уже известны такие проекты как «Проект 100 000 геномов», целью которого является создание наборов данных по геномике, транскриптомике и эпигеномике для 100 000 пациентов с раком или редкими заболеваниями [3], «100K Foodborne Pathogen Genome Project» направлен на секвенирование геномов сотен тысяч болезнетворных бактериальных и вирусных изолятов [4] или секвенирование 1000 геномов грибов [5]

Эксперты в данной области прогнозируют, что количество генетических данных уже к 2025 году будут сопоставимы с большими данными астрономии, YouTube и Twitter [6].

В связи с этим развитие методов *in silico* генотипирования с использованием полногеномных данных является актуальной задачей, особенно для патогенов, для которых MLVA и MLST рассматривался в качестве конвенции генотипирования включая такие патогены как: *Francisella tularensis*, *Brucella spp*, *Bacillus anthracis*, *Yersenia pestis*.

Долгое время для определения использовалось классическое генотипирование на основании тандемных повторов с использованием ПЦР и капелярного электрофареза, однако в настоящее время для таких опасных патогенов проводится полногеномное секвенирование, в результате чего появилась необходимость сопоставления результатов старых данных с новыми. К сожалению прямая *de novo* сборка может содержать ошибки, поскольку программа сборщик может не включить риды в сборку, поэтому существует необходимость в разработке программы которая может работать не только со сборками, а непосредственно с сырыми данными, что значительно снижает риск интерпретировать результат неверно. В связи с переходом от классической типизации к подходам, основанным на WGS, недавно были разработаны скрипты Python для быстрого извлечения *in silico* аллелей *Brucella* MLVA. Это позволит назначать типы MLVA в геномную эру, избегая нежелательной потери генетической информации, которая предоставлялась в течение более 10 лет с использованием золотого стандарта типирования MLVA на основе ПЦР. Это также важно, поскольку сроки технологического перехода будут значительно различаться в разных странах. Поскольку до сих пор не проводилось экспериментальной оценки одного из этих сценариев, наша главная цель состояла в том, чтобы получить доступ к его производительности по методу золотого стандарта.

#### **Материалы и методы**

В качестве материалов были использованы геномные данные 39 штаммов *F. tularensis subsp. holarctica*, секвенированных с использованием MiSeq(Illumina) с длиной ридов 300 пар нуклеотидов.

*in silico* MLVA запускали на сборках генома, применяя скрипт MLVA\_finder.py [7] с настройками по умолчанию и праймерами, перечисленными.

Поскольку размер полиморфных массивов VNTR меньше, чем полученные здесь риды размером 300 п.н., также были использованы сырые данные ридов. Был разработан скрипт на основе инструмента bbduk [8] для вычленения ридов, которые содержат VNTR и обе фланкирующие последовательности [9]. Были подобраны праймеры, расположенные как можно ближе к целевым локусам VNTR. При выборе праймеров для *in silico* вместо обнаружения обращали внимание только на уникальность и на формирование только целевого фрагмента. Скрипт объединяет все риды, содержащие интересующий локус VNTR, в один файл в формате .fasta, после чего файл был готов в качестве входных данных для скрипта MLVA\_finder.py.

Скрипт был написан с использованием Python в качестве языка программирования. Софт intelliJ idea применялся в качестве среды разработки.

#### **Результаты и дискуссия**

Применение MLVA является важным инструментом для установления путей распространения особо опасных штаммов в окружающей среде. Существующие протокола подтипирования для различных патогенов, облегчили описание моделей распространения среди коллекции. Выбор переменных, повторяющихся локусов является важной составляющей эпидемиологического контроля, поскольку эти локусы позволяют разделять изоляты на кластеры, тем самым облегчая задачи по контролю и выявлению цепочки распространения вспышек заболевания.

Чтобы оценить совместимость между генотипированием MLVA *in silico* и *in vitro*, мы использовали два разных подхода к выводу генотипа MLVA из данных последовательности. В первом подходе чтения секвенирования были собраны при помощи програм сборщиков. Сравнивались два ассемблера SPAdes и SKESA. Второй метод



заклучался в вычленении необработанных чтений секвенирования, соответствующие локусам VNTR, и непосредственно проанализированы для получения *in silico* генотипа MLVA. В данном подходе используются чтения длиной 300 пар оснований.

Результаты, полученные с помощью анализа прямого считывания, были идентичны данным *in vitro* (Таблица 1). Ft-3 не собирался в одиннадцати и пяти штаммах со SPADES и SKESA соответственно (Таблица 1). Более того, сборки Ft-3 SPADES показали неправильный размер в нескольких образцах. Сборки SKESA предсказали неправильный размер *insilico*-FT-4 во всех штаммах, за исключением одного.

Таблица 1 – Анализ представления вариабельных локусов с использованием ридов и сборок

Анализ прямых считываний	Локусы		SPADES сборка	Локусы		SKESA сборка	Локусы	
	Ft-3_9bp_248bp_25u	<i>insilico</i> -FT-4_6bp_97bp_2u		Ft-3_9bp_248bp_25u	<i>insilico</i> -FT-4_6bp_97bp_2u		Ft-3_9bp_248bp_25u	<i>insilico</i> -FT-4_6bp_97bp_2u
Наименование			Наименование			Наименование		
Tul-106_KZ	10	4	Tul-106_KZ	10	4	Tul-106_KZ	10	3*
Tul-11_KZ	9	4	Tul-11_KZ	9	4	Tul-11_KZ	9	3*
Tul-112_KZ	15	4	Tul-112_KZ	15	4	Tul-112_KZ	15	3*
Tul-128_KZ	21	4	Tul-128_KZ	19*	4	Tul-128_KZ		3*
Tul-13_KZ	20	4	Tul-13_KZ	18*	4	Tul-13_KZ	20	3*
Tul-132_KZ	23	4	Tul-132_KZ		4	Tul-132_KZ		3*
Tul-133_KZ	21	4	Tul-133_KZ		4	Tul-133_KZ		3*
Tul-135_KZ	22	4	Tul-135_KZ		4	Tul-135_KZ	22	3*
Tul-139_KZ	20	4	Tul-139_KZ		4	Tul-139_KZ	20	3*
Tul-149_KZ	19	4	Tul-149_KZ	18*	4	Tul-149_KZ		3*
Tul-15_KZ	23	4	Tul-15_KZ		4	Tul-15_KZ	23	3*
Tul-151_KZ	21	4	Tul-151_KZ		4	Tul-151_KZ	21	3*
Tul-153_KZ	9	4	Tul-153_KZ	9	4	Tul-153_KZ	9	3*
Tul-154_KZ	16	4	Tul-154_KZ	16	4	Tul-154_KZ	16	3*
Tul-155_KZ	20	4	Tul-155_KZ	17*	4	Tul-155_KZ	20	3*
Tul-161_KZ	16	4	Tul-161_KZ	16	4	Tul-161_KZ	16	3*
Tul-17_KZ	20	4	Tul-17_KZ		4	Tul-17_KZ		3*
Tul-18_KZ	20	4	Tul-18_KZ		4	Tul-18_KZ	20	3*
Tul-19_KZ	17	4	Tul-19_KZ	17	4	Tul-19_KZ	17	3*
Tul-2_KZ	10	4	Tul-2_KZ	10	4	Tul-2_KZ	10	3*
Tul-20_KZ	10	3	Tul-20_KZ	10	3	Tul-20_KZ	10	3
Tul-30_KZ	12	4	Tul-30_KZ	12	4	Tul-30_KZ	12	3*
Tul-52_KZ	23	4	Tul-52_KZ		4	Tul-52_KZ	23	3*

Tul-6_KZ	10	4	Tul-6_KZ	10	4	Tul-6_KZ	10	3*
Tul-61_KZ	9	4	Tul-61_KZ	9	4	Tul-61_KZ	9	3*
Tul-66_KZ	14	4	Tul-66_KZ	14	4	Tul-66_KZ	14	3*
Tul-67_KZ	16	4	Tul-67_KZ	16	4	Tul-67_KZ	16	3*
Tul-68_KZ	14	4	Tul-68_KZ	14	4	Tul-68_KZ	14	3*
Tul-7_KZ	21	4	Tul-7_KZ		4	Tul-7_KZ	21	3*
Tul-71_KZ	17	4	Tul-71_KZ	17	4	Tul-71_KZ	17	3*
Tul-76_KZ	9	4	Tul-76_KZ	9	4	Tul-76_KZ	9	3*
Tul-78_KZ	9	4	Tul-78_KZ	9	4	Tul-78_KZ	9	3*
Tul-86_KZ	10	4	Tul-86_KZ	10	4	Tul-86_KZ	10	3*
Tul-87_KZ	10	4	Tul-87_KZ	10	4	Tul-87_KZ	10	3*
Tul-92_KZ	16	4	Tul-92_KZ	16	4	Tul-92_KZ	16	3*
Tul-93_KZ	21	4	Tul-93_KZ	19*	4	Tul-93_KZ	21	3*
Tul-97_KZ	16	4	Tul-97_KZ	16	4	Tul-97_KZ	16	3*
Tul-98_KZ	10	4	Tul-98_KZ	10	4	Tul-98_KZ	10	3*
Tul-99_KZ	21	4	Tul-99_KZ		4	Tul-99_KZ	21	3*

Таким образом, прямой анализ с использованием ридов 300 п.н. в качестве входных данных, показал значительно лучшие результаты, чем анализ сборок. Данный скрипт генотипирования *in silico* MLVA, основанный на прямом анализе сырых прочтений, позволил правильно идентифицировать все аллели в 39 штаммах.

#### Заключение

Был разработан скрипт позволяющий проводить MLVA анализ с использованием сырых данных (риды полученные с секвентора нового поколения. В качестве апробации было использовано 39 геномов *Francisella tularensis*. Преимущество нашего метода является возможность использования сырых данных, таким образом значительно снижается вероятность ошибок так как при использовании сборок, программы сборщики образуют консенсус в которой как правило образуются разрывы и последовательность с тандемным повтором может быть исключена из сборки или засчет нарушения размера k-мера или за счет гомоплазии данного региона.

#### Литературный обзор

1. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., ... Holt, R. A. (2001). The Sequence of the Human Genome. *Science*, 291(5507), 1304–1351. doi:10.1126/science.1058040
2. Check Hayden, E. Technology: The \$1,000 genome. *Nature* 507, 294–295 (2014). <https://doi.org/10.1038/507294a>
3. Wheway, G., Genomics England Research Consortium, & Mitchison, H. M. (2019). Opportunities and challenges for molecular understanding of ciliopathies—the 100,000 Genomes Project. *Frontiers in genetics*, 10, 127.
4. Weimer, B. C. (2017). 100K pathogen genome project. *Genome announcements*, 5(28), e00594-17.
5. Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otilar, R., ... & Shabalov, I. (2014). MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic acids research*, 42(D1), D699-D704.
6. Stephens, Z. D., Lee, S. Y., Faghri, F., Campbell, R. H., Zhai, C., Efron, M. J., ... & Robinson, G. E. (2015). Big data: astronomical or genetical?. *PLoS biology*, 13(7), e1002195.
7. Скрипт MLVA\_finder.py: [https://github.com/i2bc/MLVA\\_finder](https://github.com/i2bc/MLVA_finder) (Дата обращения: 2022, май 20)

8. BBTools : <https://sourceforge.net/projects/bbmap/> (Дата обращения: 2022, сентябрь 14)
9. <https://github.com/Vladislav-Shevtsov/search-primers-in-reads/> (Дата обращения: 2022, апрель 24)
10. Среда разработки intellij idea: <https://www.jetbrains.com/idea/> (Дата обращения: 2022, февраль 21)

## КӨПӨЛШЕМДІ МОДЕЛЬДЕРДІ КЛАССИФИКАЦИЯЛАУ, СИПАТТАУ ЖӘНЕ САЛЫСТЫРУ ОРТАСЫ

*Муратова Жансая Муратовна, аспирант*

*Российский университет дружбы народов, г. Москва, [jansik.86@mail.ru](mailto:jansik.86@mail.ru)*

*Жетписов Кабылда, ф.-м.ғ.к., доцент*

*Евразийский национальный университет имени Л. Н. Гумилёва, г. Астана,*

*[jetpisov\\_K54@mail.ru](mailto:jetpisov_K54@mail.ru)*

*Иксебаева Жанна Сагинтаевна, аға оқытушы*

*Западно-Казахстанский университет имени М. Утемисова, г. Уральск, [iksebaeva@mail.ru](mailto:iksebaeva@mail.ru)*

**Аннотация.** Көпөлшемді модельдерді классификациялаудың әртүрлі тәсілдері бар. Мұндай классификациялау туралы жұмыстардың көпшілігінде көпөлшемді модельге қойылатын арнайы талаптар тізімі бастапқыда бар модельдердің осы талаптарға сәйкестігін зерттеу мақсатында анықталады. Бұл мақалада айтылып отырған классификациялар модельдердің әлсіз жақтарын анықтауға бағытталған емес, оның мақсаты көпөлшемді модельдеудің әртүрлі тәсілдерін салыстыру болып табылады. Бұл мақалада қарастырылатын көп өлшемді модельдерді сипаттау үшін классификациялау белгілерінің екі жиынтығы енгізілген.

Модельдер пәндік аймақты сипаттау кезеңіне сәйкес қарастырылады. Осылайша, концептуалды, логикалық және физикалық модельдерді ажыратуға болады. Модельдердің осы үш түрінен басқа жобалау кезеңдерінің ешқайсысына тікелей қатысы жоқ, бірақ пәндік саланы теориялық сипаттауға арналған модельдер де қарастырылады. Мұндай модельдер математикалық модельдер деп аталады.

Модельдер мәліметтерді бейнелеудің детализациялау деңгейіне сәйкес: жоғарғы, орта және төменгі деңгейлерінде қарастырылады. Осы деңгейлердің әрқайсысы үшін модельде домендік концепциялардың қайсысы ұсынылғаны анықталады. Детализациялау деңгейлері әртүрлі конструкцияларды, ақпараттың көп өлшемді көрінісін зерттеу үшін қолданылады.

Деректер модельдерін жобалаудың әртүрлі кезеңдерінде пәндік аймақты сипаттау үшін пайдаланылатын ұғымдар жиынтығы ретінде қарастыруға болады: бизнес-аналитиктерге арналған жоғары деңгейлі сипаттамадан бастап, ақпаратты сақтау құрылымдарын анықтайтын дерекқор схемасын құруға дейін. Деректер модельдері адамның ақпаратты қалай қабылдайтынына жақын ұғымдарды білдіретін болса, «Жоғары деңгей» немесе «Концептуалды» деп аталады; «Төмен деңгей» немесе «Физикалық» ретінде болады, егер олар деректердің компьютерде сақталуын сипаттайтын ұғымдарды білдірсе; және «Орындау» деп аталады, егер олар соңғы пайдаланушыға түсінікті болатын, бірақ сонымен бірге деректердің сақтауда қалай ұйымдастырылғанынан тым алыс емес ұғымдарды пайдаланса.

Сол сияқты бұл мақалада пәндік аймақты сипаттау кезеңдеріне сәйкес модельдердің үш тобын ажыратамыз.

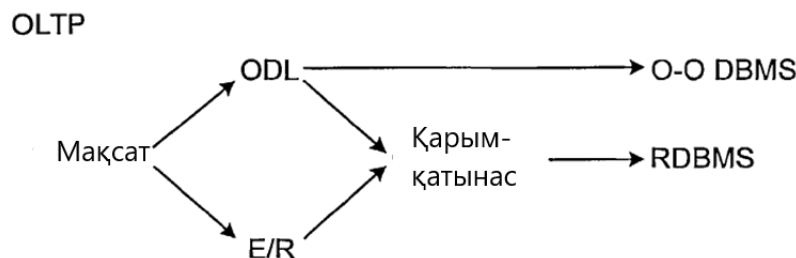
- Концептуалды модельдер соңғы пайдаланушыға ең жақын және ең түсінікті, олар адам ойлауына жақын ұғымдармен жұмыс істейді. Бұл модельдер деректердің қалай сақталатындығынан тәуелсіз.

- Логикалық модельдер әлі де соңғы пайдаланушыға түсінікті, бірақ олар қазірдің өзінде деректерді сақтау тәсіліне, атап айтқанда ДҚБЖ таңдауына байланысты.

- Физикалық модельдер деректердің ДҚБЖ-да сақталу тәсіліне толығымен тәуелді. Олардың бейнеленуі тек нақты және пайдаланушы деректерінің көрсетіліміне еш қатысы жоқ. Бұл модельдер деректердің компьютерде қалай сақталатынын сипаттауға арналған.

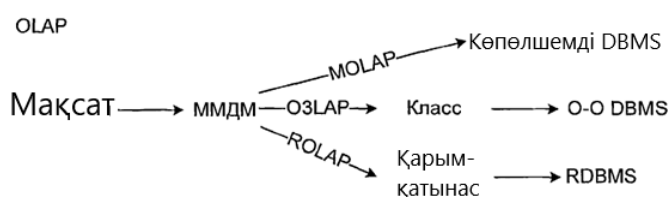
Мысалы, Online Transaction Processing (OLTP) ортасында модельдеу процесін қарастыруға болады (1-сурет). Концептуалды деңгейінде модельдеудің бірінші қадамы пайдаланушының мақсатын көрсету үшін ODL (объектіні анықтау тілі) немесе E/R (нысан

қатынасы) сипаттамасын пайдаланады; логикалық деңгейдегі келесі қадамда әдетте реляциялық модель пайдаланылады, бірақ иерархиялық немесе желілік модель (суретте көрсетілмеген) де пайдаланылуы мүмкін; және соңғы қадамда физикалық деңгейде іске асыру нақты ДҚБЖ (мысалы, Oracle, Informix, ObjectStore және т.б.) тәуелді болады.



Сурет 1. OLTP ортасында модельдеу процесі.

Сол сияқты, OLAP (On-Line Analytical Processing) ортасын пайдаланған жағдайда (2-сурет) көпөлшемді деректер моделі концептуалды деңгейде, содан кейін тәсілге байланысты (реляциялық - ROLAP, нысанға-бағдарланған - O3LAP немесе жай ғана көпөлшемді - MOLAP) логикалық деңгейде әртүрлі модельдерді және физикалық деңгейде іске асыру үшін әртүрлі ДҚБЖ қолданады.



Сурет 2. OLAP ортасында модельдеу процесі.

Жоғарыда аталған модельдердің үш түрінен басқа пәндік аймақтың жобалау кезеңдерінің ешқайсысына жатпайтын, бірақ оның түсініктерін формализациялауға арналған көптеген модельдер бар. Мұндай модельдер математикалық модельдер болып табылады, олар пәндік саланы сипаттау үшін математикалық ұғымдарды, формулаларды және анықтамаларды пайдаланады. Бұл модельдер әдетте концептуалды концепциялардың көп санын қамтымайды, бірақ алгебра мен есептеулерді қолдану олардың семантикалық және сипаттау күшін кем емес қамтамасыз етеді. OLTP-де мұндай модельдің мысалы реляциялық алгебра болып табылады. Модельдердің төртінші тобы дизайнға жақсы сәйкес келмейтін, бірақ зерттеу саласын формальды етіп, оны математикалық түрде сипаттайтын модельдерді көрсету үшін қосылды.

Көпөлшемді диаграмманы әр түрлі мәліметтер дәрежесінен қарауға болады, осылайша деректер қаншалықты детализацияланған түрде бейнеленгенін баса көрсетеді. Мысалы, нысанға-бағытталған модельдеудегі кластардың сипаттамасын қарастыруға болады: әрбір класс үшін диаграмма атрибуттарды, әдістерді және шектеулерді көрсете алады немесе тек класс атын ғана көрсетеді. Кластардың атауларына қарап, модельдеу туралы қорытынды жасауға болады, бірақ детализациялану туралы ақпарат қарастырылмаса, онда мәселенің мәнін толық түсіну мүмкін болмайды.

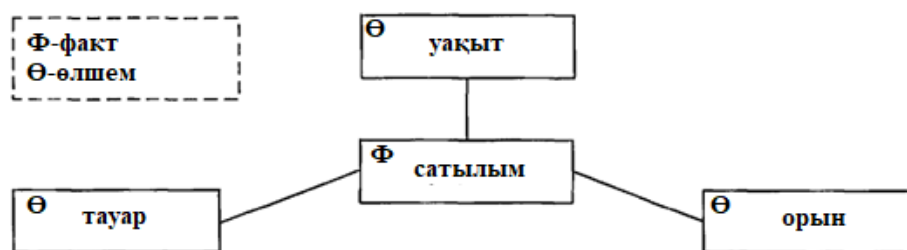
Болашақта көпөлшемді модельдерді сипаттау кезінде біз детализациялаудың үш деңгейін қолданамыз.

- Жоғары: өлшемдер мен фактілер осы деңгейде қарастырылады. Өлшемдер зерттелетін фактілердің аналитикалық кеңістігін анықтау үшін қолданылады.

- Аралық: өлшемдер мен фактілер сәйкесінше деңгейлер мен ұяшықтарға бөлінеді. Өлшемдердің әртүрлі деңгейлері жинақтау иерархияларын құрайды. Әрбір ұяшықта сәйкес фактісі бар әрбір өлшемге арналған детализациялаудың нақты деңгейіндегі деректер бар.

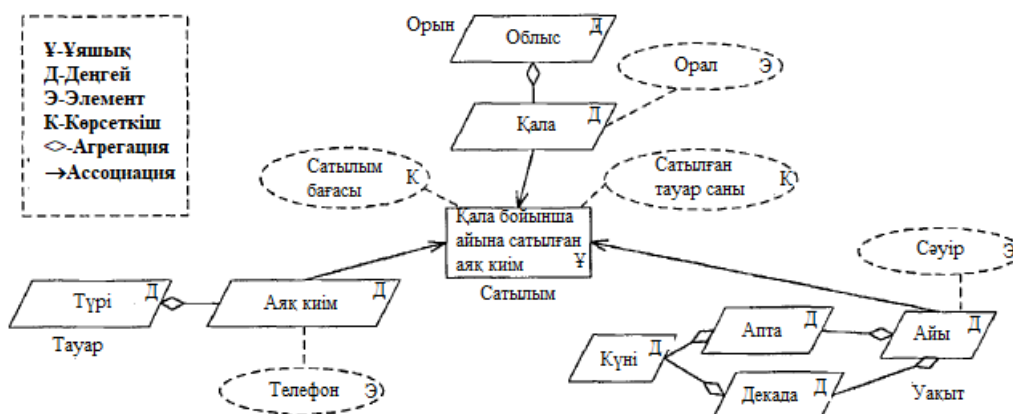
- Төменгі: Ең детализацияланған деңгей өзгерту деңгейлерінің элементтерін және олардың сәйкес көрсеткіш мәндерін көрсетеді.

3-суретте детализацияның ең жоғары деңгейіндегі көп өлшемді диаграмма көрсетілген. Сату фактісін талдау жағдайында сату уақытын, сату орнын және сатылатын нақты өнімді қарастыру қызықты болуы мүмкін. Осылайша, 3 өлшемді кеңістік алынады, оның әрбір нүктесі нақты сатылым туралы мәліметтерді қамтиды және сатылған өніммен, уақыт нүктесімен және сату орнымен анықталады.



Сурет 3. Детализацияның жоғарғы деңгейіндегі көп өлшемді диаграмманың мысалы.

4-суретте сол жоғарыдағы суреттегі диаграмма берілген, бірақ мәліметтері толығырақ көрсетілген. Өлшемдердің әрқайсысы әртүрлі деңгейлердің иерархиясы арқылы сипатталған. Мысалы, уақыт өлшемі күндерге, апталарға, онжылдықтарға және айларға сәйкес деңгейлерді қамтиды. Сол сияқты, таңдалған өлшем деңгейлеріне сәйкес индикаторлардың мәндерін топтастыратын ұяшықтарды бөлектеу арқылы сату фактісін толығырақ қарастыруға болады.



Сурет 4. Аралық және төменгі детализация деңгейлеріндегі көп өлшемді схеманың мысалы.

Соңында төменгі детализация деңгейінің конструкциялары 4-суретте нүктелі сызықпен көрсетілген. Өлшем деңгейлері талдау қажет болатын ұяшық өлшемі мәндерін анықтайтын мүшелерді қамтиды.

Қорыта келе, физикалық модельдер деректерді сақтау технологиясын сипаттайды және аз дәрежеде пайдаланушы деңгейіндегі түсініктерді ұсынуға назар аударады.

Логикалық модельдер соңғы пайдаланушыға жақын, бірақ олар әлі де деректерді сақтау әдісіне байланысты. Концептуалды модельдер басқа модельдерге қарағанда тақырыптық аймақтың бай сипаттамаларын құруға мүмкіндік береді, өйткені олар пайдаланушылардың идеяларын қамту және ұсыну үшін семантикалық дизайнның бай жиынтығын қамтамасыз етуі керек. Өз кезегінде, математикалық модельдер концептуалды конструкцияларды аз қамтамасыз етеді, бірақ олар алгебралар мен есептеулерді қолданады, олардың экспрессивтілігі концептуалды модельдерден кем түспейді және негізгі көп өлшемді конструкцияларды да, олардағы операцияларды да математикалық сипаттауға мүмкіндік береді.

Деректердің көп өлшемді көрінісін сипаттауға арналған көптеген жұмыстардың болуына қарамастан, ашық көздерде реляциялық ДҚБЖ сияқты көп өлшемді сақтау және деректерді талдау жүйелері үшін қатаң, бірыңғай математикалық негіз жоқ. Осылайша, деректерді көп өлшемді ұсынудың математикалық моделін және оны талдау әдістерін жасау қазіргі уақытта өте өзекті мәселе болып табылады.

#### Пайдаланылған әдебиеттер тізімі:

1. Асламов Е.Б., Висков А.В., Фомин М.Б. Свойства аналитических измерений // 42-я Всероссийская' конф. по проблемам математики, информатики, физики и химии: Тезисы докладов. Секции математики и информатики. - М.: Изд-во РУДН, 2006. - С. 30.
2. Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Методы и модели анализа данных: OLAP и Data Mining. - СПб.: БХВ-Петрбург, 2007. - 336 с.
3. Висков А.В., Фомин М.Б. Моделирование аналитических измерений в многомерных базах данных // Вестник ИрГТУ. 2012. № 4 (63). С. 15–19.
4. Thomsen E. OLAP Solution: Building Multidimensional Information System. NY, Wiley Computer Publishing, 2002, 688 p.
5. Фомин М.Б. Описание метаданных многомерных информационных систем с использованием кластерного метода // Вестник Иркутского государственного технического университета. 2017. Т 21. № 7. С. 78-86. DOI: 10.21285/1814-3520-2017-7-78-86
6. Информационные системы в экономике: Учеб. Пособие / Под ред. Проф. Д.В. Чистова. - М.:ИНФРА-М, 2009. - 234 с.
7. Балдин К.В., Уткин В.Б. информационные системы в экономике: Учебник. - 5-е изд. - М.: «Дашков и К », 2008. - 395 с.
8. D. Davidov, E. Gabrilovich, S. Markovitch. Parameterized Generation of Labeled Datasets for Text Categorization Based on a Hierarchical Directory // Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval, Sheffield, UK, 2004.
9. Блинов С.Ю., Коробейников А.Г., Лейман А.В., Демина Е.А. Систематизация разнородной информации в задаче фильтрации спама// В книге "Труды конгресса по интеллектуальным системам и информационным технологиям AIS-IT'12. Научное издание в 4-х томах. М.:Физматлит, 2012, - Т.2. стр. 18-22.
10. Ранганатан Ш. Р. Классификация двоеточием: основная классификация. Пер. с англ. / Под. ред. Т. С. Гомолицкой, Х.М. Зайдберга и П.И. Шифман. -М.: ГПНТБ СССР, 1970 - 421 с.

# МОДЕЛЬ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРЕДСКАЗАНИЯ ФУНКЦИЙ БЕЛКОВ

Голенко Е.С., Исмаилова А.А.

Казахский агротехнический университет им. С. Сейфуллина, Астана, Казахстан

**Аннотация.** Прогнозирование функции белков — важнейшая часть аннотации генома, которая может помочь в решении широкого круга биологических задач. Существует множество методов для предсказания функций белков. Однако, за исключением последовательности, большинство признаков трудно получить или они недоступны для многих белков, что ограничивает область их применения. Кроме того, производительность методов прогнозирования признаков на основе последовательностей часто ниже, чем у методов, включающих несколько признаков, а предсказание признаков белков может занимать много времени. Последние достижения в этой области связаны с развитием машинного обучения, что свидетельствует о большом прогрессе в решении задачи предсказания функций белков. Однако сегодня большинство белковых последовательностей имеют статус «неохарактеризованных» или «предполагаемых». Необходимость оценки точности идентификации функций белков является актуальной задачей для подходов машинного обучения, используемых для прогнозирования функций белков. В этом исследовании была реализована и оценена модель ProtCNN. В результате изучения семейств Pfam ProtCNN достигает уровня точности 0,988%. Использование набора данных Pfam позволило повысить точность классификации за счет большого набора обучающих данных. Качество прогноза повышается при большом количестве обучающих данных. Исследование показало, что алгоритмы машинного обучения могут быть использованы в качестве эффективного инструмента для построения моделей прогнозирования функций белков, в частности, сеть CNN может быть адаптирована в качестве точного инструмента для аннотирования функций белков при наличии больших наборов данных.

**Ключевые слова:** предсказание функции белка, классификация, нейронные сети, ProtCNN.

## 1. Введение

Белки представляют собой длинные цепочки аминокислот, которые формируются из информации, полученной от ДНК, а затем складываются в трехмерные формы. Формы, которые принимают белковые молекулы, определяются информацией, заложенной в ДНК, а то, в какую форму складывается сама молекула ДНК, зависит от состава аминокислот в цепи. В свою очередь форма в биологии определяет функцию.

Прогнозирование функции белка может быть рассмотрено как задача классификации с несколькими метками, в которой есть набор функций  $F=(F_1, \dots, F_m)$ . Для заданного набора белков  $P=(P_1, \dots, P_n)$ , где первые  $l$  белков помечены как  $y_1, \dots, y_l$ , каждый  $y_i$  является вектором с  $y_{ij}=1$  в случае, если белок  $P_i$  связан с  $j$ -й функцией  $F_j$ , иначе  $y_{ij}=0$ . Цель состоит в том, чтобы предсказать метки  $y_{l+1}, \dots, y_n$  для оставшихся немеченых белков  $P_{l+1}, \dots, P_n$ .

Можно использовать различные подходы для решения задач классификации с несколькими метками, но самый простой способ — рассматривать каждый термин Gene Ontology (GO) как независимую задачу классификации. Наиболее подходящий вариант для решения этой задачи — модели машинного обучения с лучшим соотношением качества классификации и времени расчета. Огромным преимуществом машинного обучения является то, что производительность и точность модели имеют стабильный рост при большом объеме обучающих данных. Учитывая стремительный рост количества белковых последовательностей в открытых базах данных, это обстоятельство делает использование машинного обучения для решения задачи предсказания функций белков актуальной задачей для исследований [1].



## 2. Обзор литературы

Понимание функции белков важно для изучения биологических механизмов развития болезней и открытия лекарств. Многие базы данных обновляются ежедневно, чтобы предоставить функциональные аннотации с разных точек зрения, таких как взаимодействие белок-белок, биологическая сеть и многие конкретные функциональные классы. Однако общее количество обнаруженных белковых последовательностей значительно превышает количество белков, характеризующихся известной функцией. Чтобы сократить разрыв между количеством охарактеризованных и не охарактеризованных белковых последовательностей, исследуются тысячи высокопроизводительных геномных проектов, но только 1% найденных последовательностей подтвержден экспериментальной аннотацией [2]. Это создало большую потребность в разработке теоретических методов аннотирования функций белков. Для обнаружения функций белков был разработан и широко используется широкий спектр методов. К ним относятся кластеризация последовательностей, слияние генов, сходство последовательностей, изучение эволюции, структурное сравнение, белок-белковое взаимодействие, функциональная классификация с помощью признака, полученного из последовательности и домена и интегрированные методы, которые в совокупности рассматривают несколько методов и данных для продвижения производительность предсказания функции [3]. За исключением методов, использующих подобия последовательностей, недостатком этих методов является то, что они используют особенности белков, которые трудно получить или вообще недоступны для многих белков, что сильно ограничивает область применения таких методов.

Сверточная нейронная сеть (CNN) изначально разрабатывалась для обработки 2D-изображений, однако сегодня успешно используется для обработки геномных последовательностей. CNN состоит из разных типов слоев: сверточных слоев, слоев субдискретизации и слоев «обычной» нейронной сети — персептрона [4-6].

Модель машинного обучения ProtCNN использует остаточные блоки, вдохновленные архитектурой ResNet, которая также включает расширенные свертки для обеспечения большего рецептивного поля без увеличения количества параметров модели. В качестве входного сигнала в сеть передается одна горячечодированная невыровненная аминокислотная последовательность с заполнением нулями. В общем случае формирование выходной карты признаков скрытого слоя  $l$  архитектуры CNN можно описать следующим образом:

$$h_j^l = f(\sum_i x_i^{l-1} * k_j^l + b_j^l), \quad (1)$$

где  $f$  — функция активации;  $b_j$  — коэффициент сдвига для карты объектов;  $k_j$  — номер ядра свертки  $j$ ;  $x_i^{l-1}$  — карта признаков предыдущего слоя; \* — операция свертки. На Рис. 1 показана архитектура модели CNN.

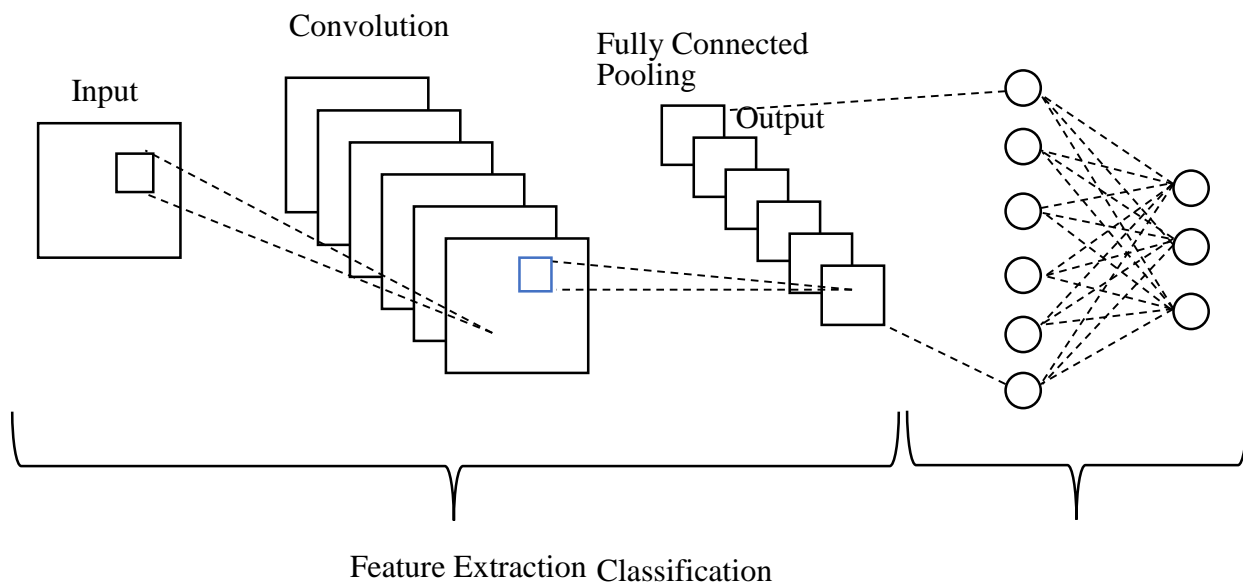


Рис.1. Архитектура CNN

### 3. Материалы и методы

При реализации модели использовались:

- алгоритм *Adam* как алгоритм оптимизации [7];
- показатель точности *Accuracy* как целевая функция;
- Бинарно-кроссэнтропийная функция, возвращающая ошибку классификации как функцию логистических потерь *Loss*:

$$Loss = -\frac{1}{N} \sum_{i=1}^N y_i * \log(\hat{y}_i) + (1 - \hat{y}_i) * \log(1 - \hat{y}_i), \quad (2)$$

где  $y_i$  — истинная метка класса;  $\hat{y}_i$  — отклик классификатора (вычисляемая метка класса) на  $i$ -й объект;  $N$  - количество классов.

Архитектура ProtCNN начинается с начальной операции свертки, которая применяется к входным данным с размером ядра 1 для извлечения основных свойств. Два остаточных блока используются для захвата сложных закономерностей в данных, что помогает проводить обучение с большим количеством эпох и повышать производительность модели. После двух остаточных блоков применяется максимальное объединение для уменьшения пространственного размера представления. Dropout добавлен для регуляризации, чтобы предотвратить переобучение модели.

В качестве основного источника данных для обучения и тестирования модели использовалась общедоступная база данных Pfam. Pfam — это база данных семейств белковых доменов. Каждое семейство в нем представлено множественным выравниванием фрагментов белковых последовательностей и скрытой марковской моделью (СММ). 77,2% из ~137 млн последовательностей в UniprotKB имеют по крайней мере одну аннотацию семейства Pfam [8]. По состоянию на ноябрь 2021 г. в Pfam насчитывалось 19 632 записи (семейства), объединенных в 657 кланов [9].

Эта база данных содержит пять функций:

- *sequence*: последовательность аминокислот для данного домена. Эта последовательность представляет собой домен, а не полноценный белок.
- *family\_accession*: номер доступа.
- *sequence\_name*: имя последовательности.

- align\_sequence: содержит одну последовательность из множественного выравнивания последовательностей.
- family\_id: фамилия из одного слова.

Эталонный тест включает в себя случайное разделение 17 929 семейств Pfam на тестовую последовательность, где 80 % последовательностей используются для обучения, 10 % для настройки модели и 10 % в качестве тестовых последовательностей.

Для оценки эффективности модели предсказания свойств белка использовался показатель точности классификации *Accuracy*, определяемый соответственно формулой (3):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

где *TP*, *TN*, *FP* и *FN* представляют собой истинно положительные, истинно отрицательные, ложноположительные и ложноотрицательные результаты [10].

#### 4. Обсуждение результатов

Таблица 1 иллюстрирует результаты тестирования сети ProtCNN, полученные в течение первых 10 эпох обучения.

Таблица 1  
Результаты обучения модели ProtCNN

Epoch	Loss	Accuracy	Val_loss	Val_accuracy
1	0.9027	0.9306	0.4603	0.9843
2	0.4362	0.9790	0.4435	0.9849
3	0.4296	0.9814	0.4326	0.9864
4	0.4216	0.9823	0.4394	0.9855
5	0.4131	0.9832	0.4263	0.9861
6	0.4051	0.9835	0.4281	0.9852
7	0.3967	0.9840	0.4088	0.9872
8	0.3907	0.9844	0.4124	0.9870
9	0.3865	0.9845	0.4034	0.9871
10	0.3787	0.9849	0.3954	0.9879

На Рис. 2 показаны графические представления результатов обучения модели ProtCNN для 10 эпох с точки зрения точности и потерь соответственно.

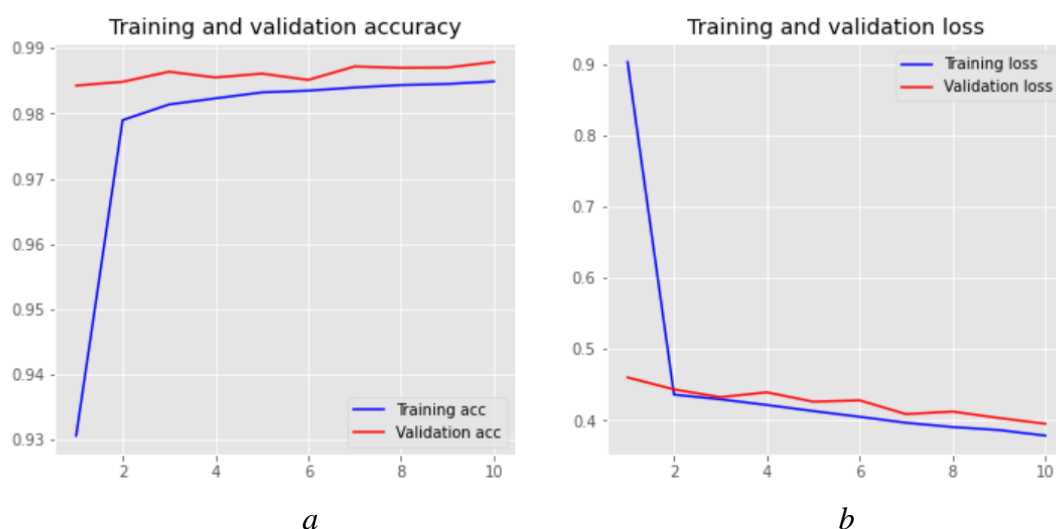


Рис. 2. Результаты обучения модели ProtCNN по показателю Accuracy: *a* – значение Accuracy, *b* – значение функции потерь Loss

На случайно разделенных данных ProtCNN достигает точности 0,988%. Полученные результаты определяются двумя факторами. Первый — это объем наборов данных, на которых обучались модели. Учитывая ~ 1,1 миллиона обучающих примеров в 17 929 выходных семействах самых разных размеров, изучаемые модели очень точны. Во-вторых, использование расширенных сверток в ProtCNN помогло модели изучить более сложные функции, расширения позволяют использовать более крупные рецептивные поля. Пропущенные соединения помогли модели сохранить важную пространственную информацию из предыдущих слоев. Применение более сложных методов машинного обучения может привести к дальнейшему повышению производительности.

## 5. Заключение

В этом исследовании была реализована и обучена модель машинного обучения ProtCNN. При реализации модели применялись следующие инструменты: алгоритм Адама в качестве алгоритма оптимизации, индикатор точности в качестве целевой функции и функция бинарной кроссэнтропии, возвращающая ошибку классификации в виде функции логистических потерь Loss. Модель прогнозирования была обучена с использованием обучающего набора и оценены на тестовом наборе. Оценка производительности реализованной модели показала, что точность ProtCNN составляет 0,9967 для набора обучающих данных, 0,9877 для набора значений и 0,9880 для тестового набора.

## Список использованной литературы

- [1] Deen A., Gayanchandani M. (2019). Protein Function Prediction using SVM Kernel Approach. International Journal of Scientific & Engineering Research, 10, ISSN 2229-5518. Available at: <https://www.ijser.org/researchpaper/Protein-Function-Prediction-using-SVM-Kernel-Approach.pdf>
- [2] Barrell D., Dimmer E., Huntley R.P., Binns D., O'Donovan C., Apweiler R. (2009). The goa database in 2009—An integrated gene ontology annotation resource. Nucleic Acids Res., 37, 396–403. doi: <https://doi.org/10.1093/nar/gkn803>
- [3] Piovesan D., Giollo M., Leonardi E., Ferrari C., Tosatto S.C. (2015). Inga: Protein function prediction combining interaction networks, domain assignments and sequence similarity. Nucleic Acids Res., 43, 134–140. doi: <https://doi.org/10.1093/nar/gkv523>
- [4] Kulmanov M., Khan M.A., Hoehndorf R. (2018). Deepgo: predicting protein functions from sequence and interactions using a deep ontology-aware classifier. Bioinformatics, 34, 660–668. doi: <https://doi.org/10.1093/bioinformatics/btx624>
- [5] Cai Y., Wang J., Deng L. (2020). Sdn2go: an integrated deep learning model for protein function prediction. Frontiers in Bioengineering and Biotechnology, 8, 391. doi: <https://doi.org/10.3389/fbioe.2020.00391>
- [6] Du Z., He Y., Li J., Uversky V.N. (2021). Deepadd: protein function prediction from k-mer embedding and additional features. Computational Biology and Chemistry, 89, 107379. doi: <https://doi.org/10.1016/j.compbiolchem.2020.107379>
- [7] Kingma D.P., Ba J. L. (2014). Adam: A Method for Stochastic Optimization. 3rd International Conference for Learning Representations, 1412.6980. doi: <https://arxiv.org/abs/1412.6980>
- [8] UniProt Consortium. (2016). Uniprot: the universal protein knowledgebase. Nucleic Acids Res., 45, 158–169. doi: <https://doi.org/10.1093/nar/gkw1099>
- [9] Pfam. Availale at: <https://xfam.wordpress.com/2021/11/19/pfam-35-0-is-released/>
- [10] Vu T.T.D., Jung J. (2019). Protein function prediction with gene ontology: from traditional to deep learning models. PeerJ., 9, e12019. doi: <https://doi.org/10.7717/peerj.12019>

## КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ВЕДЕНИЯ СКЛАДСКОГО УЧЕТА В СТРОИТЕЛЬНОМ БИЗНЕСЕ

*Дрозд В.Г., Целых У.Е., Сорокин А.Д., Ярославцев С.В.  
Карагандинский университет Казпотребсоюза, г. Караганда, vgdrozd@mail.ru*

**Аннотация.** В данной статье рассмотрена автоматизация складского учета строительного бизнеса. Было разработано приложение «Accounting». Назначение данного приложения состоит в том, чтобы автоматизировать процесс ведения складского учета. Автоматизирован процесс ведения учета товаров, учета поставщиков, формирование необходимых документов и отчетов. А также расчет оптимальных размеров товарных запасов и дальнейшего планирования размеров и номенклатуры заказов.

**Введение.** Актуальность темы научного исследования обуславливается необходимостью автоматизации складского учета с целью снижения временных и денежных затрат на выполнение стандартных рутинных операций.

Автоматизация складского учета оказывает существенное влияние на качественный показатель и скоростные параметры выполнения основных рабочих процессов выполняемых в складской деятельности, приводит к развитию и дальнейшему совершенствованию, как систем управления, так и регулирования информационных потоков на складе. Это осуществляется путем применения более современного программного обеспечения и проведение компьютерного моделирования на складе.

Внедрение информационной системы управления складской деятельности на предприятии процесс весьма сложный и многомерный. Для решения данных проблем, которые возникают в процессе разработки и внедрения ИС применяют различные подходы.

**Обзор литературы.** Многие авторы рассматривают различные аспекты внедрения компьютерного моделирования при разработке ИС. В научной статье «Внедрение информационных систем», автор Тренева Н.Н., детально рассматриваются проблемные вопросы, связанные с необходимостью автоматизации и информатизации, существующими особенностями и сложностями при внедрения ИС. Автор отмечает, что управление деятельностью предприятия основывается не на достижении максимизации ключевых показателей, а на выстраивании компромисса между множеством существенных факторов. Достоинством научной работы стало наличие большого числа примеров по их внедрению на предприятиях, которые способствовали увеличению экономической эффективности. Также проведено детальное рассмотрение требований к информационной системе [1].

В научной статье «Проблемы внедрения информационной системы управления предприятием», автор Баламирзоева Н.Л., приведены основные проблемные вопросы и задачи, появляющиеся в ряде случаев при внедрении ИС управления складского хозяйства предприятия и даются рекомендации и предложения по их решению [2].

В научной работе «Обучение персонала как часть процесса внедрения информационно-аналитических систем», автор Л. Шапиро, приведен алгоритм организации и проведения повышения квалификации в рамках внедрения ИС на предприятии [3]. Данный автор предлагает разделять процесс обучения на ряд этапов, отмечает насущную необходимость для подготовки персонала предприятия к внедрению ИС. В статье «Как выбрать программу «1С» для торговой фирмы» был произведен сравнительный анализ существующих программ «1С: Предприятие». Так же в данной статье описаны все имеющиеся возможности конфигурации программы, их цена и системные требования, исходя из которых, можно определиться с выбором программы, которая в большей степени подойдет для малого предприятия [4].

Основной задачей является создание автоматизированного рабочего места, которое позволит выдавать отчет в разделе поставщиков и договоров, которое также позволит вести контроль за поставкой товара на склады предприятий [5].

Достигаются следующие целевые задачи:

- активное управление складским хозяйством;
- получение точной и достоверной информации о месте нахождения товара на складе;
- эффективное управление товаром с ограниченным сроком годности;
- получение инструмента для повышения эффективности и развития процессов по обработке товара на складе;
- оптимизация использования складских площадей.

Для использования программного приложения предлагаются следующие прикладные решения:

- Бухгалтерия.
- Зарплата и Управление Персоналом.
- Управление производственным предприятием.
- Комплексная автоматизация.
- Управление торговлей.

**Методы.** Программное приложение разработано и полностью отвечает специфике складского учета, удобно в восприятии и обслуживании, отличается понятным интерфейсом и легко адаптируется к любым бизнес процессам.

Функциональные возможности приложения охватывают все возможные ситуации в работе склада. Приложение поддерживает ведение базы товаров, поставщиков и сотрудников, а также обеспечивает правильное взаимодействие этих баз.

Реализованы основные функции приложения для учета материалов:

- Возможность вносить товары в базу.
- Поиск товаров в приложении.
- Добавление и редактирование параметров учета.
- Импорт/экспорт товарной базы.
- Ведение данных статистики поступившей и списанной продукции.
- Получение напоминаний.
- Подбор товара по году, цвету, производителю, сорту и т.д.
- Ведение и хранение отчетов/сводок за период по движению.

Автоматизация устраняет проблемы ненужных трудозатрат, экономит время на ручном учете и формировании документации. Снижение объема хищений наблюдается после внедрения автоматизации складского учета в сочетании с системой инвентаризации. Учет движения товарно-материальных запасов склада выявляет лицо, ответственное за недостачу, к которому применяются штрафные санкции [6].

Преимущества автоматизации:

- Возможность вводить данные однократно и в дальнейшем использовать в разных видах операций.
- Автоматизация формирования отчетов и выходных документов.
- Устранение ошибок и «человеческого фактора».

**Обсуждение.** Важной причиной ведения автоматизированного складского учета является обеспечение мобильности. Автоматизация склада позволяет работать не только с одним складом, но и параллельно с несколькими. Система делит склад на несколько отделов, и для каждого отдела устанавливаются независимые параметры. Складская отчетность представлена в удобном виде, чтобы вы могли быстро разобраться во всех процессах и продумать дальнейшие действия.

Складская отчетность представлена в удобном виде, чтобы пользователь мог быстро разобраться во всех процессах и продумать дальнейшие действия. Программное приложение можно адаптировать под маленькое ИП, просто отключив ненужный

функционал. Помогает оптимизировать и эффективно управлять. Отличная функциональность, которую смогут использовать даже неопытные пользователи. При необходимости разворачивается и сворачивается без ущерба для функциональности. Не перегружает операционную систему компьютера.

Автоматизация склада также может стать основным улучшением в более широкой цепочке поставок и, что еще более важно, позволит реализовать ранее невозможные и даже невообразимые стратегические варианты и получить стратегические преимущества.

Все функции, выполняемые программным приложением по инвентаризации товаров в магазине строительных материалов, дают предпринимателю определенные экономические или организационные преимущества. Поскольку любая новая реализация, в том числе автоматизация, окупается за счет выгод, полученных от ее использования. Но понесенные затраты окупятся в течение первого года, и преимущества инноваций будут ощутимы немедленно. Поэтому инвестиции в программы автоматизации и складского учета всегда перспективны и выгодны [7].

Благодаря внедрению процессов автоматизации можно проводить подсчет остатков товара на складе, прогнозировать и планировать заказы продукции. Внедрение элементов автоматизации позволит расширить ассортимент продукции, повысить объемы поставок и продаж.

Помимо выполняемых учетных операций имеются и другие полезные достоинства:

- можно проводить комплектацию заказа товаров;
- выполнять контроль по приему продукции или сырья от поставщика;
- обеспечивать оптимальное распределение поступающих товаров по складским площадям;
- выполнять процедуры оформления возврата товаров от клиентов;
- ведение учетных данных по каталогу.

Автоматизация складской деятельности позволяет включить складское хозяйство в единое информационное пространство всего предприятия, что дает возможность оперировать актуальными данными обо всех остатках товара на складе, избегая ошибок работы с клиентом [8].

С помощью автоматизации можно минимизировать:

- создаваемый беспорядок на полках хранения товаров;
- мошеннические действия;
- ускорить процесс обслуживания заказов;
- минимизировать или полностью исключить ошибки персонала;
- снизить накладные издержки;
- выполнять списание товаров с вышедшим для него сроком годности.

Автоматические склады лишены большинства недостатков. Программное приложение включает в себя обширную базу данных, включающую в себя как уже имеющиеся документы, так и многочисленные шаблоны, с которыми заполнение любых форм не составит труда для работника любой квалификации.

В результате ведения складского учета, программное приложение по запросу клиента выдает соответствующие отчетные документы: по приходу товара, его расходу или возврату [9].

Для выполнения задачи «Учет товаров», была разработана база данных, предназначенная для хранения и учета данных о продажах. А также хранение и редактирование регистрационных данных поставщиков, сотрудников, клиентов.

Данное приложение позволит значительно ускорить процесс приемки товара на склад, а также ликвидировать длительный простой транспорта на складе. Это, в свою очередь, позволит значительно уменьшить возможность недопоставку товара при пересчете, и позволит в поставленные сроки подготовить документацию для использования товара в дальнейших процессах. Все это в дальнейшем сможет положительно сказаться на деятельности компании в целом [10].

Существенным отличием предложенного способа автоматизации торгового склада от других решений в этой области является доступность. Работа с системой не предполагает покупку программного обеспечения, внедрение его на предприятии и обучение персонала.

В программном приложении использована локальная база данных, что позволило уменьшить время работы с базой данных и уменьшило сложность настройки прикладного программного обеспечения.

Исходный программный продукт разработан на объектно-ориентированном языке программирования С#. Данное программное обеспечение позволяет собирать статистику по скорости передачи данных.

В качестве элементов интерфейса использовались визуальные формы, кнопки, переключатели и меню. Разработанный пользовательский интерфейс позволяет оптимизировать бизнес процессы, а также повысить объем продаж.

Далее рассмотрим работу кнопочных форм, которые приведены на рисунках 1-2.

	Id	Материалы	Поставщики	Производители	Количество	Цена	Номер партии	Срок гарантии (год)
▶	1	Герметик	Alibaba	Makroflex	1432	1550	3	2
	2	Окна	SAVA	Veka	832	41082	8	3
	3	Обои	СтройДекор	Aura	2569	5400	2	1
	4	Монтажная пена	Soudal	Soudal	8254	2345	1	1
	5	Ламинат	СтройДекор	Artens	7543	31	5	2
	6	Стекланная пли...	Bluwahle	BlueHale	1324	1863	2	1
	7	Жидкие гвозди	Asmaco	Asmasy	95612	1345	7	1
	8	Линолеум	Би-Плюс	Би-Плюс	7542	1234	3	45
	9	Паркет	Мир паркета	Мир паркета	15234	985	8	4
*								

Рисунок 1 – Внешний вид формы «Учет материалов»

В данной форме содержатся отчеты, создаваемые для анализа работы склада, (магазина) и предоставления данных по товарам. Данная форма предназначена для ввода данных в таблицы через формы для ввода данных (Наименование товара, поставщики, производители, количество товара на складе (по названию), цена, номер партии, срок гарантии).

	Id	Клиент	Дата	Номер счета	Скидка	Сумма	Delete
▶	1	Серафима Стел...	30427	256646466	0	250000	<a href="#">Delete</a>
	2	Анна Вест	54982	447455555	15	545785	<a href="#">Delete</a>
	3	Наталья Алексе...	55444	444455211	7	554777	<a href="#">Delete</a>
	4	Евгений Геннад...	45545	544778888	0	307815	<a href="#">Delete</a>
	5	Феликс Вест	55526	672799798	45	445555	<a href="#">Delete</a>

Рисунок 2 – Внешний вид формы «База данных клиентов»

Данная форма предназначена для ввода информации о счетах: кто и на какую сумму оплатил счет за определенное количество с определенной скидкой и какие товары были проданы на этом счете. В этой форме также есть кнопки для добавления данных клиента, выхода из этой формы и печати этой формы.



Разработанная база данных может хранить данные о продуктах, поставщиках, сотрудниках и продажах. Проектирование осуществлялось путем построения концептуальной модели базы данных.

Разработанная база данных обеспечивает эффективную работу, создает удобство для пользователей, как для сотрудников, так и для его клиентов.

**Заключение.** Преимущества внедрения программного приложения для управления складом очевидны. Автоматизация склада позволит вам разгрузить сотрудников, оптимизировать все процессы: от приема товара до его отгрузки. Результатом проведенной разработки программного приложения является программный продукт. Программное приложение включает в себя все этапы хранения, поиска и сортировки необходимой информации, предоставляет удобные инструменты, позволяющие повысить качество и скорость каждой операции.

Разработанное приложение имеет практическое применение для организаций. Внедрение данного программного продукта в сфере малых коммерческих предприятий, благодаря простоте системы, не займет много времени. Система организована таким образом, что в нее легко и удобно вносить изменения и подстраиваться под специфику отдельных организаций.

#### Список использованной литературы:

1. Тренев Н. Н. «Внедрение информационных систем» [Электронный ресурс]. – Режим доступа: <http://www.k-press.ru/comp/2000/3/trenev/trenev.asp> (дата обращения: 27.05.2019).
2. Баламирзоев Н. Л. «Проблемы внедрения информационной системы управления предприятием» [Электронный ресурс]. – Режим доступа: <https://e-koncept.ru/2014/54385.htm><https://e-koncept.ru/2014/54385.htm> (дата обращения: 27.05.2019).
3. Шапиро Л. «Обучение персонала как часть процесса внедрения информационно-аналитических систем» [Электронный ресурс]. – Режим доступа: <http://bit.samag.ru/archive/article/1348> (дата обращения: 27.05.2019).
4. Как выбрать программу «1С» для торговой фирмы [Электронный ресурс]. – Режим доступа: [http://center-compotech.ru/articles/st\\_311013-12.html](http://center-compotech.ru/articles/st_311013-12.html) (дата обращения: 01.05.2019).
5. LINQ to SQL [Электронный ресурс]. Режим доступа: <https://metanit.com/sharp/adonet/4.1.php>
6. Microsoft [Электронный ресурс]. Режим доступа: <https://www.microsoft.com/ru-ru/store/b/home?irgwc=1>
7. Троелсен Э. Язык программирования C# 5.0 и платформа .NET 4.5. 6-е изд. – М.: ООО «И.Д. Вильямс», 2015. – 1312 с. 308
8. Гуцин А.Н. Базы данных: учебно-методическое пособие. – М., Берлин: Директ-Медиа, 2015. – 311 с.
9. А.Д Хомоненко. Delphi 7. Санкт-Петербург, 2008 г.; изд. «БХВ-Петербург»
10. Джеффри Ульман. Системы баз данных. Санкт-Петербург, 2009 г.; изд. «Вильямс».

## АУЫЛ ШАРУАШЫЛЫҚ БАҒЫТЫН АВТОМАТТАНДЫРУ БОЙЫНША АҚПАРАТТЫҚ ЖҮЙЕЛЕРДІ ҚҰРУ МӘСЕЛЕЛЕРІ

К.М. Алдабергенова<sup>1</sup>, М.А. Кантуреева<sup>2</sup>

<sup>1,2</sup>Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Нұр-Сұлтан қ, Қазақстан  
(E-mail: <sup>1</sup>[kamar\\_suli\\_9028@mail.ru](mailto:kamar_suli_9028@mail.ru), <sup>2</sup>[ma\\_khantore@mail.ru](mailto:ma_khantore@mail.ru).)

**Аннотация:** Бұл мақалада ауыл шаруашылығы өндірісін техникалық қамтамасыз ету жүйесін дамытудың жолдары, аграрлық сала кәсіпорындарындағы технологиялық процестері туралы, сонымен қатар машина технологиялық станциялары сияқты ауылдық тауар өндірушілерді техникалық қамтамасыз етудің ерекшелігі мен ақпараттық қамтамасыз ету жүйесін қалыптастырудың негіздері туралы қарастырылған.

**Кілттік сөздер:** ауыл шаруашылығы өндірісі, ақпараттық жүйелер, МТС.

Қазақстандағы ауыл шаруашылығы өндірісі өткен ғасырмен салыстырғанда едәуір айырмашылығы байқалады. Дегенмен, технологиялық жарақтандыру деңгейінің төмендігі мен білікті кадрлардың жетіспеушілігінен агроөнеркәсіп кешенінің дамуы тежелуде. Бүгінгі таңда ауылшаруашылық жұмысындағы әлемдік және еуропалық тәжірибе қазірдің өзінде ақпараттық технологиямен тікелей байланысты. Аграрлық секторы дамыған жетекші елдердің тәжірибесі олардың барлығы да өзіндік «технологиялық революцияны» бастан өткергенін көрсетеді. Тәжірибенің арқасында есептелген қарапайым ауыл шаруашылығын қазір нақты ғылымдар алмастыруда. Ең бастысы, геоақпараттық технологиялар, көп операциялық энергия үнемдейтін ауыл шаруашылығы қондырғылары, жоғары өнімді өсімдік сорттарын таңдау және жоғары өнімді мал тұқымын өсіру, биологиялық белсенді жемшөп қоспаларын жасау, жануарларға арналған жаңа дәрі-дәрмектер, эпизоотиямен күресудің заманауи әдістері, карантин. жануарлар мен өсімдіктердің аурулары көбіне қолданылады. Көптеген ұқсас тапсырмалар қазір ТМД аумағында шешімдердің мысалдары бар. Басшылық жағдайды дер кезінде және дұрыс бағалап, ресурс үнемдейтін инновациялық технологияларға көшкен шаруа қожалықтары қолда бар түрлі ақпараттық технологиялар мүмкіндіктерін пайдалана бастайды. Ауыл шаруашылығын реформалаудың жаңа кезеңінде агроөнеркәсіптік кәсіпорындарды жаңғырту үшін техникалық және ақпараттық құралдарды өндіру мен таратудың өткір қажеттілігі туындады. Бүгінгі күні ақпараттық ресурстар қалыптастырылатын, деректер базасы, қолданбалы бағдарламалық өнімдер мен ауыл шаруашылығы өндірісінің тиімділігін арттыру бойынша ұсыныстар жинақталатын, жинақталған және бейімделетін мамандандырылған жүйелер бар.

1-кесте – ТМД елдеріндегі ауылшаруашылық бағытының бағдарламалық қамтамасыз ету мысалдары.

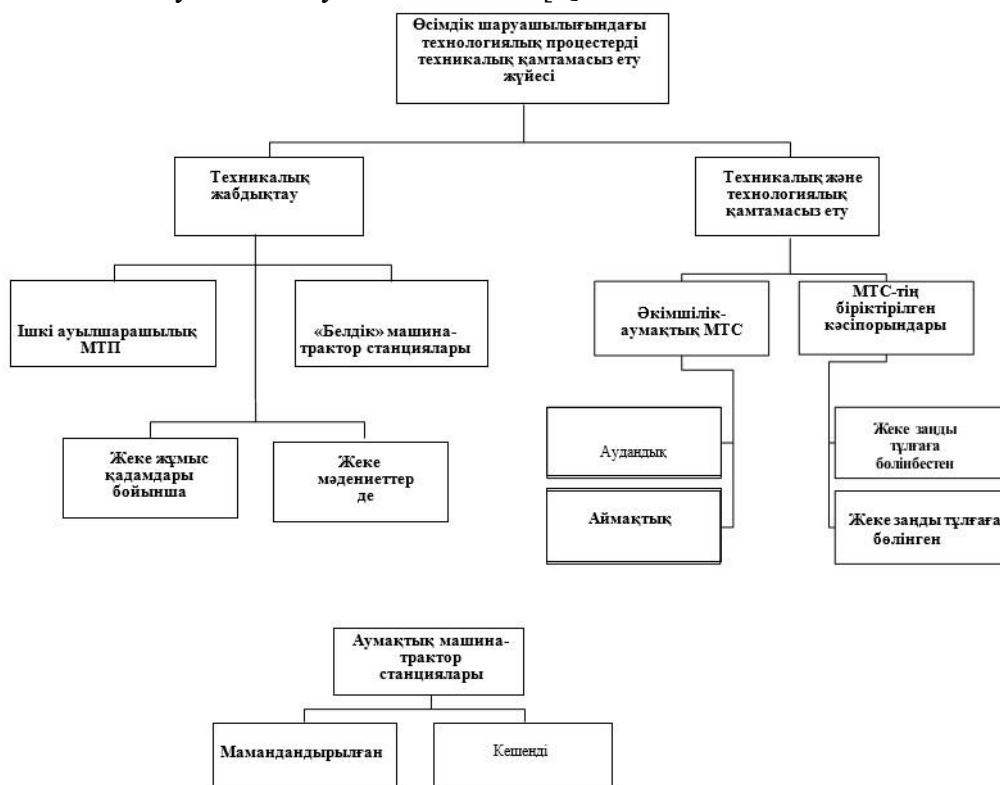
Аты	Қысқаша сипаттамасы
Ветеринария және мал шаруашылығы	Жануарларды азықтандыру, ветеринариялық сараптау, тізімі туралы ақпаратты қамтитын ақпараттық жүйе
«Ветеринария және мал шаруашылығы» кешені	Малды азықтандырудың қажетті мөлшерін есептеу. Рулық(племенного) есепті автоматтандыру.
«Механизация» бағдарламалық өнімі	Ауыл шаруашылығы техникасының деректер базасы. Өңдеу өнеркәсібіне арналған жабдықтар мен осы жабдықты өндіретін және жеткізетін кәсіпорындардың тізімі
«Агрохимиялық күтім және өсімдіктер карантині»	Минералды, органикалық және органикалық-минералды тыңайтқыштар мен химиялық мелиоранттар туралы ақпаратты қамтитын өңдеуші агенттердің тізбесі, сондай-ақ жеткізушілердің тізімі.
«Ауыл шаруашылығы дақылдарын өсірудің	Аудандық, өндірістік және технологиялық жағдайларды ескере отырып, ауыл шаруашылық дақылдарының түрлерін

дәстүрлі және озық технологиялар» бағдарламасы	өндірудің негізгі технологияларын жоспарлау. Сондай-ақ сатудан түскен қаржылық жағдайларды есепке алу.
Бағдарламалық өнім «жалпы табыс принципіне негізделген экономикалық қызметті жоспарлау»	Ауыл шаруашылығы кәсіпорындарының шаруашылық қызметін жоспарлау және талдау. Кірістерді, шығыстарды және әртүрлі пайдаларды есептеу.

Көптеген елдерде жаңа дақылдарды, сорттарды, технологиялар мен техникаларды насихаттау мен таратуға көп көңіл бөлінді. Аграрлық секторды дамытудың теориялық бөлігі әлдеқайда алға кетті, ең жаңа селекциялық технологиялар шығарыла бастады, бірақ бұл саладағы ақпараттық технологиялар айтарлықтай дамуға ие болмады. Заман талабына сай білім беретін ауылшаруашылық орталықтары күннен-күнге ашылуда. Бұл орталықтар қарапайым шаруаларды да, ауыл шаруашылығы кәсіпорындарының басшыларын да, мамандары мен жұмысшыларын да оқытады.

Аграрлық саладағы технологиялық процестерді техникалық қамтамасыз ету проблемалары ауылшаруашылық өндірісінің дамуымен бірге пайда бола бастады. Дақылдарды өсіру технологиялары неғұрлым қарқынды болса, оларды жүзеге асырудың техникалық құралдарына қойылатын талаптар соғұрлым жоғары болды

Ауыл тауар өндірушілерін техникалық қамтамасыз ету жүйесі технологиялық процестерді іске асырудың материалдық базасы болып табылады. Бүгінгі таңда бұл жүйеде техникалық және техникалық-технологиялық қолдаудың ішкі жүйелерін бөлуге болады. Техникалық қамтамасыз ету жүйесінің өзегі ауыл шаруашылығы тауарын өндірушілердің машина-трактор парктері (МТП) болып табылады. Машина-трактор паркі деп аграрлық сала кәсіпорындарында технологиялық процестерді іске асыру кезінде механикаландырылған жұмыстардың орындалуын қамтамасыз ететін ауыл шаруашылығы техникасының тоғысуы деп айтумызға болады. [1]



Сурет 1. Өсімдік шаруашылығындағы технологиялық процестерді техникалық қамтамасыз ету жүйесінің схемасы

Көбінесе МТП ауыл шаруашылығы кәсіпорындарының құрылымдық бөлімшелері (ішкі ауыл шаруашылық МТП) ретінде ұсынылады.

Соңғы жылдары шектеулі аумақта (аумақтық машина-трактор станциялары) бірнеше ауыл шаруашылығы тауар өндірушілеріне қызмет көрсететін машина-трактор станцияларын (МТРС) құру жағдайлары орын алды. Мұндай субъектілер кооперативтік негізде ауыл шаруашылығы аралық кәсіпорындар (кооперациялар) нысанында құрылуы мүмкін, бірақ әзірге олар көбінесе ауыл шаруашылығы өнімдерін өңдеу және сақтау функцияларын жүзеге асыратын кәсіпорындардың құрылымдық бөлімшелері түрінде құрылады. Бұл ретте мамандандырылған және күрделі машина-трактор станциялары ерекшеленеді. Бірінші типтегі машина-трактор станциялары не жеке жұмыс операцияларын орындауға, не жекелеген дақылдарды өсіруге және жинауға маманданған, екінші түрі – таңдалған технологияларды кешенді енгізуді қамтамасыз етеді.

МТС құрудың мақсаттылығын, олардың ауыл тауар өндірушілері үшін пайдалылығын анықтайтын басты нәрсе – экономикалық қатынастар механизмі. Ол МТС пен ауыл тауар өндірушілерінің экономикалық теңдігіне, станцияны пайдалана отырып өндірілген түпкілікті ауыл шаруашылығы өнімдерін (мысалы, астық) сатудан түскен табысты әділ бөлуге, баға диспропорциясын тегістеуге кепілдік беруі керек [2].

Бүгінгі күні қайта жаңғырып жатқан МТС-тің негізгі мақсаты – ауыл тауар өндірушілерімен бірлесіп, жоғары және қарқынды технологияларды пайдалана отырып, ауыл шаруашылығы өнімдерін өндіру бойынша жұмыстар кешенін өзара тиімді шарттарда жүзеге асыруды қамтамасыз ету, осы қызметтен барынша нәтиже алу, сондай-ақ елді отандық өндірістің ауыл шаруашылығы өнімдерінің барлық түрлерімен, ал мал шаруашылығын толық жеммен қамтамасыз ету үшін қолда бар өндірістік ресурстарды (табиғи, материалдық-техникалық, еңбек, қаржылық) пайдалану деңгейін арттыру туралы.

Осы мақсатқа жету үшін МТС келесі міндеттерді шешуі керек:

- өндірістік қызметте – қаражаттың, еңбектің, материалдардың қолайлы шығындары және адам ресурстарының жоғары уәждемесі кезінде жоғары өнімділік пен өнімділікке негізделген жоғары түпкілікті нәтижелермен ауыл шаруашылығы өндірісінің бүкіл технологиялық циклі бойынша өндірісті қамтамасыз ету;

- өндірістік-техникалық пайдалануда – механикаландырылған егістік кешеніне және озық технологияларды қолдануға және еңбекті ұйымдастыруға негізделген басқа да жұмыс түрлеріне шарттар бойынша орындау;

- техникалық пайдалануда – диагностикалау және жөндеу әдістері мен құралдарын пайдалана отырып, машиналарды барынша тиімді пайдалану, техникалық қызмет көрсету, ақауларды жою, жабдықты сақтау, техникалық қызмет көрсетудің жаңа технологияларын енгізу;

- нормативтік-технологиялық, консалтингтік және кадрлық процестерде – орындаушыларды нормативтік-техникалық құжаттамамен қамтамасыз ету, перспективті технологиялар мен машиналарды пайдалану бойынша кеңес беру, оңтайландырудың бірқатар мәселелерін шешу (қызмет бағытын таңдау, МКК құрамын анықтау, персоналды даярлау және олардың біліктілігін арттыру, диспетчерлік басқаруды енгізу және т.б.. Р.);

- МТС қызметінде ақпараттық технологияларды кеңінен қолдану [3].

Машина-технологиялық станциялардың алдында тұрған міндеттерді шешу ғылыми-техникалық прогрестің жетістіктерін барынша енгізуді қамтамасыз етеді; ауыл шаруашылығы техникасының жұмысқа дайындығын арттыру; ақпараттық қызмет көрсетуді жетілдіру, шаруашылықтарда машиналарды, жаңа технологияларды, машиналар мен жабдықтарды пайдаланудың озық тәжірибесін енгізу; механизаторлар мен мамандарды іріктеу, оқыту және олардың біліктілігін арттыру, олардың саны жұмыстың қызу кезеңдерінде қосымша механизаторларды тарта отырып, жыл бойына жұмысқа орналасумен анықталады; механизаторлар мен техникалық қызмет көрсетуші персоналдың өз қызметінің нәтижелеріне, еңбек заңнамасын сақтауға, экологиялық қауіпсіздікке материалдық қызығушылығын арттыру бойынша шараларды әзірлеу және

жүзеге асыру; машина-технологиялық станция ұжымының әлеуметтік даму жоспарларын орындау.

Қолданылған әдебиеттер тізімі.

1. Алтухов А.И. Проблемы развития АПК страны и необходимость их ускоренного решения // Экономика сельского хозяйства России. 2018. № 4. С. 2-14
2. Рогожина О. В. Развитие районных МТС / О. В. Рогожина // Техника и оборудование для села. – 2000. – №7. – С. 26–27; №8. – С. 19–21.
3. Черноиванов В. И. Машинно-технологическая станция. Организация, структура, виды работ, техника, нормативы, передовой опыт / В. И. Черноиванов, А. Э. Северный, В. М. Михлин. – 2-е изд., доп., перераб. – М.: ГОСНИТИ, 2003. – 332 с.

## АНАЛИЗ АТАК ИСТОЩЕНИЯ ЭНЕРГОРЕСУРСОВ В СИСТЕМАХ ИНТЕРНЕТА ВЕЩЕЙ

*Десницкий Василий Алексеевич*

*Федеральное государственное бюджетное учреждение науки*

*«Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН), Санкт-Петербург, Россия, [desnitsky@comsec.spb.ru](mailto:desnitsky@comsec.spb.ru)*

Все большее распространение получают различные прикладные информационно-телекоммуникационные системы, включающие в свой состав помимо традиционного коммуникационного и вычислительного оборудования общего назначения различные мобильные и встроенные устройства, сенсоры, исполнительные элементы систем интернета вещей (IoT). Функционирование систем интернета вещей охватывает как программно-информационные, так и физические аспекты взаимодействия устройств и пользователей, что наряду с критически важным характером подобных систем и возможностью удаленного доступа к устройствам через сеть Интернет обуславливает важность решения вопросов информационной безопасности таких систем.

Мобильный характер устройств и возможность их функционирования в автономном режиме, а также ограничения на имеющиеся энергоресурсы делают их подверженными атакам истощения энергоресурсов. В отношении различных приложений IoT – производственных систем, систем имплантируемых медицинских устройств, систем управления на транспорте и пр., последствиями нарушения доступности в них могут оказаться масштабные техногенные катастрофы и значительные материальные издержки.

Ключевыми особенностями атак истощения является сложность их обнаружения, потенциально высокая продуктивность и их вариативность. Сложность обнаружения атак истощения определяется, во-первых, тем, что воздействие на модуль является зачастую не прямым, а опосредованным – путем отправки извне через информационно-телекоммуникационную инфраструктуру последовательностей поддельных сообщений, которые трудно идентифицируемы в качестве злонамеренных. Во-вторых, для того чтобы отслеживать атаки истощения требуется анализировать не столько процесс разряда батареи устройства, сколько изменения скорости ее истощения. В-третьих, детектирование атак истощения может сопрягаться с объективными факторами разрядки батареи в результате легитимных операционных процессов и нормальной пользовательской активности. При этом для поражения устройства нарушитель может вариативно подбирать наиболее предпочтительные и эффективные шаги атаки с постепенным или стремительным исчерпанием энергоресурса.

В известных источниках литературы тематика атак истощения энергоресурсов отражена в недостаточно полной степени, причем, в основном анализируются конкретные, частные способы анализа лишь определенного вида атак истощения на системы определенного вида, как например, Denial-of-Sleep-атаки на сенсорные сети [1]. Все это, а также тенденция к недостаточной защищенности современных IoT-систем, и отсутствие, в частности, средств отслеживания атак истощения обуславливают необходимость дополнительных исследований по данному научному направлению. Стоит отметить, что архитектура устройств может не предполагать наличия технических возможностей для детального анализа энергопотребления устройства с приемлемой точностью, при этом форм-фактор и условия функционирования IoT-модулей могут препятствовать подключению внешних компонентов для считывания системных логов и сведений об энергопотреблении.

Вкладом настоящего исследования являются, во-первых, результаты анализа атак истощения энергоресурсов на устройства интернета вещей, полученные с учетом существующих работ по данной тематике, во-вторых, модель нарушителя, выполняющего атаки истощения, в-третьих, результаты экспериментов по моделированию и оценке атак

истощения на примере модулей беспроводных сетей. К элементам новизны проводимой работы относится ориентированность на формирование единой классификации атак истощения энергоресурсов в рамках комплексного подхода к их моделированию, анализу, экспериментальной оценке эффективности для конкретных видов IoT-сетей.

В [2] анализируются причины нецелевого использования батареи мобильных устройств. Предлагается метод обнаружения неисправного процесса и уменьшение потребления ресурсов батареи путем остановки или задержки подозрительных действий. Процесс отслеживания производится до восстановления штатного потребления заряда батареи. В [3-4] рассматриваются некоторые виды атак типа Denial-of-Sleep. В [3] предложен подход к созданию промежуточного уровня для эффективной защиты от атак истощения с использованием сетевого взаимодействия на стыке физического и канального уровней. В частности, анализируются следующие виды атак: Sleep Deprivation, Barrage, Replay, Broadcast, Collision и Synchronization атаки [3].

Отдельного внимания заслуживают работы, посвященные вопросам атак истощения в беспроводных сенсорных сетях. В [5] исследуются селективные атаки глушения, которые ориентированы на выборочное влияние на физическом уровне на наиболее критически важные пакеты, передаваемые по беспроводному каналу. Существенным компонентом такого воздействия является классификация пакетов на лету на физическом уровне соединения. Предложенный авторами алгоритм предотвращения такой атаки базируется на основе сокрытия пакетов, реализуемого между канальным и физическим уровнями сетевой модели. В [6] предложена ролевая интеллектуальная методика по выявлению атак глушения в беспроводной сенсорной сети, которая в случае обнаружения такого глушения позволяет учитывать данные от сенсоров и в качестве средств защиты проводить динамическую смену характеристик используемого сетевого канала связи.

В настоящей работе предложена обобщенная аналитическая модель нарушителя применительно к атакам истощения энергоресурсов, представляющая собой модель знаний, необходимых для моделирования и анализа таких воздействий. Модель представляется следующим формальным кортежем  $(G, O, A, R, C, F, E, P)$ .  $G$  – задает цели атак истощения, которые включают нарушение доступности некоторого устройства, работающего от автономного источника питания путем воздействия с постепенным равномерным, скачкообразным или максимально быстрым достижением данной цели в зависимости от специфики атакуемого устройства и мотивов атакующего.  $O$  задает объекты прямого и опосредованного воздействия атак, в том числе конкретные сенсоры физического окружения, аппаратные средства, коммуникационные каналы, программные компоненты, процессы операционной системы.  $A$  – шаги, которые нужно совершить атакующему для достижения целей  $G$ .  $R$  описывает системные ресурсы, оборудование и инструменты, используемые в процессе воздействия, а также стартовые ресурсы атакующего, знания и практические навыки, которыми он должен обладать. Кроме того  $R$  описывает временные затраты и материальную стоимость атаки для нарушителя.

$C$  специфицирует свойства атакующего в части возможностей осуществления им доступа и его расположения по отношению к атакуемому устройству.  $F$  описывает специфичную область или условия применимости атаки истощения, например, возможность при определенных условиях перехода целевого устройства в режим сна и обратно.  $E$  задает показатель эффективности атаки – усредненный коэффициент увеличения скорости разряда батареи при условии выполнения атаки с сравнении со штатным режимом работы устройства.  $P$  задает средства защиты от атаки истощения энергоресурсов.

На основе анализа опубликованной литературы в качестве наиболее значимых были выделены следующие четыре класса атак истощения энергоресурсов:

- 1) принудительный вывод устройств из сна – режима пониженного энергопотребления. Цель атаки – истощение батареи устройства для его отключения путем увеличения среднего энергопотребления устройства или его модулей;
- 2) увеличение объема беспроводного трафика, проходящего через устройство;
- 3) формирование электромагнитных помех на беспроводные каналы передачи данных, определяющих увеличение мощности электронного передатчика [5];
- 4) нештатное использование программного обеспечения устройства, в том числе множественный запуск приложений, неверные конфигурации, нарушение оптимизаций.

На фрагментах разработанных прототипов систем интернета вещей проведено моделирование и оценка некоторых видов атак истощения энергоресурсов. В частности, смоделирована атака принудительного перевода мобильного устройства в более энергозатратный режим путем эксплуатации его беспроводного коммуникационного интерфейса. Была смоделирована атака на модуль Bluetooth смартфона с использованием средств операционной системы Kali Linux. Схема атаки основана на выполнении утилит Bluetoothctl, используемой для обнаружения атакуемого устройства и получения информации о нём, и l2ping, при помощи которой выполняется последовательная отправка ping-запросов.

В эксперименте измерения энергопотребления проводились при нормальном функционировании устройства и при атаке. В обоих случаях измерения осуществлялись сразу после предваряющей зарядки смартфона и достижения полного заряда. Период измерений в обоих случаях составлял 4 часа. Во время эксперимента были активны модули сотовой связи и Bluetooth. Экран, а также другие коммуникационные модули были выключены. В процессе проведения эксперимента никаких сторонних манипуляций с устройством и обновления операционной системы и приложений не проводилось. Данные по расходу батареи получались программным способом путем чтения значений Android-переменной BatteryManager.EXTRA\_LEVEL, фиксирующей текущий уровень заряда.

Также была продемонстрирована комбинированная атака на оконечный узел самоорганизующейся XВее сети, функционирующий в режиме энергосбережения с запланированной периодичностью перехода из режима сна в режим полноценной работы. Атака моделировалась путем отправки запросов со стороны ложного XВее-узла. В условиях эксперимента на целевом узле были заданы следующие настройки, определяющие параметры энергосбережения:  $SM = 4$  (режим циклического сна),  $ST = 1000$  мс (время до сна) и  $SP = 10000$  мс (продолжительность цикла сна коммуникационного модуля). Действия нарушителя, эксплуатирующего ложный XВее узел, состоят в периодичной отправке на атакуемый узел сообщений (2 шт. в секунду), которые будут обнулять таймер  $ST$ , не давая узлу перейти в состояние сна. Комбинированный характер данной атаки состоит в том, что помимо препятствования переходу атакуемого узла в режим сна, дополнительное энергопотребление осуществляется также в результате приема данных вместо нахождения в режиме простоя. Эксперимент проводился циклично на интервале времени  $t_2 - t_1 = 600$  с. Значение показателя эффективности атаки было получено равным  $E = 4.488$ , что означает, что такая атака способна более чем в 4 раза быстрее исчерпать оценочное время работы батареи XВее-узла в условиях заданных параметров режима сна. Точность датчика тока, который использовался в экспериментах, составляет  $\pm 0.5$  мА.

Отметим, что по большей части атаки истощения выполнимы применительно к аппаратным модулям, использующим автономные, исчерпаемые источники энергии. Вариативность атак данного вида ограничивается разнообразием способов воздействий на отдельные наиболее энергозатратные компоненты устройства с целью нелегитимного увеличения интенсивности их использования. При этом проведение атаки истощения нарушителем имеет смысл только в случае фактора важности работоспособности устройства, на которое воздействие нацелено. Так, если допустима краткосрочная приостановка функционирования устройства, или оно оперативно может быть замещено,



то такое воздействие может не оказать значительного эффекта. В целом проблематика атак истощения энергоресурсов видится крайне актуальной в различных практических областях приложения. В случае малых радиоуправляемых беспилотных летательных аппаратов (дронов), эксплуатация атак истощения позволяет нарушительно воздействовать на дрон во время его полета и исчерпать заряд, затрачиваемый, в том числе и на функционирование моторов. При этом скрытый характер таких воздействий увеличивает вероятность того, что владелец дрона не будет обладать достаточным временем для того, чтобы вовремя аварийно приземлить дрон, и он может разбиться.

Проведенные в работе эксперименты по моделированию атак для различных устройств показывают потенциал таких атак по снижению времени жизни батареи устройств в несколько раз, что оказывается достаточно важным для критически-важных систем. Обнаружение атак истощения представляется во многом процессом, специфичным конкретным практическим областям приложения. Поэтому комплексный подход к детектированию таких атак должен охватывать детальное отслеживание процесса энергопотребления, а также анализ логов перехода устройств и их аппаратных модулей между различными режимами их работы, логов запуска и приостановки программных приложений, обращений к модулям хранения. Все эти данные необходимо использовать для формирования потоков инцидентов и верификации правил корреляции с целью выявления конкретных признаков атак истощения энергоресурсов.

#### Список использованной литературы

1. Balueva A., Desnitsky V., Ushakov I. Approach to Detection of Denial-of-Sleep Attacks in Wireless Sensor Networks on the Base of Machine Learning // 13th International Symposium on Intelligent Distributed Computing (IDC 2019), Saint-Petersburg, Russia, 7-9 October 2019. Intelligent distributed computing XIII p. 350-355, 2019.
2. Shin S., Lee T., In. H.P. Defending Battery Exhaustion Attacks on Mobile Systems // In: Proceedings of 33rd Annual IEEE International Computer Software and Applications Conference, p. 347-352, 2009.
3. Boubiche D.E., Bilami A. A Defense Strategy against Energy Exhausting Attacks in Wire-less Sensor Networks // Journal of Emerging Technologies in Web Intelligence, vol. 5, no. 1, 2013.
4. Krishnan M. Intrusion Detection in Wireless Sensor Networks // <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.89.3793&rep=rep1&type=pdf>.
5. Karpagam R., Archana P. Prevention of Selective Jamming Attacks Using Swarm intelligence Packet-Hiding Methods // International Journal Of Engineering And Computer Science, vol. 2, p. 2774-2778, 2013.
6. Periyamayagi S., Sumathy V., Kulandaivel R. A Defense Technique for Jamming Attacks in Wireless Sensor Networks Based on Sensor Networks // International Conference on Process Automation, Control and Computing, p. 1-5, 2011.

## ВОПРОСЫ ПРИМЕНЕНИЯ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ В БАЗЕ ДАННЫХ ОБУЧАЮЩИМИСЯ ВУЗОВ

*Серік Меруерт<sup>1</sup>, С.К. Жумагулова<sup>1</sup>, А.В.Копыльцов<sup>2</sup>*

*<sup>1</sup>ЕНУ им.Л.Н.Гумилева, Астана, saulesha\_81@mail.ru*

*<sup>2</sup>Санкт-Петербургский государственный университет аэрокосмического приборостроения, г.Санкт-Петербург*

Сегодня повсеместно можно наблюдать активное развитие сферы программного и аппаратного обеспечения. Практически во всех сферах деятельности общества происходит внедрение многопроцессорных вычислительных систем. При этом регулярно совершенствуются методы решения задач, а также технологии программирования.

Особо сильно повлияли возникшие в процессе развития общества изменения на область информатики. На сегодня в результате ее активного развития выросла роль параллельной обработки данных любой компьютерной системы. Постоянный рост информации привело к тому, что возникает необходимость универсальных инструментов и решений, охватывающих помимо сферы информатики, также остальные направления. Одним из наиболее актуальных тем является параллельные вычисления в базах данных. Они позволяют получить результаты, которые невозможно получить по иным технологиям, а также составить модели, расчеты и генерировать данные.

В целом, параллельные вычисления представляют собой не только реализацию больших объемных вычислений, но и также параллельную обработку большого объема данных за небольшой интервал времени. Применение параллельных вычислительных систем служит одним из ключевых направлений развития аппаратного обеспечения, что определено ограничением максимально возможного быстродействия последовательных машин. Кроме того, всегда имеется необходимость в решении задач, где недостаточно возможностей имеющихся средств аппаратного обеспечения [1].

Параллельные вычисления (одновременная реализация нескольких операций обработки данных) выполняются чаще всего путем введением избыточности функциональных устройств. Здесь ускорения процесса решения вычислительной задачи можно достичь при разделении используемого алгоритма на информационно независимые части, каждая из которых реализуется на разных процессорах. Путем применения такого подхода возможна минимизация затрачиваемого на вычисления времени, причем эта возможность ограничивается лишь числом имеющихся процессоров и реализуемых на них частей вычислений.

Сегодня однако параллелизм применяется довольно редко. Это можно объяснить тем, что стоимость высокопроизводительных систем до сих пор остается достаточно высокой. Параллельные вычисления в базах данных играют ключевую роль в научно-методических исследованиях. Это указывает на потребность внедрения в образовательный процесс вуза определенной методики обучения параллельным вычислениям с целью подготовки будущих программистов. Также важно отметить, что параллельное программирование и параллельные вычисления постепенно становятся важными частями в содержании дисциплин информатики и программирования [2]. Технология параллельного программирования в существенной степени изменяет алгоритмическую деятельность ИТ-специалиста, вносит в нее новые действия и этапы. Кроме способов построения программы, программирование посредством параллельной технологии формирует параллельный стиль мышления человека, а также его способность параллельно обрабатывать данные [3].

Программирование баз данных является значительной частью системы предметной подготовки специалистов информационного профиля. Не обеспечив изучение современных парадигм и технологий программирования, включая базы данных, невозможно полноценно подготовить таких специалистов. В результате изучения новой

технологии параллельных вычислений в системе подготовки будущих специалистов информационного профиля возникают вопросы, касающиеся специфики методики преподавания параллельных вычислений в базе данных. Как показал опыт обучения студентов основам параллельного программирования в базе данных, сегодня можно говорить о недостаточности методической и педагогической проработки данной проблемы в целом, а также для высших учебных заведений в частности.

Таким образом, сложность организации взаимодействия параллельных процессов в базе данных, а также возникающие при этом информационные процессы определяют целесообразность выбора информационного подхода к обучению в качестве основы для создания методической системы обучения будущих специалистов информационного профиля параллельным вычислениям в базе данных [4].

Формирование и развитие параллельного стиля мышления, а также успешность усвоения материала в сфере вычислительных технологий у обучающихся будут обеспечены при условии определения стадии формирования параллельных вычислений на базе информационного и деятельностного подходов, анализа программно-аппаратных средств реализации параллельных вычислений в базе данных, а также разработки методики обучения реализации параллельных вычислений в базе данных в вузах.

#### Список использованной литературы

1. Лупин С.А., Посыпкин М.А. Технологии параллельного программирования. Серия: Высшее образование. - М.: Форум, Инфра-М, 2008. - 208 с.
2. Миллер Р., Боксер Л. Последовательные и параллельные алгоритмы: Общий подход. - М.: БИНОМ. Лаборатория знаний, 2006. - 406 с.
3. Носов М.Т., Ерахтин В.М. Параллельные вычисления. Основы параллельного программирования. М.: Форум, Инфра-М, 2011. - 277 с.
4. Хаблецкий И.В., Веренюк С.Т. Параллельные алгоритмы решения вычислительных задач. - СПб.: БХВ-Петербург, 2009. – 298 с.

## КОМПАНИЯДАҒЫ ТИІМДІ ІШКІ КОММУНИКАЦИЯ ҮШІН ІТ ИНФРАҚҰРЫЛЫМЫН ОҢТАЙЛЫ ТАҢДАУ

*Сағатова А.С.<sup>1</sup>, Алимжанова Л.М.<sup>2</sup>, Сарбасова А.К.<sup>3</sup>*

<sup>1</sup>*Әл-Фараби атындағы ҚазҰУ, Алматы, sagatovaidana13@gmail.com*

<sup>2</sup>*Әл-Фараби атындағы ҚазҰУ, Алматы, dimkim\_01@mail.ru*

<sup>3</sup>*Әл-Фараби атындағы ҚазҰУ, Алматы, aluasar@mail.ru*

**Абстракт.** Зерттеу ішкі цифрлық инфрақұрылымды дамыту аспектілерін және оны ұйымдарда қалыптастыру мәселелерін қарастырады. Компанияның цифрлық инфрақұрылымын қалыптастыру мен дамытудың маңызды алғышарты ретінде тұрақты және дамыған технологиялық инфрақұрылымды қалыптастыру қажеттілігі негізделген. Цифрлық инфрақұрылымды дамытуды қамтамасыз ету процесін жетілдіру үшін көпфункционалды есептеу ресурстары ұсынылған.

**Кілт сөздер:** ІТ, инфрақұрылым, цифрлық, кәсіпорын, мәліметтер базасы, ақпараттық жүйе.

**Кіріспе.** Қазіргі уақытта көп нәрсе сандық инфрақұрылымға байланысты: бизнес қаншалықты сәтті, ол қанша табыс әкеледі, ұйым қаншалықты тиімді жұмыс істейді. Сондықтан бұл жүйе сенімді, тұрақты және қауіпсіз болуы керек. Бұл ұйымның ең маңызды активі деп айта аламыз. Персоналды басқару жүйелері, клиенттерге қызмет көрсету процестері және нарықты талдау міндеттері үнемі күрделене түседі және ақпараттық технологиялардың көмегінсіз жүзеге асырыла алмайды. Сондықтан бизнестің жетістігі көбінесе компанияның ІТ инфрақұрылымының қаншалықты жақсы қалыптасқанына және оның қаншалықты жақсы жұмыс істейтініне байланысты. Сондай-ақ, жүйе өзіне жүктелген тапсырмаларды орындай алмайтын немесе деректердің қауіпсіздігін қамтамасыз ете алмайтын жағдайларда өте күрделі мәселелер туындайтыны сөзсіз. Компанияның ішкі цифрлық инфрақұрылымы әртүрлі ақпараттық ресурстарды біріктіреді, онсыз ұйым дұрыс жұмыс істей алмайды және қызметкерлер өз жұмысын тиімді орындай алмайды. Сондықтан ұйымдағы бөлімдер мен әртүрлі модульдер арасындағы сенімді байланыс құралын қамтамасыз ету маңызды. Ішкі цифрлық инфрақұрылымдардың тұрақты дамуын қамтамасыз ету компанияның бизнестегі технологиялық мүмкіндіктерін кеңейтеді, осылайша олардың бәсекеге қабілеттілігін арттырады.

Зерттеу ішкі цифрлық инфрақұрылымды дамыту аспектілерін және оны ұйымдарда қалыптастыру мәселелерін қарастырады. Компанияның цифрлық инфрақұрылымын қалыптастыру мен дамытудың маңызды алғышарты ретінде тұрақты және дамыған технологиялық инфрақұрылымды қалыптастыру қажеттілігі негізделген. Цифрлық инфрақұрылымды дамытуды қамтамасыз ету процесін жетілдіру үшін көпфункционалды есептеу ресурстары ұсынылған.

### **Негізгі бөлім**

Ішкі ІТ инфрақұрылымы үшін оңтайлы бағдарламалық жасақтама қандай болуы керек?

Зерттеу материалын зерттеу кезінде біз бүгінгі күнге дейін қолданылған және қолданылған әдістерді іздеу және талдау арқылы өзіміздің жаңа шешімімізді қарастырамыз. Яғни, қолданыстағы жүйелерді қарастырып, оларды өзіңізге бейімдей аласыз. Алынған мәліметтер мен зерттеулер веб-қызмет арқылы ұйымдастырылған.

ІТ-инфрақұрылым — бұл ақпараттық-коммуникациялық инфрақұрылым: телекоммуникациялық және ақпараттық желілердің, соңғы құрылғылардың, ақпараттық ресурстардың жиынтығы, кез-келген уақытта және кез-келген жерде пайдаланушылар арасында байланыс орнатуға және кез-келген ақпаратқа қол жеткізуге болады [1]. Тұрақты цифрлық инфрақұрылымды дамыту қоғамға пайда әкелетін технологиялық прогресті қамтамасыз ету үшін өте маңызды [1].

Қызметке бағытталған архитектураны (SOA) енгізу қазіргі уақытта компанияларға икемділік пен тиімділіктің жоғарылауы, қолда бар ресурстарды тиімді пайдалану және икемділік сияқты маңызды артықшылықтарды алуға мүмкіндік беретін танымал тәсіл болып табылады және осылайша нарықтағы өзгерістерге жауап берудің жақсы мүмкіндіктеріне ие. Алайда, осы күрделі шешімдерді әзірлеу және қолдау үшін қажетті адами және техникалық ресурстар кең таралған және компанияларға да, қоршаған ортаға да пайда әкелмейді. Қызметтерге назар аудару SOA-ны бірегей етеді және көптеген ескірген жүйелер мен қара корпустағы деректер көздерінде ашықтықты қамтамасыз етеді [2, 3].

Көп жағдайда ұйымдар өздерінің мұрагерлік жүйелерін сақтағысы келеді, өйткені олар бизнес үшін әлі де құнды және белгілі бір бизнес функциялары мен процестерін қолдайды; мұрагерлік жүйесі тек технологиялық артефакт емес — бұл IT мамандары үшін де, соңғы пайдаланушылар үшін де маңызды ұйымдастырушылық мәдениеттің бөлігі. SOA мұрагерлік жүйелерді үздіксіз пайдалануды жеңілдетеді алады және цифрлық инфрақұрылым қолдайтын шешімдер қабылдауды жеңілдету үшін бизнес-аналитикалық қосымшаларды қосу арқылы осы жүйелерді тиімді пайдалануға ықпал ете алады [4]. Цифрлық инфрақұрылым жаңа цифрлық технологиялардың әлеуетін кеңейтуді қалайтын ұйымдардың цифрлық трансформациясын қамтамасыз етудің маңызды негізі болып табылады. Сандық инфрақұрылым техникалық және ұйымдастырушылық компоненттерден, процестерден және желілерден тұрады. Ол цифрлық құралдарды пайдаланушылардың, сондай-ақ инфрақұрылымға қосылған жүйелерді жобалаушылар мен әзірлеушілердің әлеуметтік ортасын қамтиды [5].

Бұл зерттеуде эмпирикалық талдау әдісі, атап айтқанда салыстырмалы талдау қолданылды. Ұсынылған гипотезаны негіздеу үшін әртүрлі корпоративтік бағдарламалық жасақтамаға салыстырмалы талдау жүргізілді. Бұл нысандар әртүрлі функциялардан тұратын корпоративтік порталдың бағдарламалық жасақтамасы болып табылады. Мысалы, бұлтты есептеу, деректерді сақтау, қауіпсіздік, сақтық көшірме жасау, құжаттарды басқару, іздеу жүйесі, корпоративтік байланыс, ынтымақтастық және басқалар.

*1 кезең.*

Зерттеу үшін sourceforge.com сайтынан жоғары рейтингі бар 6 корпоративтік бағдарламалық қамтамасыз етуші порталдар таңдалды. Кесте сайттың тәуелсіз пайдаланушыларының рейтингтері мен шолуларына негізделген. Бағалаудың жалпы нәтижесі көрсетіледі (1 кесте).

*2 кезең.*

Зерттеудің екінші кезеңінде техникалық сипаттамалардан тұратын кесте жасалды. Өндіруші ел, компания, құрылған жылы, өнімнің бастапқы бағасы, интеграцияланған өнімдер саны, бағдарлама тілдері сияқты мәліметтер келтірілген (2-кесте).

*3 кезең.*

Зерттеудің үшінші кезеңінде функционалдық сипаттамалар қарастырылды. Жүйелік талаптар, API және мобильді қосымшаның қол жетімділігі, мазмұнды басқару, іздеу жүйесі, ішкі байланыс, ынтымақтастық, құжаттарды басқару және файлдарды ортақ пайдалану сияқты мүмкіндіктеріне салыстырмалы талдау жасалды (3-кесте).

Зерттеу нәтижесінде 6 корпоративті бағдарлама бойынша 3 кесте құрылды. Үш салыстырмалы кестені қорытындылай келе, барлық жүйелер көптеген танымал веб-порттармен және портативті платформалармен үйлесімді екендігі анықталды, деректер қоймасы, мобильді қосымшалар және ішкі байланыс, ынтымақтастық және деректермен бөлісу құралдары бар. Функционалды сипаттамаларға сүйене отырып, ең оңтайлы корпоративтік платформалар OnBoard Board Management Software және Ntranet бағдарламалық жасақтамасы. Себебі бұл бағдарламалар пайдаланушылардан ең жоғары баға алды және кең функционалдылыққа ие болды. Алайда, басқа қосымшалар да аса кем түспейді.

## Кестелер

### 1-кесте. Рейтингтер/Пікірлер [6]

Корпоративтік порталдың атауы	Support	Design	Features	Ease	Overall
Intrexx	4.6/5	4.3/5	4.5/5	4.6/5	4.7/5
SmartVault	4.4/5	4.4/5	4.6/5	4.6/5	4.7/5
Bitrix24	4.7/5	4.5/5	4.4/5	4.3/5	4.7/5
OnBoard Board Management Software	4.9/5	4.5/5	4.7/5	4.8/5	4.9/5
Yammer	4.4/5	4.3/5	4.3/5	4.7/5	4.3/5
Ntranet	5.0/5	5.0/5	5.0/5	4.0/5	5.0/5

### 2-кесте. Техникалық сипаттамалары [6]

Корпоративтік порталдың атауы	Компания	Шығарушы ел	Құрылды	Бағасы	Интеграциялар	Тілдер	Веб-сайт
Intrexx	United Planet GmbH	Germany	1998	Тегін тест нұсқасы қол жетімді.	10 API	Ағылшын, Француз, Неміс.	www.intrexx.com
SmartVault	SmartVault	United States	2008	Бастапқы баға: Айына 20,00 / пайдаланушы	24 API	Ағылшын,	smartvault.com
Bitrix24	Bitrix24	United States	1998	Бастапқы баға: Айына 24 доллар	96 API	Ағылшын,	www.bitrix24.com
OnBoard Board Management Software	OnBoard	United States	2003	Тегін сынақ нұсқасы қол жетімді.	14 API	Ағылшын,	www.onboardmeetings.com
Yammer	Microsoft	United States	1975	Бастапқы баға: Бір пайдаланушыға айына 3 доллар	143 API	Ағылшын,	www.yammer.com
Ntranet	Ncontracts	United States	2009	-	5 API	Ағылшын,	www.ncontracts.com

### 3-кесте. Функционалдық сипаттамалары [6]

Корпоративтік порталдың атауы	Жүйелік талаптар	API	мобильді қосымша	Бірлескен жұмыс	Чат	Контент басқару	Құжаттарды басқару	Файлдармен бөлісу	Іздеу
Intrexx	SaaS, Windows, Mac, Linux, iPhone, iPad, Android	+	+	+	+	+	+	+	+
SmartVault	SaaS, Windows, Mac, Linux, iPhone, iPad, Android	+	+	-	-	-	+	+	+
Bitrix24	SaaS, Windows, Mac, iPhone, iPad, Android, On-Premises	+	+	+	-	+	+	+	+

OnBoard Board Management Software	SaaS, Windows, Mac, iPhone, iPad, Android		+	+	+	+	+	+	+
Yammer	SaaS, Windows, iPhone, iPad, Android	+	+	-	-	-	-	+	+
Ntranet	SaaS, Windows, Mac, iPhone, iPad, Android	+	+	+	+	+	+	+	+

Осы зерттеу барысында рейтингтер мен сипаттамаларды салыстыра отырып, салыстырмалы талдау үшін қол жетімді алты үздік платформа таңдалды. Ішкі корпоративтік порталдың оңтайлы бағдарламасы үшін компоненттер мен функционалдық талаптардың тізімі анықталды.

Бірлескен жұмыс - қызметкерлерге командада, қашықтан жұмыс істеуге және жұмыс нәтижелерін біріктіруге мүмкіндік береді. Чат - бөлімдер мен модульдер арасында іскерлік хат алмасуға, қабылданған шешімдерді үйлестіруге көмектеседі. Құжаттарды басқару жүйесі дерекқорда сақталған деректерді басқаруға, сақтауға және ортақ пайдалануға мүмкіндік береді. Іздеу жүйесі қолданушыға бүкіл компания бойынша қажетті ақпаратты табуға көмектеседі. Бағдарламалық жасақтаманы қолдаудың жетілдірілген мүмкіндіктері әртүрлі бағдарламалық өнімдерді пайдаланатын пайдаланушыларды тартуға мүмкіндік береді.

### Қорытынды

Кәсіпорынның корпоративті порталдары фирмаларға өндірісті жоспарлау және адам ресурстарын басқару сияқты басқа қызмет түрлерімен біріктірілген бизнестранзакцияларды өңдеудің модельдерін ұсынады. Компанияның стандартты процестерін енгізу және оның барлық қызметі мен барлық қызметкерлерін қамтитын бірыңғай мәліметтер базасын ұсыну арқылы корпоративтік портал бағдарламалары оның көптеген бөлінген бөлімшелері мен функционалды салаларын біріктіруді қамтамасыз етеді. Нәтижесінде, бұл жүйелер штаттың қысқаруы, іскерлік шешім қабылдау процесін жеделдету, деректердің бұзылуын бақылау және т. б. сияқты көптеген көрсеткіштерде көрінетін шешім қабылдау мүмкіндіктерін жақсартуға әкелді.. Осылайша, корпоративтік порталдарды фирмаларға еңбек өнімділігін арттыруға көмектесу үшін пайдалануға болады.

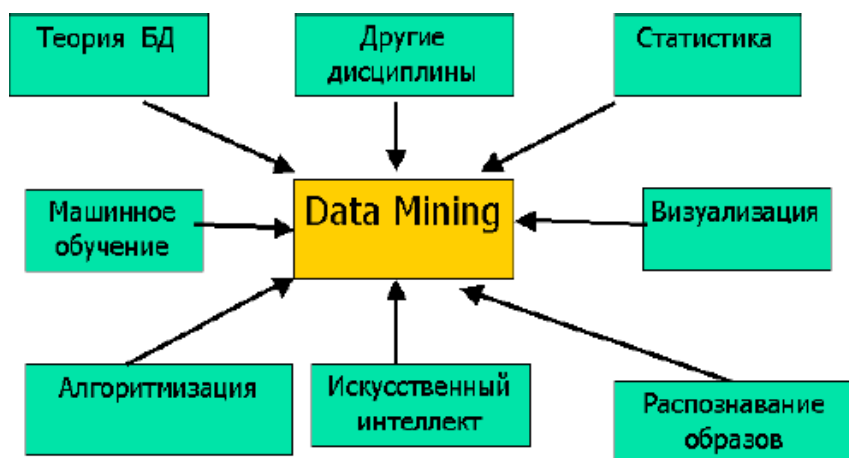
### Пайдаланылған әдебиеттер тізімі

1. Hirschheim R., R. Welke and A. Schwarz (2010) "Service - Oriented Architecture: Myths, Realities, and a Maturity Model." MIS Quarterly Executive 9(1): 37–48.
2. Dedeke A. (2012) "Improving legacy-system sustainability: A systematic approach." IT Professional 14(1): 38–43.
3. Leonardi P. M., M. Huysman and C. Steinfield (2013) "Enterprise social media: Definition, history, and prospects for the study of social technologies in organizations." Journal of Computer- Mediated Communication 19(1): 1-19.
4. Hustad E., and L. Staverløkk (2013) Implementing a Service-Oriented Architecture: A Technochange Approach. Information Systems Development, Springer: 527-538.
5. Eli Hustad Dag H.Olsen "Creating a sustainable digital infrastructure: The role of service-oriented architecture". Procedia Computer Science. Volume 181, 2021, pages 597-6.
6. [https://sourceforge.net/software/portal/for-enterprise/?sort=rating\\_count](https://sourceforge.net/software/portal/for-enterprise/?sort=rating_count)

## ДЕРЕКТЕРДІ ЗИЯТКЕРЛІК ТАЛДАУ (DATA MINING)

Некесова Анаргуль Аймуратовна  
Л.Н. Гумилев атындағы Еуразия ұлттық университеті,  
«Ақпараттық жүйелер» кафедрасының аға оқытушысы, Астана қаласы,  
[aimurat\\_anara@mail.ru](mailto:aimurat_anara@mail.ru)

Бүгінгі таңда нарықта көптеген IT құралдары бар, соның ішінде көптеген компаниялар үшін Data Mining табысты бизнеске айналдыратын әртүрлі әдістері қарастырылған. Data Mining – қолданбалы статистика, үлгіні тану, жасанды интеллект, деректер қоры теориясы және т.б. сияқты ғылымдардың негізінде пайда болған және дамып келе жатқан көппәнді сала (Сурет. 1 қараңыз).



Сурет.1 Data Mining көппәнді сала

Статистика - зерттелетін құбылысқа тән заңдылықтарды анықтау үшін деректерді жинау, өңдеу және талдау әдістері туралы ғылым.

Статистика – экспериментті жоспарлау, мәліметтерді жинау, оларды ұсыну және қорытындылау, сондай-ақ осы мәліметтер негізінде талдау және қорытынды жасау әдістерінің жиынтығы. Статистика бақылаулар немесе эксперименттер нәтижесінде алынған мәліметтерге негізделген.

Машиналық оқытуды бағдарлама арқылы жаңа білімді алу процесі ретінде сипаттауға болады. Машиналық оқыту алгоритмінің мысалы нейрондық желілер болып табылады. Жасанды интеллект – бұл ғылыми бағыт, оның шеңберінде интеллектуалды деп саналатын адам әрекетін аппараттық немесе бағдарламалық модельдеу міндеттері қойылған және шешілетін. Жасанды интеллект дәстүрлі түрде адамның артықшылығы болып саналатын шығармашылық функцияларды орындау үшін интеллектуалды жүйелердің қасиеті болып табылады.

Data Mining - бұл деректердегі жасырын үлгілерді табуға негізделген шешімдерді қолдау процесі. Data Mining технологиясы - бұрын белгісіз, тривиальды емес, практикалық пайдалы және адам қызметінің әртүрлі салаларында шешім қабылдау үшін қажетті білімді түсіндіру үшін қолжетімді бастапқы деректерді табу процесі. Data Mining технологиясы деректердің үлкен көлемдерінде айқын емес, объективті және практикалық үлгілерді іздеуге арналған. Айқын емес - бұл табылған үлгілер ақпаратты өндеудің стандартты әдістерімен немесе сарапшылармен анықталмағанын білдіреді.

Объективті - бұл әрқашан субъективті болып табылатын сараптамалық пікірден айырмашылығы, ашылған заңдылықтардың шындыққа толығымен сәйкес келетінін білдіреді. Практикалық пайдалы – бұл тұжырымдардың нақты мағынаға ие екендігін білдіреді, оны тәжірибеде қолдануға болады. Data Mining технологиясы адамға түсінікті



пішінде көрсетілуі мүмкін деректердің ішкі үлгілеріне тән үлгілер болып табылатын үлгілер (үлгілер) тұжырымдамасына негізделген.

Үлгілерді іздеудің мақсаты - қажетті процестерді көрсететін пішінде деректерді ұсыну. Болжамдық үлгілерді құру да үлгілерді табу мақсаты болып табылады. Деректерді өңдеу құралдары статистикалық құралдардан айырмашылығы ретроспективті деректердің қатаң анықталған көлемін қажет етпейді. Бұл мүмкіндік сенімсіз, жалған үлгілерді анықтауға және нәтижесінде олардың негізінде дұрыс емес шешімдер қабылдауға себеп болуы мүмкін. Ашылған білімнің статистикалық маңыздылығын бақылау қажет. Кең мағынада деректер фактілер, мәтін, графика, суреттер, дыбыстар, аналогтық немесе цифрлық бейне сегменттері болып табылады. Мәліметтерді өлшеулер, тәжірибелер, арифметикалық және логикалық операциялар нәтижесінде алуға болады.

Мәліметтер сақтауға, жіберуге және өңдеуге жарамды пішінде ұсынылуы керек. Басқаша айтқанда, деректер деректер жеткізушілері беретін және тұтынушылар деректерден ақпаратты қалыптастыру үшін пайдаланатын шикізат болып табылады. Data Mining негізгі ерекшелігі кең математикалық құралдар жинағының (классикалық статистикалық талдаудан жаңа кибернетикалық әдістерге дейін) және ақпараттық технологияның соңғы жетістіктерінің үйлесімі болып табылады. Data Mining технологиясы бейресми талдаудың қатаң формалды әдістері мен әдістерін, яғни деректерді сандық және сапалық талдауды үйлесімді біріктіреді.

Data Mining әдістері мен алгоритмдеріне мыналар жатады: жасанды нейрондық желілер, шешім ағаштары, символдық ережелер, ең жақын көрші және k-ең жақын көрші әдістері, тірек векторлық машина, Байес желілері, сызықтық регрессия, корреляциялық-регрессиялық талдау; кластерлік талдаудың иерархиялық және иерархиялық емес әдістері, ассоциация ережелерін іздеу әдістері; шектеулі іздеу әдісі, эволюциялық бағдарламалау және генетикалық алгоритмдер, деректерді визуализациялаудың әртүрлі әдістері және т.б.

Data Mining технологиясында қолданылатын аналитикалық әдістердің көпшілігі белгілі математикалық алгоритмдер мен әдістер болып табылады. Аппараттық және бағдарламалық қамтамасыз етудің пайда болу мүмкіндіктеріне байланысты оларды белгілі бір нақты мәселелерді шешуде пайдалану мүмкіндігі олардың қолдануындағы жаңалық болып табылады.

Айта кету керек, Data Mining әдістерінің көпшілігі жасанды интеллект теориясы аясында әзірленген.

Әдіс (әдіс) – бұл теориялық, практикалық, танымдық, басқарушылық сипаттағы мәселені шешудің белгілі бір жолы, әдісі, шешімі норма немесе ереже.

Алгоритм (алгоритм) – бастапқы деректерді қажетті нәтижеге түрлендіретін әрекеттердің (қадамдардың) реттілігіне қатысты нақты нұсқау.

Алгоритм ұғымы электронды есептеуіш машиналарды жасаудан көп бұрын пайда болды. Қазір алгоритмдер адам қызметінің әртүрлі салаларындағы көптеген қолданбалы және теориялық есептерді шешудің негізі болып табылады, олардың көпшілігі компьютердің көмегімен шешілетін тапсырмалар болып табылады.

Заманауи аналитикалық құралдар, соның ішінде Data Mining, жоғары сапалы визуализациясыз мүмкін емес. Көрнекі құралдарды қолдану нәтижесінде әртүрлі құралдарды қолдану арқылы көрнекі және мәнерлі, анық және қарапайым бейнелер алынуы керек: түс, контраст, жиектер, пропорциялар, масштаб және т.б.

Визуализация құралдарына қойылатын талаптардың өсуіне, сондай-ақ оларды бір-бірімен салыстыру қажеттілігіне байланысты соңғы жылдары ақпаратты жоғары сапалы көрнекі түрде көрсетудің бірқатар принциптері қалыптасты.

Рекрутингтік агенттіктердің бірінің нақты бизнес мәселесін шешуге мысал келтірейік, онда түйіндемелерді талдау үшін Data Mining технологиясы қолданылды. Бұл агенттік ІТ-компаниялар үшін кадрларды іріктеуге маманданған, жұмыс барысында ол әртүрлі бос орындарға үміткерлердің бірнеше мың түйіндемелерінің дерекқорын жинақтай алды.

Біріншіден, барлық түйіндемелердің әртүрлі форматындағы мәселе шешілді, ал әзірленген бағалау стандарты жаңадан жіберілген түйіндемелерді қабылдау кезінде бірден белгілеуге мүмкіндік берді. Осылайша, компанияның тағы бір міндеті шешілді - жинақталған ақпаратты жүйелеудің тиімді стандартын құру. Мұндай жүйе агенттікке Data Mining көмегімен түйіндемелерді талдау үшін ғана емес, сонымен қатар мәліметтер базасында тиімдірек іздеу, статистикалық өңдеу және т.б. үшін қажет. Түйіндемеге белгі қойғаннан кейін компания мамандары деректерді талдауға дайындауға кірісті. Келесі қадам арнайы әзірленген Data Mining құралдар жинағы арқылы тікелей деректерді талдауды жүргізу болып табылады.

Талдау нәтижесінде адал қызметкерлердің және жылына бір реттен жиі жұмысын өзгертуге бейім адамдардың егжей-тегжейлі профильдерін (портреттерін) құру мүмкін болды; әртүрлі жастағы және басқа әлеуметтік-демографиялық топтардың профильдері, әртүрлі бөлімшелер қызметкерлерінің, түлектер, түрлі университеттер және тағы басқалар салынды. Мәселен, 20-25 жас аралығындағы әйелдер жұмыс орнын үнемі ауыстыруға бейім екені анықталды. Маркетинг бөлімінің қызметкерлері де жұмыс орнын жиі ауыстырумен ерекшеленеді.

Агенттік жұмысқа қабылдау кезінде осыдан қандай қорытынды шығарады? Клиент үшін алты айдан кейін жұмыстан кетпейтін адал және берілген қызметкерді табу маңызды болса, агенттік іздеуді 35-45 жас аралығындағы ер адамдарға бағыттайды. Егер клиентке уақытша жұмысқа тез арада адамды табу маңызды болса, агенттік оған 20-25 жас аралығындағы қызды ұсына алады.

Осылайша, Data Mining технологиясының арқасында агенттік үміткерлерді іздеуді алдын ала қысқарта алады және, демек, жұмысқа қабылдауды тиімдірек - жылдамырақ және аз шығындармен жүргізе алады.

Нәтижесінде, Data Mining құралын және оны жүзеге асыру әдісін таңдау нақты мақсаттар мен міндеттерге сәйкес жүзеге асырылуы керек, компанияның қаржылық мүмкіндіктерінің деңгейін, пайдаланушылардың біліктілігін және басқа да бірқатар талаптарды ескере отырып жүргізілуі керек. факторлар. Data Mining енгізу әрқашан дерлік елеулі қаржылық шығындарды талап етеді. Сондай-ақ, компанияның бүгінгі алдында тұрған міндеттерін ескеріп қана қоймай, сонымен қатар болашақта жаңа міндеттердің туындау мүмкіндігіне сену керек.

#### Әдебиеттер тізімі:

1. Паклин Н.Б., Орешков В. И. Бизнес-аналитика: от данных к знаниям. – СПб.: Питер, 2009.
2. Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Методы и модели анализа данных: OLAP и Data Mining. 2-е изд.– СПб.: БХВ – Петербург, 2008.
3. Кацко И.А., Н.Б. Паклин. Практикум по анализу данных на компьютере. – М.: КолосС, 2009.
4. Дюк В.А., Самойленко А.П. Data Mining: учебный курс. – СПб.: Питер, 2001

## КЕСКІНДЕРДІ ТАНУДА ҚОЛДАНЫЛАТЫН ТЕРЕҢ ОҚЫТУДЫҢ ӘДІСТЕРІНЕ ШОЛУ

*Ш. Темиргазиева, И. Е. Базаркулова, Б.С. Омаров*  
*әл-Фараби атындағы Қазақ Ұлттық Университеті, Алматы, Қазақстан*  
*e-mail: [temirgazievash@gmail.com](mailto:temirgazievash@gmail.com)*

### **Аңдатпа**

Терең нейрондық желілер қазіргі уақытта дыбысты тану, табиғи тілді өңдеу, компьютерлік көру және т.б. сияқты жасанды интеллект жүйелерін құрудың ең танымал тәсілдерінің біріне айналуда. Мақалада терең нейрондық желінің даму тарихы мен қазіргі жағдайы, оқыту әдістері туралы шолу берілген. Жасанды нейрондық желінің моделі, нейрондық желіні оқыту алгоритмдері, соның ішінде терең нейрондық желілерді оқыту үшін қолданылатын қатені кері тарату алгоритмі қарастырылады. Жасырын қабаттары көп терең нейрондық желілер жоғалып бара жатқан градиент мәселесіне байланысты жаттықтыру қиын. Мақалада жүзден астам қабаттары бар терең нейрондық желілерді сәтті оқытуға мүмкіндік беретін осы мәселені шешу әдістері талқыланады. Қазіргі уақытта конволюционды нейрондық желілер компьютерлік көру тапсырмалары үшін пайдаланылады, ал қайталанатын нейрондық желілер, ең алдымен ұзақ мерзімді жады желілері және басқарылатын қайталанатын нейрондық желілер, реттіліктерді, соның ішінде табиғи тілді өңдеу үшін қолданылады.

### **Кіріспе**

Терең нейрондық желілер қазіргі уақытта машиналық оқытудың ең танымал әдістерінің біріне айналуда. Олар сөйлеуді тану, табиғи тілді өңдеу, компьютерлік көру, медициналық информатика және т.б. салалардағы балама әдістермен салыстырғанда жақсы нәтиже көрсетеді. Терең нейрондық желілерді сәтті қолданудың бір себебі - желі автоматты түрде мәселені шешуге қажетті деректерден маңызды мүмкіндіктерді таңдайды. Мақалада терең нейрондық желі архитектурасының дамуының тарихи шолуы және оларды оқыту тәсілдері берілген. Мұндай шолуды құрастыру міндеті терең нейрондық желілердің көптеген нұсқаларының ұсынылғандығымен және терминологияның уақыт өте өзгергенімен айтарлықтай қиындады. Жасанды нейрондық желілер моделі 1943 жылы ұсынылды, ал терең оқыту термині тек 2006 жылдан бастап кеңінен қолданыла бастады [1]. Бұған дейін терең желілерді жүктеу және терең естеліктерді үйрену терминдері қолданылды. Соңғы бірнеше жылда орын алған терең нейрондық желілердің танымалдылығының өсуін үш фактормен түсіндіруге болады. Біріншіден, компьютерлердің, соның ішінде GPU есептеу үдеткіштерінің (Graphics Processing Unit) өнімділігінің айтарлықтай артуы байқалды, бұл желінің терең нейрондық желілерін әлдеқайда жылдам және жоғары дәлдікпен оқытуға мүмкіндік берді. Екіншіден, терең нейрондық желілерді үйрету үшін қажетті деректердің үлкен көлемі жинақталды. Үшіншіден, нейрондық желіні оқыту әдістері әзірленді, олар жүз және одан да көп қабаттардан тұратын желілерді жылдам және сапалы оқытуға мүмкіндік береді [2].

### **Нейрондық желілерді оқыту**

Нейрондық желіні оқыту – бұл нейрондар арасындағы байланыстардың салмақтарын желі берілген дәлдікпен қажетті функцияға жақындататындай етіп анықтау процесі. Нейрондық желілерді оқытудың үш тәсілі бар: бақыланатын оқыту, бақылаусыз оқыту және күшейтетін оқыту. Мұғаліммен оқыту кезінде желінің кірісі дұрыс жауабы алдын ала белгілі болатын кіріс сигналдарының (объектілердің) жиынтықтарымен қамтамасыз етіледі (оқу жинағы). Желінің дұрыс шығыс сигналын шығарғанына байланысты салмақтар белгілі бір ережелерге сәйкес өзгереді. Бақылаусыз оқытуда дұрыс шығыс сигналы алдын ала белгісіз объектілер желілік кіріске беріледі.

Нейрондық желілерді оқыту идеясын алғаш рет 1949 жылы Дональд Хабб ұсынған. Хаббтың пікірінше, бірге жанатын нейрондардың байланысын күшейту керек, ал бір-

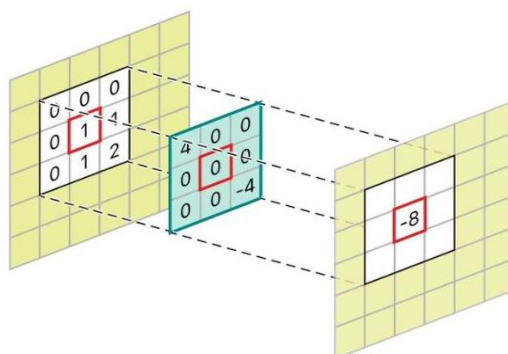
бірінен бөлек жанатын нейрондардың байланысын әлсірету керек. Хабб нейрондардың кіріс сигналдарының салмағын желі дұрыс жауап берген немесе бермегеніне сәйкес өзгерту ережелерін ұсынды [3] (бақылаумен оқыту). А.В. Новиков ұсынылған нейрондық оқыту әдісінің конвергенциясын Хабб ережелеріне негізделгенін [4] дәлелдеді, бұл жағдайда объектілер үлгісі сызықты түрде ажыратылады. Кейіннен мұғаліммен бірге және мұғалімсіз [5] оқу үшін де бірнеше ұқсас ережелер ұсынылды.

Қазіргі уақытта нейрондық желілерді, соның ішінде тереңдетіп оқыту үшін градиентті түсіру әдісіне негізделген қатені кері тарату алгоритмі қолданылады. Алгоритм 1970 жылы магистрлік диссертацияда [6] нейрондық желілерге қосылмай ұсынылды. Бұл алгоритмнің нейрондық желілерді оқытуға арналған алғашқы қолданылуы 1981 жылы жарияланған мақаласында сипатталған. Осыдан кейін осы тақырып бойынша тағы бірнеше жұмыс пайда болды [7].

Кері таралу алгоритмі арқылы нейрондық желілерді оқытуды жүзеге асырудың бірнеше нұсқалары бар. Толық жаттығу кезінде градиент оқу жинағындағы барлық нысандар үшін есептеледі. Дегенмен, бұл тәсіл көбінесе оқу жинағы үлкен болған кезде тиімді болмайды және оның барлық элементтерін өңдеуге көп уақыт кетеді. Альтернативті нұсқа стохастикалық градиентті түсіру әдісін пайдалану болып табылады, онда жаттығулар жиынтығының бір элементін (онлайн оқыту) немесе бірнеше элементтерді (партиялар немесе шағын үлгілерде оқыту) өңдеу кезінде салмақтар өзгереді. Іс жүзінде нейрондық желілерді жаттықтыру үшін жиі қолданылатын стохастикалық градиентті түсіру әдісі немесе оның модификациясы болып табылады [8].

### Терең нейрондық желілердің архитектурасы

1-суретте көрсетілген желі толық қосылған нейрондық желі деп аталады. Мұндай желіде келесі қабаттың әрбір нейроны алдыңғы қабаттың барлық нейрондарымен байланысады. Дегенмен, бұл нейрондарды желіге қосудың жалғыз жолы емес. Бұл бөлімде нейрондық желі архитектурасының дамуы қарастырылады. 1980 жылы Кунихико Фукусима неокогнитрон деп аталатын нейрондық желі архитектурасын ұсынды [9]. Архитектурада мысықтың көру қыртысының күрделі және қарапайым жасушаларына ұқсастық қолданылған [10]. Қарапайым ұяшықтар шекаралық бағдар сияқты қарапайым көрнекі белгілерге жауап ретінде жанады. Күрделі жасушалар сигналдардың кеңістіктік орналасуына аз тәуелді және жалпы белгілерді басшылыққа алады. Неокогнитронда жай ұяшықтар конволюционды қабаттарға, ал күрделі ұяшықтар ішкі іріктеу қабаттарына сәйкес келеді. Конволюционды қабаттарда салмақтардың берілген жиыны (конволюциондық ядро) бар конволюционды бірліктің (конволюционды бірлік) терезесі кіріс деректерінің екі өлшемді массиві бойынша, мысалы, кескін пикселдері бойынша жылжиды (1-сурет).



**1-сурет** – Неокогнитрондық нейрондық желінің конволюционды қабатының схемасы

Барлық конволюция түйіндері бірдей конволюциялық ядроларды пайдаланады, сондықтан конволюция желісін сипаттау үшін салыстырмалы түрде аз параметрлерді қажет етеді. Әдетте, конволюциялық қабаттар бір емес, бірнеше конволюция ядроларын пайдаланады.

Неокогнитрондағы конволюционды қабаттарды оқыту оқытушысыз жергілікті оқыту алгоритмдері арқылы орындалады немесе салмақтар алдын ала қойылады. Кеңістіктік орташалау қабаттардың ішкі үлгілерін алу үшін қолданылады. Осылайша, неокогнитрон терең нейрондық желі болғанымен, онда терең оқыту қолданылмайды.

### Терең нейрондық желілерді оқытуға арналған бағдарламалық жүйелер

Қазіргі уақытта терең нейрондық желілерді оқытуға арналған бағдарламалық жүйелердің көптеген нұсқалары жасалған. Олардың ішінде ең танымалдары: Caffe, Theano, TensorFlow, Torch және CNTK. Олардың негізгі сипаттамалары кестеде келтірілген.

Caffe Library [11] ең алғашқы терең оқытуды әзірлеу жүйелерінің бірі болып табылады. Ол Беркли компьютерлік көру және оқу орталығында (Berkeley Vision және Learning Center) әзірленген, бастапқы кодтар 2014 жылы ашылды. Caffe пайдалануға дайын алдын ала дайындалған модельдердің ең көп санын қамтиды. Theano жүйесі [12] Канаданың Монреаль университетінде құрылды. Theano Python тілінде жасалған, бірақ Python бағдарламасы компиляцияланатын және кейін орындалатын C++ бағдарламасына автоматты түрде түрленетіндіктен жоғары өнімділікті қамтамасыз етеді. TensorFlow [13] Google компаниясымен 2015 жылы жасалған және тензорлармен тиімді жұмыс істеуге және графикте ағынды деректерді өңдеуге арналған жүйелерді қамтиды. Torch кітапханасы [14] Луада әзірленген және MATLAB сияқты машиналық оқыту бағдарламаларын жасау үшін ыңғайлы жоғары деңгейлі интерфейсті қамтамасыз етеді. Си тілімен интеграцияның арқасында Theano сияқты жоғары өнімділік қамтамасыз етіледі. Torch авторлары Си және Lua интеграциясының қарапайымдылығына байланысты Python орнына Луаны пайдалануды жөн көрді. Microsoft – CNTK (Cognitive Toolkit) жүйесін жасап, 2016 жылы өзінің бастапқы кодтарын ашты.

**Кесте-1.** DL бағдарламалық жүйелерін салыстыру

Сипаттамалары	Caffe	Theano	Tensorflow	Torch	CNTK
Негізгі тілі	C++	Python	C++	Lua	C++
API	C++ Python	Python	C++ Python	Python Lua	C++, C# Python
Көпядролы CPU	+	+	+	+	+
GPU	+	+	+	+	+
Әзірлеуші	Berkeley Vision and Learning Center	Monreale university	Google	Ronan Collobert, Samy Bengio, Johnny Mariéthoz	Microsoft
Ашық бастапқы код	+	+	+	+	+
Алдын ала дайындалған желілер	+	-	+	+	+

### Қорытынды

Нейрондық желілер пайда болғаннан бері олардың архитектурасы мен оқыту әдістерінде көптеген өзгерістер болды. Қазіргі уақытта архитектураның екі түрі басым: компьютерлік көру тапсырмалары үшін сәтті қолданылатын конволюционды желілер

және табиғи тілді өңдеу тапсырмалары үшін белсенді қолданылатын қайталанатын желілер.

Конволюционды желілер автокодерлерді және терең сенім желілерін қолдана отырып, мұғаліммен және мұғалімсіз оқытудың үйлесімі арқылы оқытылды. Қалдық оқыту сияқты заманауи әдістер тек мұғаліммен оқытуды қолдануға және алдын ала дайындықтан бас тартуға мүмкіндік береді, бұл оқу процесін жылдамдатады және жеңілдетеді. Трансферттік оқыту да конволюционды нейрондық желілерді дамытудың маңызды бағыты болып табылады. Бұл тәсіл есептердің басқа түрлерін шешу үшін кейбір деректерге үйретілген нейрондық желілерді пайдалануды қамтиды. Нәтижесінде оқыту уақыты қысқарады және алдын ала дайындалған нейрондық желілерді қолдану аясы кеңейеді. Оқытуды күшейтумен конволюционды және қайталанатын нейрондық желілерді бірлесіп пайдалану да перспективалы [15].

Табиғи тілді өңдеу тапсырмалары үшін және ретті өңдеудің жалпы жағдайы үшін қазіргі уақытта қайталанатын нейрондық желілер қолданылады.

Олардың ішінде ең тиімдісі ұзақ мерзімді жады желілері және басқарылатын қайталанатын нейрондық желілер болып табылады, өйткені олар қызықты оқиғаларды ұзақ уақыт есте сақтауға мүмкіндік береді. Қайталанатын желілердің қосымша артықшылығы - мұғалімсіз және алдын ала белгіленген деректер жиынтығынсыз оқыту мүмкіндігі.

Нейрондық желілерді практикада қолданудың кең таралуы терең нейрондық желілерді оқытуға арналған дайын шешімдердің үлкен санының болуына байланысты мүмкін болды [16], оның ішінде заманауи көп ядролы процессорларды, GPU, сондай-ақ таратылған жады бар есептеу кластерлері.

#### Әдебиеттер тізімі

1. LeCun Y., Bengio Y., Hinton G. “Deep Learning” *Nature* Vol. 521, (2015): P. 436–444. DOI: 10.1038/nature14539.
2. Schmidhuber J. “Deep Learning in Neural Networks: an Overview” *Neural Networks* Vol. 1, (2015): P. 85–117, DOI: 10.1016/j.neunet.2014.09.003.
3. He K., Zhang X., Ren S., et al. “Deep Residual Learning for Image Recognition” *2016 IEEE Conference on Computer Vision and Pattern Recognition (Las Vegas, NV, USA, 27–30 June 2016)*, (2016): P. 770–778. DOI: 10.1109/CVPR.2016.90.
4. Novikoff A.B. “On Convergence Proofs on Perceptrons” *Symposium on the Mathematical Theory of Automata* Vol. 12. (1962): P. 615–622
5. Rosenblatt F. “The Perceptron: a Probabilistic Model for Information Storage and Organization in the Brain” *Psychological Review* (1958): P. 65–386. DOI: 10.1037/h0042519.
6. Widrow B., Hoff M. “Associative Storage and Retrieval of Digital Information in Networks of Adaptive Neurons” *Biological Prototypes and Synthetic Systems* Vol. 1. 160 (1962): p. DOI: 10.1007/978-1-4684-1716-6\_25.
7. Werbos P.J. “Applications of Advances in Nonlinear Sensitivity Analysis” *Lecture Notes in Control and Information Sciences* Vol. 38, (1981): P. 762–770. DOI: 10.1007/BFb0006203.
8. Qian N. “On the Momentum Term in Gradient Descent Learning Algorithms” *Neural Networks: The Official Journal of the International Neural Network Society* Vol. 12, No. 1. (1999): P. 145–151. DOI: 10.1016/s0893-6080(98)00116-6.
9. Fukushima K. “Neocognitron: a Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position” *Biological Cybernetics* Vol. 36, No. 4. (1980): P. 193–202. DOI: 10.1007/BF00344251.
10. Wiesel D.H., Hubel T.N. “Receptive Fields of Single Neurones in the Cat’s Striate Cortex” *The Journal of Physiology* Vol. 148, No. 3. (1959): P. 574–591. DOI: 10.1113/jphysiol.1959.sp006308.

11. Jia Y., Shelhamer E., Donahue J., et al. “Caffe: Convolutional Architecture for Fast Feature Embedding” *Proceedings of the 22nd ACM International Conference on Multimedia (Orlando, FL, USA, November 03–07, 2014)*: P. 675–678. DOI: 10.1145/2647868.2654889
12. Bergstra J., Breuleux O., Bastien F., et al. “Theano: a CPU and GPU Math Expression Compiler” *Proceedings of the Python for Scientific Computing Conference (SciPy) (Austin, TX, USA, June 28 – July 3, 2010)*: P. 3–10.
13. Abadi M., Agarwal A., Barham P. “TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems” *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16) (Savannah, GA, USA, November, 2–4, 2016)*: P. 265–283.
14. Collobert R., Kavukcuoglu K., Farabet C. “Torch7: a Matlab-like Environment for Machine Learning” *BigLearn, NIPS Workshop (Granada, Spain, December 12–17, 2011)*.
15. Weiss K., Khoshgoftaar T.M., Wang D. “A Survey of Transfer Learning” *Journal of Big Data* Vol. 3, No. 1. (2016): P. 1–9. DOI: 10.1186/s40537-016-0043-6
16. Seide F., Agarwal A. “CNTK: Microsoft’s Open-Source Deep-Learning Toolkit” *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16) (San Francisco, California, USA, August 13–17, 2016)*: P. 2135–2135. DOI: 10.1145/2939672.2945397.

**ОБЪЕКТНО-ОРИЕНТИРОВАННЫЙ ЛОГИЧЕСКИЙ ЯЗЫК  
ПРОГРАММИРОВАНИЯ ДЛЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И  
РОБОТОТЕХНИКИ.**

Сергей Гончаров, Андрей Нечесов  
Институт математики им. Соболева, Новосибирск, Россия.  
[s.s.goncharov@math.nsc.ru](mailto:s.s.goncharov@math.nsc.ru), [nechesov@math.nsc.ru](mailto:nechesov@math.nsc.ru)

**Аннотация:** В статье представлен новый высокоуровневый объектно-ориентированный язык программирования  $L^*$ , который является консервативным расширением  $r$ -полного языка  $L$ . Таким образом  $L^*$ -программы могут применяться для реализации любых алгоритмов полиномиальной вычислительной сложности в таких направлениях как искусственный интеллект, робототехника и умные контракты. Более того синтаксис данного языка максимально приближен к синтаксису таких популярных языков программирования как C++, PHP, JavaScript, что обеспечивает программистам быстрый вход в разработку.

**Ключевые слова:** полиномиальная вычислимость, семантическое программирование, искусственный интеллект, робототехника, объяснимый ИИ, нейронные сети

**Введение:** Семантическое программирование как отдельное направление было представлено в 70-80-х годах прошлого века академиками Ершовым и Гончаровым, а также профессором Свириденко. Основная идея заключалась в создании логического языка программирования в котором бы базовыми конструкциями выступали объекты математической логики такие как формулы и термы. Основной моделью, в которой происходит вся работа, была выбрана наследственная конечно-списочная надстройка  $NW(M)$  конечной сигнатуры  $\sigma$  [12]. За основу был взят логический язык (назовем его  $L_0$ ), формулы которого представляли собой  $\Delta_0$ -формулы сигнатуры  $\sigma$ , а термы - стандартные термы логики предикатов первого порядка [2]. В 2017 году Гончаров предложил синтаксическую конструкцию условного терма [3], термальное расширение (назовем его  $L_1$ ) с помощью которого языка  $L_0$  не выводит за рамки полиномиальной вычислительной сложности [9] проверки истинности  $\Delta_0$ -формул. Но вопрос представления любой полиномиально вычислимой функции с помощью  $L_1$ -программ оставался открытым. Только в 2021 году нам удалось построить синтаксическую конструкцию  $r$ -итерационного терма [1] такую, что расширение (далее  $L$ ) языка  $L_1$  с помощью  $r$ -итерационных термов не выводило за рамки полиномиальности. В работе [1] показано, что класс  $L$ -программ совпадает с классом  $P$  -- всех алгоритмов полиномиальной вычислительной сложности. В этом же языке также выразимы все конструкции, представленные Гончаровым и Свириденко в их работе [4].

В этой статье будет представлен новый объектно-ориентированный язык  $L^*$ , являющийся консервативным расширением логического языка  $L$ .  $L^*$ -программы можно будет использовать в таких направлениях как объяснительный искусственный интеллект (далее ХАИ), робототехника, смарт-контракты. На данный момент большинство алгоритмов, используемых в искусственном интеллекте (далее АИ) и машинном обучении (далее МЛ), являются черными ящиками [8] и поэтому не могут применяться во многих направлениях, где требуется объяснить результат человекупонятным языком. Более того, если алгоритм реализован на Тьюринг полном высокоуровневом языке [7], то, скорее всего, этот алгоритм может выйти за пределы полиномиальности [5-6].

Базовые  $r$ -вычислимые конструкции применяемые в машинном обучении [10], искусственном интеллекте [13] могут быть реализованы с помощью  $L^*$ -программ.



Например, большинство нейронных сетей [11], используемых в алгоритмах машинного обучения, имеют полиномиальную вычислительную сложность.

### **P-итерационные термы и их модификации**

В работе [1] было определено индуктивное понятие  $p$ -итерационного терма вида  $\text{Iteration}_{g,\varphi}(x, h(|x|))$  термальное расширение с помощью которого языка  $L_1$  помогло построить  $p$ -полный язык  $L$ . В  $p$ -итерационном терме базовыми конструкциями выступают  $L$ -программа  $g$  и  $L$ -формула  $\varphi$ . Итерационный терм считает значение  $g^i(w)$  и подставляет его в  $\varphi$ , если  $\varphi(g^i(w))$  истина на  $NW(M)$ , то вычисление останавливается и выдается результат  $g^i(w)$ , иначе вычисление повторяется пока  $i$  не превысит некоторого полинома  $h(w)$ , тогда значение  $p$ -итерационного терма будет false. В  $p$ -итерационном терме требуется, чтобы для любого  $w$  принадлежащего основному множеству модели  $NW(M)$  выполнялось  $|g(w)| \leq |w| + C$ . И в 2021 году была доказана следующая важная теорема:

#### **Теорема 1: (решение проблемы равенства $P = L$ ) [1]**

Пусть модель  $NW(M)$  сигнатуры  $\sigma$  —  $p$ -вычислима. Тогда верно следующее:

- 1) Вычислительная сложность любой  $L$ -программы является полиномиальной.
- 2) Для любой  $p$ -вычисляемой функции существует подходящая  $L$ -программа реализующая ее.

В этой работе мы ослабим это условие и представим конструкцию модифицированного  $p$ -итерационного терма. Предположим, что для фиксированного  $n$   $L$ -программа  $g(x)$  задана так

$$g(x) = w^*, \text{ если } w = \langle w_1, \dots, w_n \rangle, \text{ где } w^* = \langle w_1^*, \dots, w_n^* \rangle$$

что выполняются следующие условия:

- 1)  $|w_1^*| \leq |w_1| + C \cdot \sum |w_i|^p, i \in [2, \dots, n]$
- 2)  $|w_i^*| \leq |w_i|$ , для всех  $i \in [2, \dots, n]$
- 3) для любого  $w = \langle w_1, \dots, w_n \rangle$  либо существует такое  $i \leq h(|w|)$ , что  $\varphi(g^i(w))$  истина, либо  $\varphi(g^i(w))$  ложна для любого натурального  $i$ .

Тогда можно сформулировать следующую лемму, которая, по-сути, повторяет доказательство Теоремы 1 из [1].

**Лемма 1:** Терм  $\text{Iteration}_{g,\varphi}$  с условиями 1)-3) на  $g$  и  $\varphi$  является  $p$ -вычислимым.

### **L\*-программы и виртуальные машины**

Пусть как и раньше  $NW(M)$  сигнатуры  $\sigma$  —  $p$ -вычисляемая модель. Тогда под виртуальной машиной  $V$  будем понимать следующую тройку:

$$V = (NW(M), \sigma, \langle Fl, Cl, Hl, Sl, Ll \rangle)$$

где  $Fl$  – список  $L^*$ -определимых функций,  $Cl$  – список  $L^*$ -определимых классов,  $Hl$  – вспомогательный список,  $Sl$  – список, являющийся своеобразным аналогом экрана для виртуальной машины,  $Ll$  – список, являющийся своеобразным аналогом ленты машины  $V$ .

Взаимной индукцией определим понятия  $L(V)$ -программы,  $L(V)$ -формулы, а также  $L^*$ -программы:

#### **Индуктивное определение $L(V)$ -программы:**

- 1)  $t(x)$  –  $L(V)$ -программа, если  $t(x)$  –  $L$ -программа
- 2)  $f(t_1, \dots, t_n)$  –  $L(V)$ -программа, если  $t_1, \dots, t_n$  –  $L(V)$ -программы и  $f \in \sigma$
- 3)  $f(t_1, \dots, t_n)$  –  $L(V)$ -программа, если  $t_1, \dots, t_n$  –  $L(V)$ -программы и  $f$  –  $L^*$ -определимая функция, которая задается следующей синтаксической конструкцией:

$[mod] \text{ function } f(x_1, \dots, x_n) \{p\}$

где  $p$  –  $L^*$ -программа, являющаяся телом функции,  $mod \in \{\text{public, protected, private}\}$  – модификатор, квадратные скобки означают, что указывать модификатор не обязательно.

- 4)  $\text{Cond}(t_1, \varphi_1, \dots)$  –  $L(V)$ -программа, где  $\text{Cond}$  -условный терм из [3], при этом все  $t_i$  –  $L(V)$ -программы,  $\varphi_j$  –  $L(V)$ -формулы.

5)  $\text{Iteration}_{g,\varphi}(t,h(|t|))$  – L(V)-программа, где  $\text{Iteration}_{g,\varphi}$  – p-итерационный терм,  $g$  – L-программа,  $\varphi$  – L-формула,  $t$  – L(V)-программа,  $h$  – полином.

**Индуктивное определение L(V)-формулы:**

- 1)  $\varphi(x)$  – L(V)-формула, если  $\varphi(x)$  – L-формула
- 2)  $\varphi: P(t_1, \dots, t_n)$  – L(V)-формула, если  $t_1, \dots, t_n$  – L(V)-программы
- 3)  $\varphi: t_1 = t_2$  – L(V)-формула, если  $t_1, t_2$  – L(V)-программы
- 4)  $\varphi: \varphi_1 \delta \varphi_2$  – L(V)-формула, если  $\varphi_1, \varphi_2$  – L(V)-формулы и  $\delta \in \{\&, \vee, \rightarrow\}$
- 5)  $\varphi: \neg \varphi_1$  – L(V)-формула, если  $\varphi_1$  – L(V)-формула
- 6)  $\varphi: \exists x \delta t \varphi_1$  и  $\forall x \delta t \varphi_1$  – L(V)-формулы, если  $\delta \in \{\in, \subseteq, \leq\}$ ,  $t$  – L(V)-программа,  $\varphi_1$  – L(V)-формула

**Индуктивное определение L\*-формулы:**

- 1)  $y := t(x)$ ; – L\*-программа, если  $t(x)$  – L(V)-программа
- 2)  $\text{this.y} := t(x)$ ; – L\*-программа, если  $t(x)$  – L(V)-программа,  $\text{this}$  — вспомогательное слово [14], которое доступно, если метод вызывается в контексте объекта. [14]
- 3)  $\text{return } t(x)$ ; – L\*-программа, если  $t(x)$  – L(V)-программа
- 4)  $\text{print}(t(x))$ ; – L\*-программа, если  $t(x)$  – L(V)-программа
- 5)  $p_1 \cdot p_2$  – L\*-программа, если  $p_1, p_2$  – L\*-программы и  $\cdot$  – символ конкатенации
- 6) (задание функции как в [14]) если  $p^*$  – L\*-программа определяющая функцию, тогда  $[\text{mod}] \text{function } \text{Func}(x_1, \dots, x_n) \{p^*\}$  – L\*-программа где  $\text{mod}$  -модификатор,  $\text{mod} \in \{\text{public}, \text{protected}, \text{private}\}$  и квадратные скобки означают необязательный параметр вхождения. Модификатор чаще всего используется для ограничения видимости методов классов, при их задании.
- 7) (задание класса как в [14]) если  $p^*$  – L\*-программа, тогда  $\text{class } \text{ClassName} [\text{extends } \text{parentClass}] \{p^*\}$  – L\*-программа  $\text{extends } \text{parentClass}$  – специальная конструкция, которая описывает наследование свойств и методов класса родителя. Этот параметр необязательный.
- 8) (инициализация объекта как в [14]) если  $t_1, \dots, t_k$  – L(V) — программы, тогда  $\text{object} := \text{new } \text{ClassName}(t_1, \dots, t_k)$ ; – L\*-программа

Будем говорить, что L(V)-формула или L(V)-программа является нестандартной, если в ее построении в качестве параметров участвуют L\*-определимые функции или L\*-определимые методы классов. Под нестандартными конструкциями будем понимать нестандартные L(V)-формулы, нестандартные L(V) -программы и инициализацию объекта некоторого L\*-определимого класса.

**Компиляция L\*-программ**

Пусть  $V$  – виртуальная машина определенная выше,  $p$  – L\*-программа. Тогда зададим функцию компиляции следующим образом:

$$C: \langle \langle Fl, Cl \rangle, p \rangle \rightarrow \langle \langle Fl^*, Cl^* \rangle, p^* \rangle$$

которая переводит L\*-программу  $p$  в L\*-программу  $p^*$  следующим образом:

- 1) все куски кода, которые описывают функцию или класс удаляются из программы  $p$  и заносятся в соответствующие списки  $Fl$  или  $Cl$  в виде списочного кода вида:

$$\langle \text{funcName}, [\text{mod}], \langle x_1, \dots, x_k \rangle, p^* \rangle$$

и

$$\langle \text{className}, \langle x_1, \dots, x_k \rangle, \langle \langle f_1, [\text{mod}], \langle x_1, \dots, x_{k1} \rangle, p_1^* \rangle, \dots, \langle \langle f_m, [\text{mod}], \langle x_1, \dots, x_{km} \rangle, p_m^* \rangle \rangle \rangle$$

где  $x_1, \dots, x_k$  – имена свойств класса,  $f_1, \dots, f_m$  – названия методов класса.

Заметим, что в полученной программе  $p^*$  больше не содержится никаких конструкций, описывающих класс или задающих функцию.

**Замечание:**  $C$  – p-вычислимая функция.

**Исполнение L\*-программ.**

Пусть  $p^*(x_1, \dots, x_n)$  – L\*-программа, полученная в результате компиляции L\*-программы  $p$ . Индуктивно определим понятие исполнения программы  $p^*$  виртуальной машиной  $V$  следующим образом:

1) в список  $L1$  машина  $V$  первым элементом добавит список  $\langle\langle V1_{p^*}, O1_{p^*}, H1_{p^*}, R1_{p^*} \rangle, p^*\rangle$ , где  $V1_{p^*}$  будем называть списком проинициализированных переменных,  $O1_{p^*}$  – списком проинициализированных объектов,  $H1_{p^*}$  – вспомогательный список,  $R1_{p^*}$  – список, который будет хранить возвращаемое значение  $L^*$ -программой  $p^*$ . Первоначально  $V1_{p^*}$  имеет вид:  $\langle\langle x_1, w_1 \rangle, \dots, \langle x_n, w_n \rangle \rangle$ , где входящие переменные  $x_i$  программы  $p^*$  означены как  $w_i$ . Списки же  $O1_{p^*}, H1_{p^*}, R1_{p^*}$  вначале пустые. Далее, в процессе работы будут появляться и другие списки вида  $\langle\langle V1_{q^*}, O1_{q^*}, H1_{q^*}, R1_{q^*} \rangle, q^*\rangle$ , для  $L^*$ -программ  $q^*$  (пункты 2) и 3) ниже).

2) На каждом шаге машина  $V$  строчка за строчкой считывает программу  $q^*$  из последнего элемента списка  $L1$ . Если в процессе работы  $V$  встречается нестандартная  $L(V)$ -конструкция, в которой в качестве параметров не присутствуют никакие другие нестандартные  $L(V)$ -конструкции, то:

а) если данная  $L(V)$ -конструкция имеет вид  $f(t_1, \dots, t_n)$ , где  $t_1, \dots, t_n$  –  $L$ -программы, тогда она заменяется на ранее неиспользуемую переменную  $z$ , эта переменная записывается в список  $H1_{q^*}$  и следующий элемент:

$$\langle\langle V1_{q^{**}}, O1_{q^{**}}, H1_{q^{**}}, R1_{q^{**}} \rangle, q^{**} \rangle$$

добавляется в конец списка  $L1$ , где  $q^{**}$  – тело функции  $f$ . При этом вычисленные значения  $w_i$   $L$ -программ  $t_i$  последовательно добавляются в список  $V1_{q^{**}}$  в виде элементов  $\langle x_i, w_i \rangle$ .

б) если данная  $L(V)$ -конструкция имеет вид  $obj.m(t_1, \dots, t_n)$ , где  $t_1, \dots, t_n$  –  $L$ -программы,  $m$  – метод класса, которому принадлежит объект  $obj$ , тогда она заменяется на ранее неиспользуемую переменную  $z$ , эта переменная записывается в список  $H1_{q^*}$  и следующий элемент:

$$\langle\langle V1_{q^{**}}, O1_{q^{**}}, H1_{q^{**}}, R1_{q^{**}} \rangle, q^{**} \rangle$$

добавляется в конец списка  $L1$ , где  $q^{**}$  – тело метода класса. При этом вычисленные значения  $w_i$   $L$ -программ  $t_i$  последовательно добавляются в список  $V1_{q^{**}}$  в виде элементов  $\langle x_i, w_i \rangle$ . В  $O1_{q^{**}}$  из  $O1_{q^*}$  копируется объект  $\langle obj, ClassName, \langle\langle x_1, w_1 \rangle, \dots, \langle x_n, w_n \rangle \rangle \rangle$ , где  $x_i$  – инициализированные свойства объекта  $obj$ .

в) если данная  $L(V)$ -конструкция является инициализацией объекта класса  $ClassName$

$$obj := new ClassName(t_1, \dots, t_n);$$

где  $t_1, \dots, t_n$  –  $L$ -программы, тогда она заменяется на  $obj := new z$ ;

При этом переменная  $z$  до этого нигде не использовалась, эту переменную записывают в список  $H1_{q^*}$  и следующий элемент:

$$\langle\langle V1_{q^{**}}, O1_{q^{**}}, H1_{q^{**}}, R1_{q^{**}} \rangle, q^{**} \rangle$$

добавляется в конец списка  $L1$ , где  $q^{**}$  – тело конструктора класса  $ClassName$ . При этом список  $V1_{q^{**}}$  содержит вычисленные значения  $w_i$   $L$ -программ  $t_i$  в виде элементов  $\langle x_i, w_i \rangle$ .

3) рассмотрим остальные случаи исполнения  $L^*$ -программы  $q^*$  машиной  $V$

а) если  $y := t(x)$ ; – машина  $V$  добавляет элемент в  $V1_{q^*}$ , где  $w_1$  – вычисленное значение  $t(x)$

б)  $this.y := t(x)$ ; – берет первый элемент списка

в)  $return t(x)$ ; – машина  $V$  вычисляет значение  $w_i$   $L$ -программы  $t(x)$  в список  $R1_{q^*}$  и далее, машина  $V$  в списке  $L1$  заменяет  $q^*$  на пустой список, иначе, удаляет последний элемент из  $L1_{q^*}$  и подставляет значение из  $R1_{q^*}$  в последний элемент списка  $L1$  вместо переменной, которая находится в списке  $H1_{q^*}$  вместо этой переменной в программе  $q$ .

г)  $print(t(x))$ ; – машина  $V$  вычисляет значение  $w$   $L$ -программы  $t(x)$  с означиваниями из списка  $V1_{q^*}$  и  $O1_{q^*}$  и печатает вычисленное значение в список  $S1$  машины  $V$ .

Машина  $V$  останавливает свою работу только в том случае, когда в списке  $L1$  остался последний элемент и список  $R1_{p^*}$  не пуст. При этом возвращаемое значение при работе над программой  $p^*$  будет этим элементом из  $R1_{p^*}$ .

Конфигурацией  $c_i$  машины  $V$  будем называть список  $L1$  в некоторый момент работы машины  $V$  над  $L^*$ -программой  $p^*$ :

$$c_i: \langle\langle\langle V1_{p^*}, O1_{p^*}, H1_{p^*}, R1_{p^*} \rangle, p^* \rangle, \dots, \langle\langle V1_{q^{**}}, O1_{q^{**}}, H1_{q^{**}}, R1_{q^{**}} \rangle, q^{**} \rangle \rangle$$

Ходом работы машины  $V$  будем называть последовательность конфигураций  $c_0, \dots, c_k$ .

В силу того, что ход работы машины  $V$  может быть смоделирован подходящим модифицированным  $p$ -итерационным термом  $\text{Iteration}_{g,\varphi}$ , и с учетом леммы 1 можно сформулировать две заключительных теоремы, причем вторая теорема автоматически следует из первой в силу  $p$ -полноты языка  $L$ .

**Теорема 2:** Вычислительная сложность любой  $L^*$ -программы полиномиальна.

**Теорема 3:** Язык  $L^*$  является консервативным расширением языка  $L$ .

**Заключение:** В работе был описан новый объектно-ориентированный язык программирования  $L^*$ . Все программы в этом языке также как и в языке  $L$  имеют полиномиальную вычислительную сложность. Тем самым, при реализации алгоритмов объяснительного искусственного интеллекта с помощью  $L^*$ -программ нам гарантирован быстрый и четкий ответ, причем этот ответ будет логически обоснован и понятен человеку. Более того, в язык  $L^*$  можно добавлять в качестве встроенных библиотек любые  $p$ -вычислимые конструкции типа нейронных сетей, порождающих грамматик и т. д.

#### Список литературы:

- [1] Goncharov, S.S.; Nechesov, A.V. *Solution of the problem  $P = L$* . MDPI Mathematics, **2021**, 10(1) 113, DOI:10.3390/math10010113.
- [2] Goncharov, S.S.; Nechesov, A.V. *Polynomial analogue of Gandy's fixed point theorem*. MDPI Mathematics, **2021**, 9(17) 2102, DOI:10.3390/math9172102.
- [3] Goncharov, S.S. *Condition terms in semantic programming*. Sib Math J, **2017**, 58, 794-800 DOI:10.1134/S0037446617050068
- [4] Goncharov, S.S.; Sviridenko, D.I. *Logical Language of Description of Polynomial Computing*. Doklady Mathematics, **2019**, 485(1), 11-14, DOI:10.31857/S0869-5652485111-14
- [5] Nechesov, A.V. *Some questions of polynomial computable representations for generative grammars and Backus-Naur forms*. Matematicheskie Trudy, **2022**, 25(1), 134-151, DOI:10.33048/mattrudy.2022.25.106
- [6] Alaev, P.E. *Structures Computable in Polynomial Time. I* Algebra Log. **2017**, 55, 421--435, DOI:10.17377/alglog.2016.55.601
- [7] Michaelson, G. *Programming Paradigms, Turing Completeness and Computational Thinking*. The Art, Science, and Engineering of Programming. **2020**, 4(3). DOI:10.22152/programming-journal.org/2020/4/4
- [8] Bathaee, Y. *The artificial intelligence black box and the failure of intent and causation*. Harvard Journal of Law & Technology. **2018**. 31(2)
- [9] Cenzer, D.; Remmel, J. *Polynomial-time versus recursive models*. Ann. Pure Appl. Log. **1991**. 54, 17--58.
- [10] Deisenroth, M.P.; Faisal, A.A.; Ong, C.S. *Mathematics for Machine Learning*. Published by Cambridge University Press, **2020**.
- [11] Hagan, M.T.; Demuth, H.B.; Beale, M.H.; Jesus, O.D. *Neural Network Design*. Martin Hagan, **2014**
- [12] Ospichev, S.S.; Ponomaryov, D.K. *On the complexity of formulas in semantic programming*. Sib. Electron. Math. Rep. **2018**, 15, 987--995, DOI:10.17377/semi.2018.15.083
- [13] Russel, S.J.; Norvig, P. *Artificial Intelligence - A Modern Approach (3rd Edition)*. Prentice Hall, **2010**
- [14] PHP classes and objects. URL: <https://www.php.net/manual/en/language.oop5.php>

## МАРКЕТИНГОВЫЕ ВОЗМОЖНОСТИ ПРОДВИЖЕНИЯ БРЕНДА С ПОМОЩЬЮ ИНТЕРНЕТ-МАРКЕТИНГА

Сериков М.К., Алимжанова Л.М., Сербин В.В.  
*Евразийский технологический университет, Казахстан, Алматы*  
e-mail: 7014547793@mail.ru

В последние годы социальные сети изменили повседневную жизнь миллиардов людей по всему миру. Благодаря таким социальным сетям VK, Facebook, Instagram, Twitter, TikTok, Youtube люди могут участвовать в международном диалоге и расширять свои возможности, чтобы выражать и делиться своими идеями, вести бизнес или просто оставаться в курсе событий в мире, перегруженном информацией [1]. Несмотря на некоторую критику в отношении отсутствия конфиденциальности, напрасной траты времени или распространения усиленного невежества в интернет-пространстве, которые трудно контролировать, социальные сети продолжают расти и развиваться.

Казахстанское пространство социальных сетей не является исключением. С момента своего появления социальные сети привлекли миллионы пользователей, и люди в Казахстане являются членами некоторых местных и мировых сетей. Таким образом, в определенных областях человеческой деятельности формируются целевая аудитория.

На сегодняшний день люди используют Интернет больше, чем когда-либо прежде, и когда потребители хотят найти информацию о товаре или услуге, они прибегают к помощи поисковых систем и социальных сетей. Компании, которые не занимаются маркетингом в Интернете, им очень трудно будет стимулировать продажи или развивать свой бренд [2].

Social Media Marketing (SMM) – одна из технологий интернет-маркетинга, которая основана на рекламе бренда в известных социальных сетях: VK, Facebook, Instagram, Twitter, TikTok, Youtube и других [3]. Продвижение бренда в социальных сетях - это часть интернет-маркетинга, а также является методом коммуникации с целевой аудиторией, работы с лояльностью и вовлечения аудитории. SMM-продвижение обеспечивает узнаваемость бренда, находится в постоянной коммуникации с клиентами через социальные сети, создание интереса к рекламируемым продуктам. Основными методами SMM является постоянное заполнение профиля в социальных сетях, ведение тематического контента, общение с аудиторией, вирусный маркетинг, создание положительного имиджа [4]. В результате чего бренд набирает свою первую аудиторию, первых клиентов.

Анализ текущего состояния исследования по рынку Интернет-маркетинга. Казахстан занимает лидирующее положение в Центральной Азии и СНГ по уровню охвата Интернетом населения страны. В качестве новой области, вызывающей озабоченность развивающихся стран, таких как Казахстан, исследования новых медиа в основном включают две категории: история развития Интернета и новых медиа, включая социальные сети, и роль социальных сетей в казахстанском обществе.

Общая численность населения Казахстана в январе 2022 года составляла 19,10 миллиона человек. Из них доля пользователей социальных сетей по сравнению с прошлым годом выросли на 1,80 млн. человек (15,0%), и составила 72,3%, что эквивалентно 13,80 млн. человек. Важно отметить, что пользователи социальных сетей могут не представлять собой уникальных лиц [5]. А самыми популярными мобильными приложениями по-прежнему остаются WhatsApp и Instagram, причем и на iPhone, и на Android [6].

По данным DataReportal по состоянию на январь 2022 года в Казахстане среди социальных сетей лидирует Youtube (12 млн. пользователей), Instagram (11,75 млн. пользователей), TikTok (7,26 млн. пользователей), VK (7,20 млн. пользователей), Facebook (2,30 млн. пользователей), Snapchat (1,45 млн. пользователей), LinkedIn (1,10 млн. пользователей), Twitter (0,20 млн. пользователей) [5-7].

Основные моменты рекламы в социальных сетях [8-9]:

- По прогнозам, расходы на рекламу в сегменте рекламы в социальных сетях достигнут 54 млрд. тенге (\$126 млн) в 2022 году.
- Ожидается, что расходы на рекламу будут демонстрировать годовой темп роста на уровне 10,49%, в результате чего к 2026 году прогнозируемый объем рынка составит 80 млрд. тенге (\$187,80 млн).
- Для глобального сравнения большая часть расходов на рекламу будет приходиться на США (\$80 670 млн. в 2022 году).
- В сегменте рекламы в социальных сетях в 2026 году общие рекламные расходы в размере \$130,60 млн. будут получены за счет мобильных устройств.
- Ожидается, что к 2026 году в сегменте рекламы в социальных сетях число пользователей составит 16,2 млн пользователей.
- Рыночная доля Meta Platforms (Facebook, WhatsApp, Instagram) оценивается в 90% сегмента рекламы в социальных сетях Казахстана в 2021 году.

Определение рынка. Реклама в социальных сетях включает весь доход от рекламы, полученный от социальных сетей или бизнес-сетей, таких как Facebook, Twitter, Instagram, LinkedIn или VK. Объявления в социальных сетях могут отображаться как рекламные посты в органическом контенте или помимо ленты новостей.

Расходы на рекламу в социальных сетях через ПК и мобильные устройства, не включая рекламу в онлайн-играх, а также доходы, полученные от членских подписок или премиальных взносов приведены в соответствии с рисунком 1.

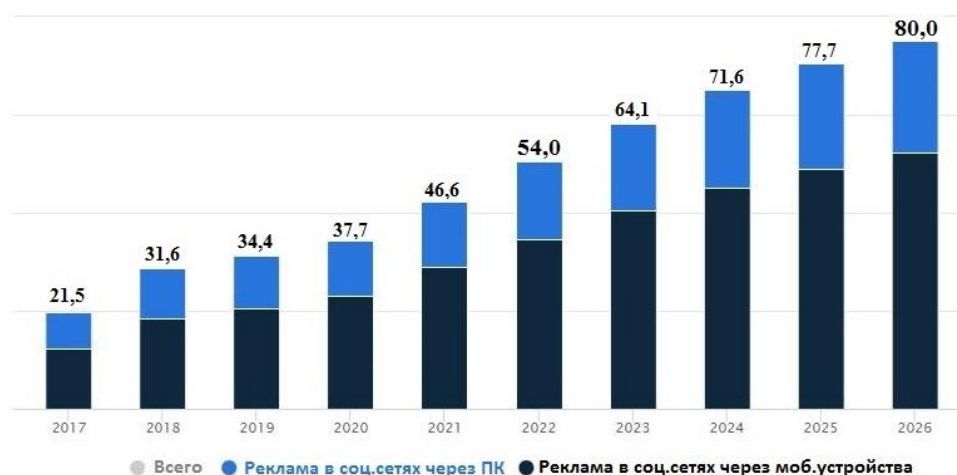


Рис.1 - Расходы на рекламу (в млрд. тенге)

По мнению аналитиков, немецкой компаний Statista [8], специализирующейся на рыночных и потребительских данных, одним из основных трендов рекламы в социальных сетях является дальнейшая монетизация социальных сетей и мессенджеров. Интеграция торговых и платежных решений в социальные сети в сочетании с точной локализацией повысит вовлеченность пользователей, конверсию и эффективность расширенного таргетинга.

Социальные сети в последнее время приобрели огромную известность как высокоэффективный канал коммуникации в наше современное время цифровой жизни. Он был поставлен на пьедестал в различных потоках для облегчения совместного взаимодействия между предприятиями, группами, обществами, организациями, потребителями, сообществами, форумами и т.п. Средние расходы на рекламу на пользователя Интернета показаны в соответствии с рисунком 2.

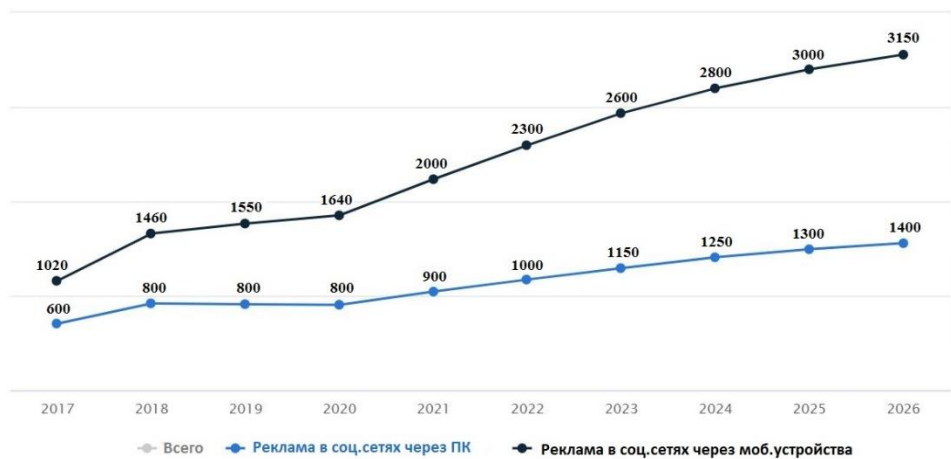


Рис.2 - Средние расходы на рекламу на пользователя Интернета (в тенге)

Помимо потенциала роста рекламы в социальных сетях, интеграция или распространение рекламных площадок наблюдается в приложениях для обмена сообщениями, таких как WhatsApp и Telegram, быстро увеличивая потенциальный доход текущих ключевых игроков рынка, таких как Meta, VK и другие. Охват аудитории рекламы через социальные сети (в млн. человек) приведен в соответствии с рисунком 3.

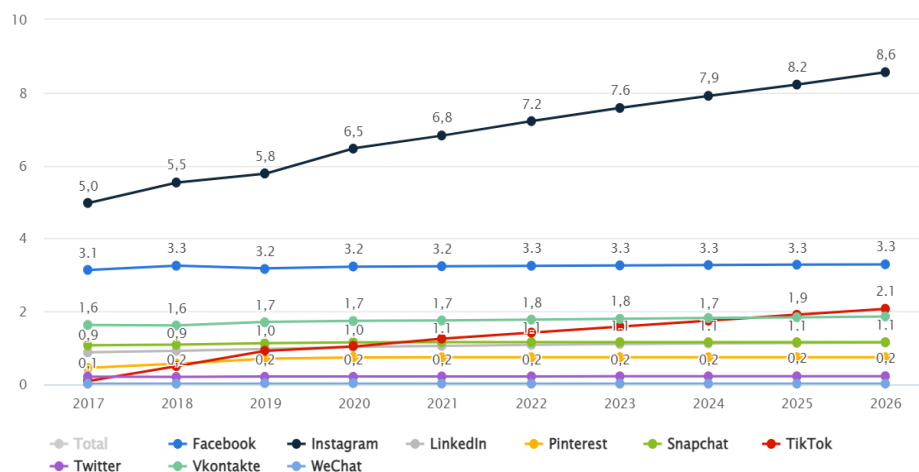


Рис.3 - Охват аудитории через социальные сети

В глобальном рейтинге Казахстан занимает 56 место по получению среднего дохода от рекламы в социальных сетях, и приведен в соответствии с рисунком 4, с выручкой 54 млрд. тенге (\$126 млн.) [8].

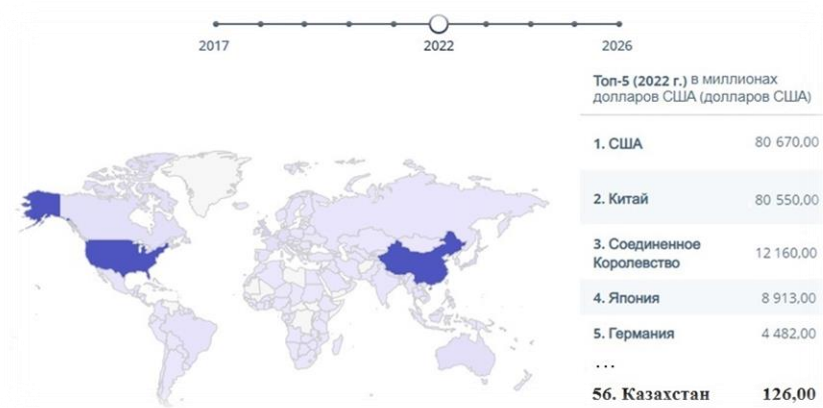


Рис.4 – Глобальный рейтинг по получению среднего дохода от рекламы в социальных сетях

В процессе исследования также необходимо описать целевую аудиторию, лояльных к SMM для продвижения бренда компаний.

Также стоит отметить, про WhatsApp – один из платформ с базой пользователей более 2 миллиардов человек по всему миру, что является важной вехой, которую они достигли в 2020 году. Недавние темпы роста приложения были поразительными: последний миллиард пользователей присоединился к нему всего за четыре года!

Поскольку WhatsApp отказывается интегрировать традиционную рекламу в свою модель доходов, платформа не предлагает такого широкого спектра маркетинговых возможностей, как многие другие социальные сети. Но это не означает, что приложение не имеет огромной маркетинговой ценности для предприятий, использующих его в качестве средства коммуникации.

Более 70% предприятий в Казахстане активно используют Instagram в качестве маркетингового канала. Instagram предлагает невероятные показатели вовлеченности по сравнению с другими социальными сетями. Несмотря на относительно схожую функциональность и механизмы вовлечения, посты в Instagram в среднем имеют на 23% более высокий уровень вовлеченности, чем изображения, опубликованные в других социальных сетях [9].

#### Заключение

Дальнейшее развитие цифровой экономики в Казахстане принесет стране дополнительные экономические преимущества. Развитие электронной коммерции способно стимулировать деловую активность путем предоставления потребителям дополнительных каналов к предприятиям малого и среднего бизнеса. Цифровая повестка дня становится обязательным элементом любой стратегии, выдвигаемой любой компанией в наши дни.

Для компаний во всех отраслях казахстанской экономики открываются реальные возможности, и в этом направлении стоит активно поработать.

Рынок цифровой рекламы основан на нисходящем моделировании, основанном на экономической мощи местоположения, измеряемой его валовым внутренним продуктом. Сегмент рекламы в социальных сетях разделен на категории для персональных компьютеров и мобильных устройств.

Различные функции и особенности делают определенные платформы социальных сетей более подходящими для одного бренда, чем для другого. В то же время различное наследие и культуры также привлекают определенные демографические группы аудитории.

В заключении можно констатировать динамичный рост пользователей сети, а как следствие есть необходимость в продвижении товаров и услуг компании через социальные сети. В Казахстане есть множество проблем, мешающих развитию интернет-маркетинга: большая территория и низкая плотность населения, низкое количество квалифицированных специалистов в области Интернет-маркетинга и т.д. Но в связи с увеличением количества пользователей сети появилась возможность продвижения бренда с помощью Интернет-маркетинга.

#### Использованные литературы

1. Appel G. et al. The future of social media in marketing //Journal of the Academy of Marketing Science. – 2020. – Т. 48. – №. 1. – С. 79-95
2. Sedgwick R. et al. Social media, internet use and suicide attempts in adolescents //Current opinion in psychiatry. – 2019. – Т. 32. – №. 6. – С. 534
3. Felix R., Rauschnabel P. A., Hinsch C. Elements of strategic social media marketing: A holistic framework //Journal of Business Research. – 2017. – Т. 70. – С. 118-126.
4. Ольшевский Д. SMM-продвижение как эффективный инструмент интернет-маркетинга //Наука и инновации. – 2017. – Т. 9. – №. 175.



5. HOOTSUITE W. E. A. R. E. S. Y. Digital 2022. Global Digital Overview. – 2022.
6. Statcounter G. S. Search engine market share worldwide. – 2022.
7. Sairambay Y. Internet and Social Media Use by Young People for Information about (Inter) National News and Politics in Russia and Kazakhstan //Studies of Transition States and Societies. – 2022. 10
8. Statista.com. Number of smartphone users worldwide from 2017 to 2022 (in billions). – 2022. 12
9. Орынбасарова А. Т. Гражданско-правовое регулирование интернет-рекламы в Республике Казахстан. – 2021. 13

## АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР ТЕРМИНДЕРІН МЕМЛЕКЕТТІК ТІЛГЕ АУДАРУ ЖӘНЕ ЖҮЗЕГЕ АСЫРУ

*Бөрібаев Б.Б.<sup>1</sup> т.ғ.к., доцент, Кенесова Н.А.<sup>2</sup> магистрант*  
*<sup>1</sup>Әл-Фараби атындағы Қазақ Ұлттық университеті, Алматы қ.*  
*E-mail: [b.buribaev@gmail.com](mailto:b.buribaev@gmail.com)*

*<sup>2</sup>Әл-Фараби атындағы Қазақ Ұлттық университеті, Алматы қ.*  
*E-mail: [nazirakenessova@gmail.com](mailto:nazirakenessova@gmail.com)*

### **Аңдатпа**

Зерттеудің мақсаты – ақпараттық технологиялар терминдерін мемлекеттік тілге аудару және жүзеге асыруда қолданылатын әдістерге шолу жасау және термин мағынасын сақтай отырып дұрыс аудармасын беретін әдісті таңдау. Баяндамада электрондық сөздіктерде машиналық аударуда қолданылатын әдістер қарастырылады: тікелей аударма әдісі, жанама аударма әдісі, кері аударма әдісі. Олардың ішінде тікелей аударма мен жанама аударма бірнеше ішкі әдістерден тұрады және ол әдістердің әрқайсысына тоқталып кеттік. Бұл әдістерден бір ғана терминологиялық сөз тіркесінің әр түрлі әдістерді қолдану арқылы, бірнеше мағыналы аудармасын ала алатынымыз көрінді. Жұмыстың ғылыми жаңалығы- қазіргі қарқынды дамып келе жатқан Ақпараттық технологиялар саласындағы терминдер мен сөз тіркестерінің электрондық сөздігінде қолданылатын әдістерді анықтау. Нәтижесінде тікелей аударма мен кері аударма әдістері – Ақпараттық технологияларда пайдаланылатын терминдер мен сөз тіркестерінің аудармаларының бастапқы мағынасы мен мәнін сақтайтын әдістер екендігі анықталды.

**Түйін сөздер:** ақпараттық технологиялар, сөз тіркестері, тікелей аударма, жанама аударма, кері аударма, сөзбе-сөз аударма, электрондық сөздік.

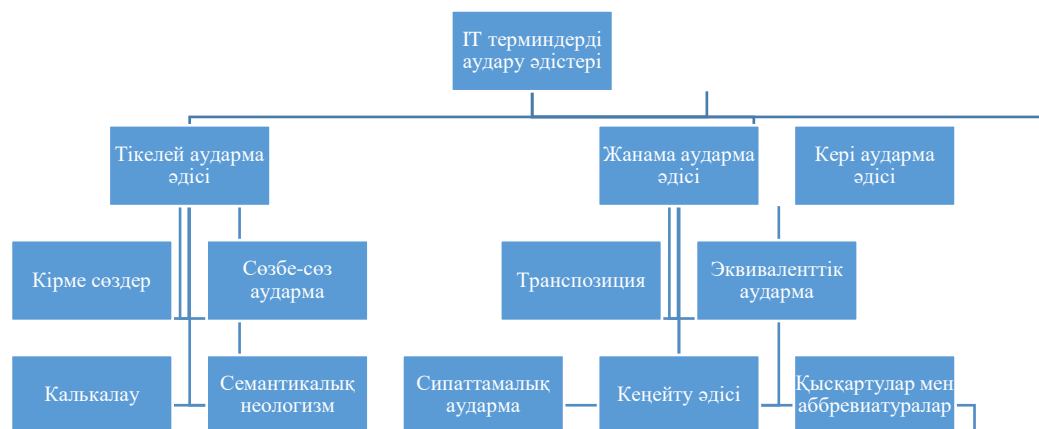
### **Кіріспе**

Әрбір саланың өзіндік міндеттерін шешу үшін қажет болатын қасиеттеріне байланысты терминнің мазмұны әрбір салада әртүрлі айқындалады. Терминді не ұғымды сөзбен белгілеу кезінде өзге ұғымдардан ажырататын ұғымы нақтылануы және саладағы қамтитын шекарасы анықталуы, терминнің дефинициясы дәл берілуі керек. Әйтпегенде, ұғымның мазмұны да толық ашылмай қалады. Сондай-ақ термин мағынасының нақты берілмеуі мәтіннің коммуникативтік функциясының бұрмалануына алып келеді [1]. Термин – бұл арнайы ұғымды дәл білдіруге қызмет ететін және анықтамаға негізделген арнайы қолдану аймағының арнайы сөзі немесе тіркесі. Кеңестік және ресейлік лингвист, терминология саласындағы маман С. В. Гринев-Гриневич терминнің келесі қасиеттерін анықтайды:

- қолдану ерекшелігі;
- бір мәнділігі;
- дәлдігі;
- бейтараптығы;
- нормативтік сипаты;
- түсініктемесін беру қажеттілігі [2].

### **Зерттеу әдіснамасы**

Ақпараттық технологиялар терминдердің ерекшелігі – ақпараттық технология ғылымымен өзара әрекеттесу процесінде қолданылатын арнайы ұғымдарды мүмкіндігінше дәл білдіру болып табылады. IT терминдерді аударудың ерекшелігі- басқа тілдегі баламаларды үнемі іздеу керек, өйткені қарапайым білім жеткіліксіз, мысалы, көркем әдебиетті аудару. Осы ерекшеліктерді ескере отырып ақпараттық технологиялар терминдерін аударудың келесідей әдістерін қарастыруға болады [4]:



Сурет 1. IT терминдері мен сөз тіркестерінің электрондық сөздігінде қолданылатын әдістер диаграммасы

1. Тікелей аударма әдісі түпнұсқа мәтін екі тілде де ұқсас ұғымдардың немесе бірдей категориялардың болуына байланысты шет тіліне дұрыс аударылған кезде қолданылады. Тікелей аударма әдісінің түрлері кесте 1-де берілген:

Кесте 1 Тікелей аударма әдістері

Кіріме сөздер	Кіріме сөздер – бұл сөздердің немесе тіркестердің бастапқы мәтіннен тікелей алынып, аударма тілге ауысуы. Бұл әдіс көбінесе мақсатты тілдің баламасы болмаған кезде қолданылады және бастапқы мәтіннің мағынасы мен контекстін сақтауға көмектеседі. Мысалы: компьютер, сервер [5].
Сөзбе - сөз аударма	Сөзбе - сөз аударма дегеніміз-түпнұсқа тіліндегі сөздердің реті сақталатын және сөздер контексті ескерместен тек кең мағынада аударылатын аударма. Сөзбе-сөз аударма, аударылған мәтіннің ойларын дұрыс жеткізе отырып, түпнұсқаның синтаксистік құрылымы мен лексикалық құрамын мүмкіндігінше көбейтуге тырысады. Бірақ бұл әдіс бастапқы мәтіннің нюанстарын жіберіп алуы мүмкін. Мысалы: Хранилище данных – деректер қоймасы [6].
Калькалау	Терминнің немесе сөз тіркесінің комбинаторлық құрамын жаңғырту, онда терминнің немесе сөз тіркесінің құрамдас бөліктері кейіннен аударылған бөліктерді қандай да бір өзгеріссіз қоса отырып, аударма тілінің тиісті элементтерімен аударылады [7].
Семантикалық неологизм	Семантикалық неологизм – аудармашы ойлап тапқан және тілдік бірліктің семантикалық мазмұнын жеткізуге мүмкіндік беретін жаңа сөз немесе сөз тіркесі. Бұл әдіс калькуляциядан түпнұсқа сөзбен этимологиялық байланыстың болмауымен ерекшеленеді.

Ақпараттық технология терминдерін неологизмдерге және неолексемдерге жатқызуға болады, себебі IT терминдері – үнемі жаңартылып, толықтырылып отыратын сала. Неологизм – тілімізде жақында пайда болған сөз немесе сөз тіркесінің мағынасы. Неолексемдер сөзжасам процестерінің нәтижесі болып табылатын жаңа сөздерден тұрады. Сөзжасамдық неологизмдерге сөзжасамның дәстүрлі модельдері бойынша құрылған неологизмдер жатады. Солардың бірі, аббревиатуралар: деректер қорын басқару жүйесі; күрделі сөздер: айти-евангелист; қарапайым туынды сөздер: антивирус, видеоредактор [8].

2. Жанама аударма әдістері екі тіл грамматикалары бір-бірінен алшақ болған кезде қолданылады. Бұл әдістер туралы толық сипаттама кесте 2-де берілген:

Кесте 2 Жанама аударма әдістері

Транспозиция	Транспозиция - мағынасын сақтай отырып, бір грамматикалық категориядан екіншісіне ауысуды қамтиды. Бұл аударма әдісі әр түрлі грамматикалық құрылымы бар тілдер арасында жиі қолданылады [9].
Эквиваленттік аударма	Эквиваленттік аударма - басқа тілде толық сәйкес келетін сөздер немесе сөз тіркестері. Эквиваленттер толық (шет тіліндегі сөздің мағынасын толығымен жабады) және ішінара (сәйкестік мағыналардың біреуіне ғана қатысты); абсолютті – сол функционалды стильге жатады және бастапқы тілдің сөзімен бірдей экспрессивті функцияға ие және салыстырмалы – мағынасына сәйкес келеді. Барлық лексикалық сәйкестіктерді екі негізгі топқа бөлуге болады: эквиваленттер және түрлі сәйкестіктер. Мәтіндердің толық эквиваленттілігіне қол жеткізудегі ең айқын кедергі, әрине, сөздердің көп мағыналылығы. Негізінен терминдер құрайтын салыстырмалы түрде аз сөздерді қоспағанда, тілде бір мағыналы сөздердің іс жүзінде жоқ екені белгілі. Бірақ бұл дегеніміз, терминдер тек бірмағыналы сөздер дегенді білдірмейді [10].
Сипаттамалық аударма	Сипаттамалық аударма – аударма тіліне баламалардың болмауы салдарынан тілдік бірлікті берудің басқа мүмкіндігі болмаған жағдайларда қолданылатын аударма әдісі.
Қысқартулар немесе аббревиатуралар	Қысқартулар немесе аббревиатуралар. Қысқартылған кезде аудармашы аударма тілінде артық деп саналатын бастапқы мәтінді құрайтын кез-келген сөздерді жояды. Аббревиатуралар IT саласындағы ғылыми-техникалық мақалалардың құрамдас бөлігі болып табылады. Мысалы, көптеген оқулық пен сөздіктерде айнымалылар мен мәннен var және val, const – тұрақты, params – params сияқты аббревиатураларды таба аласыз.
Кеңейту әдісі	Қысқартудың керісінше, мағынаны сақтау үшін сөздер қосылған кезде. Бұл сөйлем құрылымындағы, грамматикадағы немесе терминологиядағы айырмашылықтарға байланысты болуы мүмкін.

3. Кері аударма әдісі дегеніміз – термин мен сөз тіркесі берілген мақсатты тілге аударылып, содан кейін бұл аударма бастапқы тілге аударылатын аударма процесі. Көптеген лингвистикалық бағыттарда түпнұсқа мен аударма мәтіндерін салыстырудың дәстүрлі әдісі өте қиын және әрдайым зерттеудің ең көрнекі құралы бола бермейді. Осы кемшіліктерді жою мақсатында біз дереккөздерді салыстыру құралы ретінде кері аударманы, яғни мәтінді аударма тілінен түпнұсқа тіліне аударуды ұсынады.

### **Зерттеу нәтижелері**

IT терминдері мен сөз тіркестерінің электронды сөздігінде қолданылатын әдістерді қарастыра отырып, ең ыңғайлы және дұрыс әдіс-сөзбе-сөз аударма және кері аударма әдісі екенін атап өтуге болады. Бұл кері аударма әдісі Тікелей және жанама аударменің қасиеттерін қамтиды, ал тікелей аударма терминдердің дәлдігін сақтауға мүмкіндік береді. Кері аударма, сонымен қатар қос аударма деп те аталады, бұл мазмұнды аударма тілінен бастапқы тілге аудару процесі. Кері аударманың мақсаты лексикалық, синтаксистік және грамматикалық деңгейлердің терминдері мен сөз тіркестерінің максималды баламасына қол жеткізу болып табылады, бұл тапсырманы орындау үшін кері аударманы орындау кезінде әр деңгейдің тілдік құралдарын таңдау принциптері мен ережелерін жасау қажет. Аударылған мәтіннің екінші аудармасында біз сөзбе-сөз немесе сөзбе-сөз аударманы қолданамыз.

## Қорытынды

Ақпараттық технологиялар терминдерін қазақ тіліне аударғанда, түсініктемесін бергенде кері аударманы қолдану терминдер мен сөз тіркестерінің мағынасының сақталып, түрлі тәржімелерін алуға мүмкіндік береді. Себебі, түпнұсқа мен аударма мәтіндерін салыстыру арқылы сөз тіркестері мен түсініктеме мағынасы сақталып қалады. Ақпараттық технологиялар ғылымындағы терминдер мен сөз тіркестерінің электрондық сөздігінде қолдануға ыңғайлы 2 негізгі әдіс талданып, салыстырылып, іріктеліп таңдалынды: тікелей аударма (кірме сөздер мен сөзбе-сөз аударма) және кері аударма әдістері. Нәтижесінде, осы әдістерді қолдану арқылы “IT.KZ” электрондық сөздігі құрылды.

## Пайдаланылған әдебиеттер тізімі

- 1 Бисенгали А.З. Процесс терминообразования в казахском и турецком языках. – Алматы: Казахский национальный университет им. аль-Фараби, 2017, с. 96–99. <https://expeducation.ru/ru/article/view?id=11700>
- 2 Turganbayeva, A., Tukeyev, U. The solution of the problem of unknown words under neural machine translation of the Kazakh language. // Journal of Information and Telecommunication. 2021. с. 214-216. <https://doi.org/10.1080/24751839.2020.1838713>
- 3 Гарипова А.А. Структурно-семантические особенности терминологии интернета татарского языка. – Елабуга, Россия: Елабужский институт Казанского федерального университета, 2020, с. 88-90. <https://cyberleninka.ru/article/n/strukturno-semanticheskie-osobennosti-terminologii-interneta-tatarskogo-yazyka>
- 4 Ефремова Л. С., Лашкова Г. В. Структурные особенности терминов сферы IT, содержащие имена собственные (на материале английских терминологических словарей). – Саратов: Саратовский национальный исследовательский государственный университет имени Н. Г. Чернышевского, 2017, с.274. <https://cyberleninka.ru/article/n/strukturnye-osobennosti-terminov-sfery-it-soderzhaschie-imena-sobstvennye-na-materiale-angliyskih-terminologicheskikh-slovarей>
- 5 Rakhimova, D.R., Turganbaeva, A.O. Normalization of kazakh language words. // Scientific and Technical Journal of Information Technologies, Mechanics and Optics. 2020.с. 545-546. <https://doi.org/10.17586/2226-1494-2020-20-4-545-551>
- 6 Зорина Д.В., Тимохов В.В., Бурматова К.А., Янковская Н.С., Проноза Е.В. IT термины и профессиональный сленг в современных сообществах социальных сетей. – Москва, Россия: Новые информационные технологии в автоматизированных системах, 2018, с. 229-230. <https://cyberleninka.ru/article/n/it-terminy-i-professionalnyy-sleng-v-sovremennyh-soobschestvah-sotsialnyh-setey>
- 7 Саметова Ф.Т. Лексические инновации в современном казахском языке. – Алматы: Евразийский Союз Ученых, 2017, с. 67-69. <https://book.ru/book/933858>
- 8 Sennrich R., Haddow B., Birch A. Neural Machine Translation of Rare Words with Subword Units. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. – Berlin, Germany: Association for Computational Linguistics, 2016, с.1715–1725. <https://aclanthology.org/P16-1162/>
- 9 Caswell I., Chelba C., Grangier D. Tagged Back-Translation. In Proceedings of the Fourth Conference on Machine Translation. – Florence, Italy: Association for Computational Linguistics, 2019, с. 53–63. <https://aclanthology.org/W19-5206/>
- 10 Tacorda, A.J., Ignacio, M.J., Oco, N., Roxas, R.E.: Controlling byte pair encoding for neural machine translation. In: 2017 International Conference on Asian Language Processing. 2017, с. 168–171. <https://doi.org/10.1109/IALP.2017.8300571>

## КИІМ ТОПТАМАСЫНЫҢ ЭСКИЗДЕРІН ЖОБАЛАУДАҒЫ КОМПЬЮТЕРЛІК ГРАФИКАНЫҢ ҚОЛДАНЫЛУЫ

А.Е Байбатшаева - п.ғ.к., доцент м.а Шымкент қ., Қазақстан. ОҚМПУ «Өнер және білім» факультетінің деканы. [1aidai@mail.ru](mailto:1aidai@mail.ru)

З.А. Иманкулова Шымкент қ., Қазақстан. ОҚМПУ «Дизайн» б.б.б магистранты 2 курс, Ә.Қастеев атындағы өнер және дизайн колледжі, арнайы пәндер оқытушысы.

[imankulova-79@bk.ru](mailto:imankulova-79@bk.ru)

Ж.Ж. Карабаева Шымкент қ., Қазақстан. №86 ЖОББМ, «Көркем еңбек» пәнінің мұғалімі. [Zhaz.63@mail.ru](mailto:Zhaz.63@mail.ru)

Түйіндеме: Мақалада жаңа киім үлгісін жобалаудың басым бағыттарының бірі ретінде костюм дизайнын ұсынудың компьютерлік мүмкіндіктері қарастырылады. Компьютерлік технологияның дамуы сән индустриясының және жалпы білімнің дамуына көмектесетін компьютерлік графиканың көмегімен коллекцияны көрнекі бейнелеу үлгісін жасау процесіне мүмкіндік береді.

Резюме: В статье рассматриваются компьютерные возможности подачи дизайна костюма в качестве одного из приоритетных направлений проектирования новой модели одежды. Развитие компьютерных технологий позволяет осуществлять процесс разработки модели для визуального представления коллекции средствами компьютерной графики, что помогает развитию индустрии моды и образования в целом.

Тігін өнеркәсібі үшін техникалық жабдықтардың дамуы, компьютерлер мен жаңа ақпараттық технологиялардың өмір мен өндірістің барлық салаларына енуі ақпараттық қоғамның даму деңгейіне сәйкес мамандарға деген сұранысты анықтайды. Дизайн саласында іргелі білім алатын студенттер үшін компьютерлік технологиялар және ақпаратты талдамалық өңдеу саласындағы білім қажет, бұл өзгермелі ортада белсенді шығармашылық пен бейімделуге ықпал етеді. [1]

Дизайнерлік қызметте тілдік бейнелеу ойы үлкен маңызға ие. CorelDraw және PhotoShop графикалық пакеттерінде эскиздерді әзірлеу мен түрлендіру студенттің көркем және конструктивті-технологиялық деңгейін практикалық және нысанды түрде шығаруға, оның өзін өзі бағалауын түсінуге негізделеді. Осы пән бойынша практикалық сабақтар костюмді әр түрлі стильде көркем безендірумен және техникалық эскиздерді орындаумен байланысты. Бұл білім болашақ мамандарға болашақ кәсіби қызметінде қажет болады. Костюмдер дизайнындағы компьютерлік технологиялар «Дизайндағы ақпараттық технологиялар» пәнімен тығыз байланысты. «Сурет», «Кескіндеме», «Перспектива», «Тұстану», «Тігін бұйымдарын құрастыру», «Тігін бұйымдарының технологиясы», «Материалтану» пәндері бойынша арнайы білімді қажет етеді

Біз жоғары технологиялар ғасырында өмір сүрудеміз, олар біздің өмірімізбен тығыз байланысты әрі кез келген қызмет саласында маңызды. Олар дамып, жаңашылданып, қандай да бір жаңа өзгерістерге ұшырап отырады. Олардың біреуі бір кездері графикалық редакторға айналды. 1950 жылдары пайда болған компьютерлік графика белгілі бір практикалық қолданбалары жоқ «таза ғылым» ретінде біраз уақыт өмір сүрді. Сол кездегі кесте «белгісіз аурудың емі», «проблеманы іздеудегі шешім» деп аталды. Компьютерлік графиканың өзіне тән функциялары бар. Бірақ бұл функциялар туралы идеялар өте түсініксіз, біржақты немесе тіпті қате болып шығады. Қазіргі кезде компьютерлік графика қарапайым жоспарлардан табиғи объектілердің шынайы кескіндеріне дейінгі суреттерге арналған аппараттық және бағдарламалық қамтамасыз ету туралы ғылым ретінде пайда болды. Бүгінгі күні ол барлық ғылыми және инженерлік пәндерде ақпаратты беру мен қабылдау үшін қолданылады. Бірақ бұл ең алдымен дизайн мен бейнелеу өнерін қолдауға бағытталған: кескіндеме, сурет,

декор, көркем және кітап графикасы, мүсін және әрине сән дизайны. Сонымен қатар, оларды ауыстыру емес, толықтыру және дамыту. Ал «компьютерлік графика» термині нені білдіреді?

Компьютерлік графика деп әдетте компьютердің көмегімен графикалық ақпаратты дайындау, түрлендіру, сақтау және көбейту процестерін автоматтандыру деп түсінеді. Графикалық ақпарат объект модельдеріне және олардың кескіндеріне жатады. Интерактивті компьютерлік графика сонымен қатар кескіндерді дайындау және көбейту үшін компьютерлерді қолдану болып табылады, бірақ сонымен бірге пайдаланушы кескінге оны көбейту кезінде тез өзгертулер енгізуге мүмкіндігі бар. Біз нақты уақыт, диалог режимінде графикамен жұмыс жасау мүмкіндігі туралы айтып отырмыз.

Пайдаланушы интерактивті басқару құралдарын пайдаланып кескіннің мазмұнын, оның пішінін, өлшемін және түсін дисплей бетіне түсіруді динамикалық түрде басқаруға мүмкіндігі болған кезде интерактивті графика компьютерлік графиканың маңызды бөлімі болып табылады. Дизайнердің қызметі адамның айналасындағы бүкіл әлемді түрлендіруге және үйлестіруге бағытталған. Дизайнердің қызметі - қоршаған әлемді өзгерту мақсатын алға қоятын жобалық қызмет. Алдымен өзгеріс дизайнердің қиялында пайда болады, содан кейін ол графикалық тілде немесе макетте жүзеге асырылатын жобада жүзеге асырылады. Бүгінгі таңда графикалық жобалардың көпшілігі компьютерлік нұсқада ұсынылған. Егер костюмдер дизайнындағы компьютерлік графика туралы айтатын болсақ, онда графиканың растрлық және векторлық түрлері ыңғайлы. Оларды пайдалана отырып, сіз компьютерде киімнің жаңа моделін жасаудың бүкіл процесін - эскизден бастап жұмыс құжаттамасына дейін жүргізе аласыз. Бірақ біз костюмді жобалаудың ең қызықты және креативті кезеңіне тоқталамыз, онсыз одан әрі белсенділік мүмкін болмайды - эскиз жасау кезеңі. Эскиз - бұл өнер туындысының идеясын бекітетін алдын-ала эскиз-дизайнердің ойының бейнесі, бұл болашақ костюмнің ашық тізбектелген бейнесі, сондықтан ол барлық бөлшектерді көрсетуі керек, сонымен ерекшелінуі керек. Мұның бәріне «қолмен» сурет салу арқылы қол жеткізуге болады. Бірақ әртүрлі графикалық бағдарламалар сызбаларды жаман емес, кейде қолмен салғаннан да жақсы жасауға мүмкіндік береді. Сонымен қатар, киімнің жаппай өндірісінде үнемдеу және уақыт тұрғысынан компьютерде жасалған эскиздер басымырақ. Шынында да, егер біз дайын киімді жобалаумен айналысатын кәсіби дизайнерлердің қызметін қарастыратын болсақ, онда бұл жағдайда эскиз жасау процесі өте қысқа мерзімде жүргізілуі керек. [2]

Әр модель үшін барлық мүмкін түсті шешімдерді әзірлеу, материалдың текстурасын және оның барлық мүмкін нұсқаларын беру қажет. Техникалық сурет туралы ұмытпаңыз, онда кесудің барлық бөлшектері нақты белгіленіп, бұйымның дизайны жақсы оқылуы керек. Енді елестетіп көріңізші, дизайнер қолмен жұмыс істеуге қанша уақыт жұмсайды, оған қанша құрал мен материалдар қажет болады. Қазіргі компьютерлік графика сурет салудың қолданыстағы әдістерін алмастыра алмайтындығына назар аударайық. Бұл костюмдер дизайны саласындағы көрнекі әрекеттерді жүзеге асырудың нақты түрі. Бұл тәсілдің қағидалары сызбаның жүйелік байланысы және графикалық белгілердің өзара әрекеттестігі сияқты аспектілері мен принциптерін одан әрі кеңейту мен байытуды білдіреді; тілдік ойлаудың, оның ішінде мағыналық және стилистикалық гетерогенді элементтерді дамытудың жаңа жолдарын көрсету; костюм бейнесі әлемінің өзгеруі және визуалды мәдениет деңгейінің жоғарылауы.

Компьютерлік бағдарламалардың мүмкіндіктері адам мүмкіндіктеріне қарағанда әлдеқайда кең. Бір модельдің бір-бірінен қолданылған маталармен және декоративті-функционалды элементтерімен (қалталар, жағалар, манжеттер және т.б.) ерекшеленетін бірқатар модификацияларын жасау қысқа мерзімде орындалуы мүмкін және тек компьютері бар бағдарлама қол жетімді. Сонымен қатар, заманауи

технологияларды қолдана отырып, киім үлгісінің эскизін жасау кезінде «қолмен» сурет салудың барлық артықшылықтарын сақтау және компьютерлік графиканың барлық бай мүмкіндіктерін дұрыс қолдану маңызды.

Растрлық графиканы өңдеуге арналған бағдарламалардың кең класында Adobe-дің Photoshop пакеті ерекше орын алады. Шындығында, бүгінде бұл компьютерлік графикада стандарт болып табылады және барлық басқа бағдарламалар онымен үнемі салыстырылады. Corel DRAW, қазіргі Corel Graphics Suite негізін құрайтын графикалық редактор 2002 жылы шығарылды.

Бағдарламаны кез-келген зерттеу қолданушы интерфейсінен басталады. Corel Draw-де ол өте тиімді, эргономикасы жоғары деңгейде салынған. Бұл жалпы графиканы немесе векторлық графиканы зерттеуде жаңадан бастаушыларға арналған алғашқы бағдарламалық құрал ретінде бағдарламаны өте тартымды етеді, сонымен қатар тек бір қызмет түріне бағыттауға көмектеседі. Сонымен, осы бағдарламалардың бірінде эскизді қалай жасауға болады? Эскиздің негізгі сызықтарын анықтаудың ең дұрыс әдісі - қарындаш эскизін жасау, содан кейін оны компьютерге енгізу. Осы суреттің негізінде сіз әртүрлі киім модельдерінің көптеген эскиздерін жасай аласыз. Ол үшін фигураның бейнесін бір қабатқа, ал киімді екінші қабатқа орналастыру ыңғайлы. [3]

Егер бұл жағдайда жиынтықтың бөліктері де бөлек қабаттарға орналасса, онда сіз сол немесе басқа қабатты сызыққа көрінетін етіп, ансамбльдің әр түрлі нұсқаларын ала аласыз. Әдетте, эскиз элементтерінің контурлық сызықтары жойылады немесе объектінің бояу түсімен толтырылады. Алайда, кейбір жағдайларда бұл контур сызықтары модельде ойластырылған әрлеуді білдіруі мүмкін. Аяқтаудың контурлық сызықтарының түсі мен қалыңдығын өзгерту арқылы сіз қызықты композициялық шешімдер ұсына аласыз.

#### **Эскиздік қатарды құру**

#### **Киім үлгісінің жиынтығын жасаудағы алгоритм**

1. Киім жинағының тақырыбын анықтау
2. Визуальді ақпараттарды жинақтау (сән журналдарынан суреттер, суреттер)
3. Болашақ жиынтықтың силуэтін анықтау
4. Киім жиынтығының түстік гаммасын анықтау
5. Материалдар таңдау
6. Графикалық редакторда жұмыс істеу

Ақпараттарды ізденіс жұмыстары мен дайындық жұмыстарын орындап болған соң, CorelDraw бағдарламасында болашақ киім жиынтығының эскиздерін салу жұмысын бастайды.

#### **Ұсынылған алгоритмді орындау нәтижесінде дайын киім жиынтығы шығады.**

Киімдер жиынтығын АКТ көмегімен жобалау кезінде оның жағымды және жағымсыз жақтары да бар. Жағымды: - барлық киім жиынтығында және әр модельде бөлек композициялық құрылымның өзгергіштігі;- модельдердің пропорциясы мен силуэтін өзгерту мүмкіндігі; - сәндік элементтерді қосу немесе алып тастау мүмкіндігі; - әртүрлі маталар фактурасын қолдану мүмкіндігі. Жағымсыз: - заманауи графикалық редакторлардың құралдары мен фактуралары өте көп (қылқалам, қарындаш, аэрограф және т.б.), экран қағаз бетін толық алмастыра алмайды, бұл ең алдымен қандай да бір ақпаратты жүйелейтін кеңістік. Ұсынылған тәжірибе қосымша білім беру жүйесінде де, кәсіптік білім беру жүйесінде де қолданылады.

#### **CorelDRAW графика өнері, бағдарламасында киім эскиздерімен жұмыс жасау.**

Компьютерлік технологиялар жобалау қызметі үшін үлкен маңызға ие. Костюм дизайнындағы компьютерлік технологиялар ең алдымен киім дизайны үшін компьютерлік дизайн жүйелерінде қолданылады. Қазіргі заманғы САПР жүйелері өлшемдік сипаттамаларға негізделген негізгі құрылымды жобалауға мүмкіндік береді;



сызбаны конструктивті модельдеу; бұйым бөлшектерінің үлгілерін безендіру; киімнің бөлшектеріне арналған өрнектерді градациялау; бөлшектердің өрнектерінің макеттерін құру; жобалық-технологиялық құжаттаманы тіркеу; материал шығынын есептеу.

Қазіргі заманғы САПР жүйелерінің көпшілігі техникалық және көркем эскизді жасауға, модель үшін түс схемасын таңдауға және т.с.с. мүмкіндік береді, тіпті Corel Draw, Illustrator және Photoshop стандартты графикалық бағдарламаларын қолдана отырып, дизайнер киім стильдерінің эскиздерін, киім үлгілерінің үлгісі, құрылысы және градациясы, үнемдеу үшін кесуге арналған суреттер. Сондай-ақ, бөлшектер мен бұйымдардың белгілі бір түрлері мен формаларының архивін жасауға болады.

Киімнің жаңа моделін жобалау кезінде дизайнерлер, конструкторлар, технологтар арасында түсінбеушіліктер туындамас үшін оның бейнесін бірнеше интерпретация жасайды: шығармашылық нобай; техникалық сурет. [3]

Костюм дизайнерлері, барлық дизайнерлер сияқты, компьютерлік технологиялар мен өзгермелі ортада маманның шығармашылық қабілеті мен бейімделуіне ықпал ететін ақпаратты талдаумен өңдеу әдістерін жетік білуі керек. Костюм дизайнындағы компьютерлік технологиялар - бұл көркемдік-техникалық нобай жасаудан, модель құрылымын құрудан корпоративті символизм мен брендтік жарнаманы дамытуға дейінгі жобалық қызметтің барлық кезеңдеріндегі әмбебап құрал.

#### Қолданылған әдебиеттер тізімі

1. Ермилова Д.Ю., Ермилова В.В., Ляхова Н.Б., Попов С.А. Композиция костюма // Учебное пособие / Москва, 2019. Сер. 58 Бакалавр. Академический курс (3-е изд., испр. и доп).
2. Ермилова В.В. Моделирование и художественное оформление одежды: учеб. пособие для сред. проф. образования / В.В. Ермилова, Д.Ю. Ермилова. М. : Академия : Московские учебники, 2010.
3. Попов С. А. Роль компьютерных технологий в дизайне костюма // Актуальные исследования. 2020. №9 (12). Ч.І. С. 108-111. URL: <https://apni.ru/article/760-rol-kompyuternikh-tekhnologii-v-dizajne-kost>

# USING SENTIMENT ANALYSIS TO ASSESS THE INVESTMENT ATTRACTIVENESS OF AN ORGANIZATION

Alikul Sholpan Maratkyzy – MA  
al-Farabi Kazakh National University, Almaty  
[sholpanalikul@gmail.com](mailto:sholpanalikul@gmail.com)

**Abstract.** The article discusses the main theoretical prerequisites for the development of methods of Sentiment Analysis and the practice of expanding their empirical application to the field of economic analysis. This linguistic computational approach to texts is based on the consideration of human emotions. This approach based on the analysis and assessment of the text emotionality based on sentiment dictionaries.

**Keywords:** investment, assessment, sentiment analysis, tonality, text tonality processing.

## Introduction

Corporate sentiment, often known as business sentiment, describes how others view your brand. In financial words, it pertains to the level of confidence investors have in a company or organization; in business terms, it shows the general view or perspective on the success of that particular company (or lack thereof).

Reputation management stresses this facet of the industry by emphasizing the importance of the study of corporate emotions. Once you comprehend how people feel about your company, you will be better able to defend your reputation against attacks and cultivate a positive attitude toward your brand.

When we refer to consumer insights, we are referring to things like as what our customers want, what they like and dislike about our products, their purchase signals, how they make decisions, etc. When we refer to sentiment analysis, we mean the following.

As more content is produced and distributed online through Social Channels, Blogs, and other such platforms, we tend to become more forthright and vocal about our internet experiences. Recent Zendesk data indicates that 45% of people share negative customer service experiences on social media, whereas only 30% share positive customer service experiences. This again illustrates that both the need and willingness of enterprises to mine this information for business insights have increased.

## Literature review

The management of the vast majority of companies are unaware of the importance of increasing the investment attractiveness of their companies, and as a result, they do not prioritise issues such as improving the methods of investment attractiveness analysis and evaluation, researching the factors that influence investment attractiveness, or exploring the possibilities of increasing investment attractiveness. In this respect, the issue of investigating and assessing an organization's investment attractiveness receives insufficient attention in the scientific and specialised literature [1].

As a promising tool for assessing the investment attractiveness of an organization, at this stage of technology development, is the use of sentiment analysis [2]s.

Sentiment analysis is a relatively new area of research and has been slowly gaining popularity since the dawn of the digital age. Sentiment analysis (sentiment analysis) is often identified with Opinion mining (opinion extraction), believing that they represent a single broad concept in which different similar conceptual subcategories can be defined. Bo Pang and Lillian Lee, on the other hand, believe that they are two different subjects of study [3]. Sentiment analysis, compared to Opinion mining, deals with a narrower, human-emotional aspect of language. It is a computational study of the opinions and feelings expressed in a particular text, which measures the direction and subjectivity of behaviour resulting from that text. The object of Sentiment analysis is sentiment (tonality), i.e., any mood or cognitive-affective state expressed in words and sentences [4]. Given the possible subtle definitions of these two concepts and their content, it is important to emphasize the growing interest in the analysis of opinions and emotions

expressed in a written text; this interest is due to numerous factors, primarily the emergence of blogs, forums and social networks [5]. Indeed, the areas of practical application of sentiment analysis today are very heterogeneous. After the advent of the Internet, the habit of expressing one's own opinions, ideas and thoughts, actively expressing them in social networks regarding goods, services or events, has become more and more common. The data they generate is used by scientists, researchers, market operators, managers, journalists, politicians and a growing number of organizations [6].

The emotional coloring of the text, as a rule, is multidimensional, and powerful specially prepared dictionaries are required to reveal it [7]. The overall tone of a literary work is comprised of three components: the topic of the tone, the judgement of the tone, and the goal of the tone. The subject of the tonality category is the creator of the text, whereas the issue or person being discussed in the text is the object of the tonality category [8]. The assessment of the tone may be presented in any of the following formats: binary, which means positive or negative; ternary, which means positive, neutral, or negative; or ranking [9]. The search for lexical sentiments in the text can be carried out on the basis of pre-prepared tone dictionaries [10].

### Methods

A research prototype of a text sentiment analyser is provided. This analyser runs a multi-stage procedure in which the text is first separated into individual sentences and then into independent words. This is done so that the analyser can then be used to evaluate the financial attractiveness of a company using sentiment analysis. After that, a morphological analysis is performed on each individual word, and if the word is already present in the dictionary, its tone is noted. There is a correlation between each word in the dictionary and a tonal score. This indicator is made up of five different values all together. Each number represents the degree to which a given word belongs to one of the following categories: very negative, negative, neutral, positive, and positive to the extreme. The sum of all the other values for a particular word is equal to one. If a word is not included in the dictionary, then the tone of the term is neutral. Determine the overarching feeling conveyed by the statement by adding up the feelings evoked by its individual words. Certain words have more significance than others and have a stronger influence on the overall mood of the text. Calculating the values of both the positive and the negative component of an estimate is required to arrive at the conclusion. One must compute the sum of the good sensations evoked by all the positive words included inside a phrase in order to ascertain the positive component of a statement. In a similar manner, the importance of the unfavourable aspect of the text is evaluated. For the final assessment of the mood of the entire text, it is necessary to calculate the ratio of these components using the following formula:

$$S^* = \frac{\sum_{i=1}^N (S_{4_i} + S_{5_i})}{\sum_{i=1}^N (S_{1_i} + S_{5_i})},$$

where  $S^*$  – is the tonal evaluation of the supply;  $S = [S_1, S_2, S_3, S_4, S_5]$  – tonal evaluation of the word;  $S_1, S_2$  – negative main component of the tonal evaluation of the word;  $S_4, S_5$  – positive component of the word's tonal evaluation;  $N$  – is the number of words in the sentence. Value of  $S^*$  is compared with some value of  $T$ , which is calculated experimentally. The text in which the value of  $S^*$  is close to the value of  $T$ , will be considered neutral, if it exceeds the value of  $T$  - positive, if it is less than the value of  $T$  - negative.

At the next stage of the study, it is planned to conduct experiments to analyze the performance of the developed algorithms for assessing the investment attractiveness of an organization using sentiment analysis.

A rudimentary sentiment analysis system, analogous to how the brain remembers descriptive words encountered throughout life and the "sentiment weight" associated with each word, uses an emotion library to comprehend the sentiment-laden sentences it encounters. Sentiment libraries are enormous collections of meticulously evaluated words and phrases, such as "good," "awesome," "bad," and "terrible" (excellent game, great tale, awful performance, and horrible show). This manual grading of emotion is challenging since all participants must agree on how strong or weak each score should be relative to the other scores. If one individual gives a

sentiment score of -0.5 to the word "bad" and another assigns the same value to the phrase "awful," your sentiment analysis algorithm will infer that both expressions are equally unpleasant.

A sentiment analysis engine that supports several languages must also have distinct libraries for each supported language. In addition, continual maintenance is necessary for each of these libraries, including the modification of ratings, the addition of new words, and the deletion of irrelevant phrases.

Most languages adhere to a set of basic rules and patterns that may be recorded in a computer program and used to power a basic part-of-speech tagger.

The bulk of the time, nouns and pronouns are employed to denote known objects, whereas adjectives and adverbs are used to describe them in emotionally charged ways. When a sentiment analysis system recognizes combinations of adjectives and nouns, such as "terrible pitching" and "mediocre hitting," it recognizes an emotive phrase. This is because such combinations imply that the sentence has emotion. It should not come as a surprise that the adjective-noun combination is not always used in emotionally charged sentences. In the past, the phrase "cost us" was composed of a noun and a pronoun with a negative meaning.

Trustworthy sentiment analysis requires accurate part-of-speech tagging; hence, a rules-based system must account for these variances.

Due to its ease of use, rules-based sentiment analysis is a suitable alternative for basic document-level sentiment scoring of predictable text documents. Survey responses having a limited scope are one example. In contrast, the disadvantages of a rules-based approach to sentiment analysis outweigh the vast majority of its benefits. A rules-based system must have a rule for each word combination in its moods library in order to function properly. Both the creation and maintenance of these rules require substantial human effort. In addition, it is impossible to anticipate that rules would keep pace with the emergence of authentic human language. Instant messaging has corrupted traditional grammar norms, and no set of rules can account for every abbreviation, acronym, double meaning, and error that may arise in a specific text composition. In addition, instant texting has distorted grammar norms.

### **Discussion**

As seen in the last section, a rules-based system that disregards negators and intensifiers is inherently ineffective. A sentiment score at the document level may lead to erroneous conclusions if it is evaluated in isolation from its context. Lastly, a rule-based system for emotion analysis is very sensitive. The algorithm is unable to provide a score because it cannot identify anything new that has been added to the text document that is not already covered by the rules. Extremely rarely, the whole program may crash, in which case an engineer must painstakingly identify the cause of the problem before creating a new rule to fix it.

A rudimentary sentiment analysis system, analogous to how the brain remembers descriptive words encountered throughout life and the "sentiment weight" associated with each word, uses an emotion library to comprehend the sentiment-laden sentences it encounters.

Sentiment libraries are enormous collections of meticulously evaluated words and phrases, such as "good," "awesome," "bad," and "terrible" (excellent game, great tale, awful performance, and horrible show). This manual grading of emotion is challenging since all participants must agree on how strong or weak each score should be relative to the other scores. If one individual gives a sentiment score of -0.5 to the word "bad" and another assigns the same value to the phrase "awful," your sentiment analysis algorithm will infer that both expressions are equally unpleasant.

A sentiment analysis engine that supports several languages must also have distinct libraries for each supported language. In addition, continual maintenance is necessary for each of these libraries, including the modification of ratings, the addition of new words, and the deletion of irrelevant phrases.

Most languages adhere to a set of basic rules and patterns that may be recorded in a computer program and used to power a basic part-of-speech tagger.

The bulk of the time, nouns and pronouns are employed to denote known objects, whereas adjectives and adverbs are used to describe them in emotionally charged ways. When a sentiment analysis system recognizes combinations of adjectives and nouns, such as "terrible pitching" and "mediocre hitting," it recognizes an emotive phrase. This is because such combinations imply that the sentence has emotion.

### Conclusion

Businesses are seeking to extract the value buried inside text to better comprehend the ideas and wishes of their customers and make more informed business decisions. Companies used to conduct customer surveys, seminars, and focus groups to get a deeper understanding of their consumers' perspectives and emotions. With the advent of technology, however, we can now harness the power of Machine Learning and Artificial Intelligence to extract meaning from text and explore consumers' thoughts outside of the usually regulated context of a survey.

The level of satisfaction that a company's customers report may determine the company's level of success or failure. The experiences of the customers might either be satisfying or upsetting, or they can be unremarkable. Because they are now at an age when Internet literacy is common, the content of their social posts and online comments is derived from their own experiences. It is possible to determine the disposition and tone of this data, and after that, the data may be categorized according to the feelings that are linked with it. This makes it easier to determine which aspects of the company's products, services, and customer support are being successfully implemented, as well as which aspects need to be improved.

### References

1. **Ендовицкий Д. А., Бабушкин В. А., Батурина Н. А. и др.** (2010) Анализ инвестиционной привлекательности организации. – М.: КНОРУС. – 376 с.
2. **Pang B., Lee L.** (2008) Opinion Mining and Sentiment Analysis. – Foundations and Trends in Information Retrieval». – 2 (1–2). – С. 6.
3. **Nugumanova A., Bessmertnyi I.** (2013) Applying the latent semantic analysis to the issue of automatic extraction of collocations from the domain texts. – Communications in Computer and Information Science. – V. 394. – P. 92–101.
4. **Cruz F. L., Troyano J. A., Pontes B., Ortega F. J.** (2014) Building layered, multilingual sentiment lexicons at synset and lemma levels. – Expert Systems with Applications. – V. 41. – N 13. – P. 5984–5994
5. **Ермаков С. А., Ермакова Л. М.** (2012) Методы оценки эмоциональной окраски текста. – Вестник Пермского университета. Серия: математика, механика, информатика. – № 1. – С. 85–90.
6. **Parau P., Stef A., Lemnaru C., Dinsoreanu M., Potolea R.** (2013) Using community detection for sentiment analysis. – Proc. IEEE 9th Int. Conf. on Intelligent Computer Communication and Processing. – P. 51–54.
7. **Chiru C.-G., Hadgu A. T.** (2013) Sentiment-based text segmentation. – Proc. 2nd Int. Conf. on Systems and Computer Science. – P. 234–239.
8. **Caviggioli, F., Lamberti, L., Landoni, P., & Meola, P.** (2020). Technology adoption news and corporate reputation: sentiment analysis about the introduction of Bitcoin. *Journal of Product & Brand Management*.
9. **Ren, J., Ge, H., Wu, X., Wang, G., Wang, W., & Liao, S.** (2013, October). Effective sentiment analysis of corporate financial reports. In *Thirty Fourth International Conference on Information Systems, Milan*.
10. **Kim, M., Park, E. L., & Cho, S.** (2018). Stock price prediction through sentiment analysis of corporate disclosures using distributed representation. *Intelligent Data Analysis*, 22(6), 1395-1413.

## БІЛІМ БЕРУ ЖҮЙЕСІНДЕГІ САНДЫҚ ТЕХНОЛОГИЯЛАР

**Ибраева Қ.Қ., Досымова К.Б., Қожабай А.Б.**  
Қорқыт Ата атындағы Қызылорда университеті,  
Жаратылыстану институты,

«Информатика және ақпараттық-коммуникациялық технологиялар» кафедрасы  
магистранттары,  
Қызылорда қаласы, [kozhabai\\_akbota@bk.ru](mailto:kozhabai_akbota@bk.ru)

**Аңдатпа.** Компьютерлік технология адам қызметінің барлық салаларына еніп кетті. Бүгінгі таңда білім беру жүйесінде ақпараттық-коммуникациялық технологияларды қолдану қажеттілікке айналып отыр. Қазіргі білім алушы электронды мәдениет әлемінде өмір сүреді. Сондықтан ақпараттық мәдениеттегі мұғалімнің рөлі өзгеруде, ол ақпарат ағынының үйлестірушісі болуы керек. Жаңа техникалық құралдарды енгізу арқылы сабақтың кез-келген кезеңін жандандыруға болады.

**Кілт сөздер:** АКТ, цифрлық технология, веб-сервис, платформа

XXI ғасыр – жоғары компьютерлік технологиялар ғасыры болғандықтан компьютерлік технология адам қызметінің барлық салаларына еніп кетті. Бүгінгі таңда білім беру жүйесінде АКТ қолдану қажеттілік болып табылады. Қазіргі таңда білім алушылар электронды мәдениет әлемінде өмір сүреді. Сондықтан ақпараттық мәдениеттегі мұғалімнің рөлі өзгеруде, ол ақпарат ағынының үйлестірушісі болуы керек. Заман көшіне ілескен ұстаз қазіргі таңда АКТ-ны оқу үдерісінде қолдануға психологиялық және техникалық жағынан дайын және жаңа техникалық құралдарды енгізу арқылы сабақтың кез келген кезеңін жандандыруға болады.

Оқу-тәрбие процесіне АКТ-ны енгізу мұғалімге сабақта оқу-танымдық іс-әрекеттің әртүрлі формаларын ұйымдастыруға, білім алушылардың белсенді және өзіндік жұмысын жасауға жол ашады. Компьютерді барлық кезеңде, сабақты дайындауда да, оқу процесінде де: жаңа материалды түсіндіруде (енгізуде), бекітуде, қайталауда, бақылауда қолдануға болады.

**Білім беру мақсатында АКТ төмендегідей жіктеледі:**

- **Оқулықтар** меңгеруге болатын шағын бөліктерге жақсы құрылымдалған материалды зерттеуде қолданылады.
- **Басқару бағдарламалары** білім алушылардың білім, білік деңгейін анықтау үшін қолданылады. Басқару бағдарламаларының ерекше түрі болып табылады **тестілеу бағдарламалары** болып табылады. Тестілеу бағдарламаларымен жұмыс барысында білім алушы дұрыс жауапты немесе жауаптар тізімінен бірнеше жауаптарды таңдайды, белгілі элементтерден дұрыс жауапты құрастырады, сәйкестендіреді.
- **Оқыту бағдарламалары** дағдылар мен дағдыларды қалыптастыруға және бекітуге, сондай-ақ білім алушылардың өзін-өзі тексеруге арналған.
- **Демо бағдарламалар** компьютер арқылы жаңа материалды түсіндіруде көрнекі құрал рөлін атқарады, иллюстрация ретінде қолданылады.
- **Ақпараттық-анықтамалық бағдарламалар** компьютердің көмегімен білім алушы ақпаратты орталықтандырылған түрде сақтайтын және бөлісетін әртүрлі деректер банктеріне қол жеткізеді.
- **Имитациялық бағдарламалар** заттар мен құбылыстарды «имитациялауға» арналған. Оларды білім беру ортасындағы белгілі бір құбылысты шынайы түрде жүзеге асыру мүмкін емес немесе қиын болған жағдайда қолданылады.
- **Проблемалық оқыту бағдарламалары** - білім алушының іс-әрекетін жанама бақылауға мүмкіндік береді.

**Білім беру жүйесінде АКТ қолданудың негізгі 3 педагогикалық мақсаты бар.**  
**Олар:**



**АКТ қолданудың негізгі мақсаты** - білім берудің тиімділігін арттыру болғандықтан, қазіргі таңда ерекше білім беру қажеттіліктері бар балаларды оқытуда тиімділігі байқалады. Себебі, иллюстрацияланған аудио-визуалданған ақпаратты беру арқылы білім берумен қатар түсінуі, меңгеруі жеңілденеді.

Педагог сабақ жоспарының негізінде білім алушыларға өздігінен оқу, оқулықтан оқу материалы мен тапсырмаларды орындау үшін цифрлық технологияларды, веб-сервистерді қолданады.

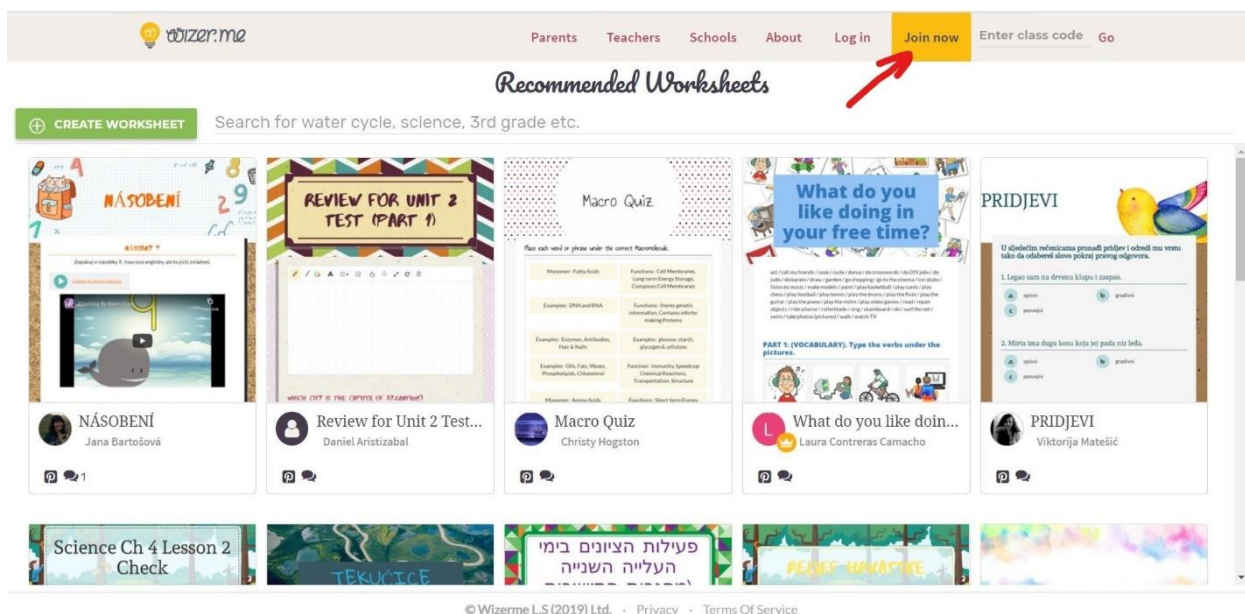
Web-сервис (қызмет) – сайт арасындағы өзара іс- қимылды ұйымдастыратын бағдарлама. Ақпарат бір порталдан екіншісіне беріледі. Веб-сервистер Интернетте көрсетілетін қызметтер деп аталады. Мұндай веб- сервистерді компьютерге, браузерге немесе Интернетке кіру орнына қарамастан пайдалануға болады.

Сайттар әртүрлі жүйелермен басқарылады. Әр түрлі тарату және шифрлау хаттамалары қолданылады. Веб-сервистер әр түрлі алаңдар арасында ақпарат алмасуды жеңілдетеді.

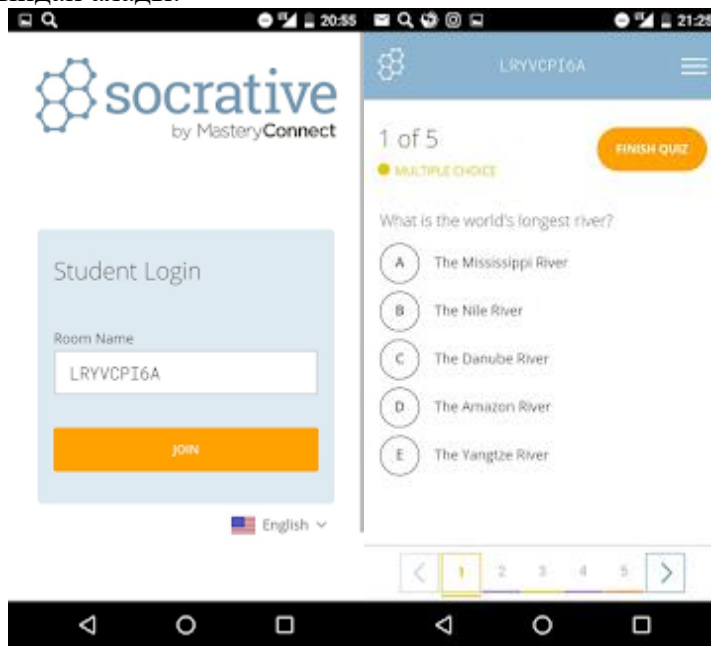
Веб-сервистердің артықшылығы платформаға қарамастан бағдарламалық компоненттердің өзара іс-қимылы үшін қажетті жағдайлар жасау.

Қазақстанда білім беру саласында жиі қолданылатын веб-сервистер – BilimLand, Kundelik.kz, Daryn Online. Цифрлық білім беру ресурстары: Mektap.OnLine; Opiq.kz; Microsoft TEAMS; GoogleClassRoom. Аталған веб-сервистердің қатарын келесілермен толықтырсақ:

- Wizer.me веб-сервисі. Бейнеге негізделген тапсырмалар мен жаттығуларды, қызықты қарапайым дидактикалық материалдарды жасауға арналған жылдам құрал. Пайдаланушылар веб-сервиске тіркеліп, ұсынылған тапсырмаларды орындайды. Мұғалім оқушының жауаптарын жеке кабинетінде көре алады.

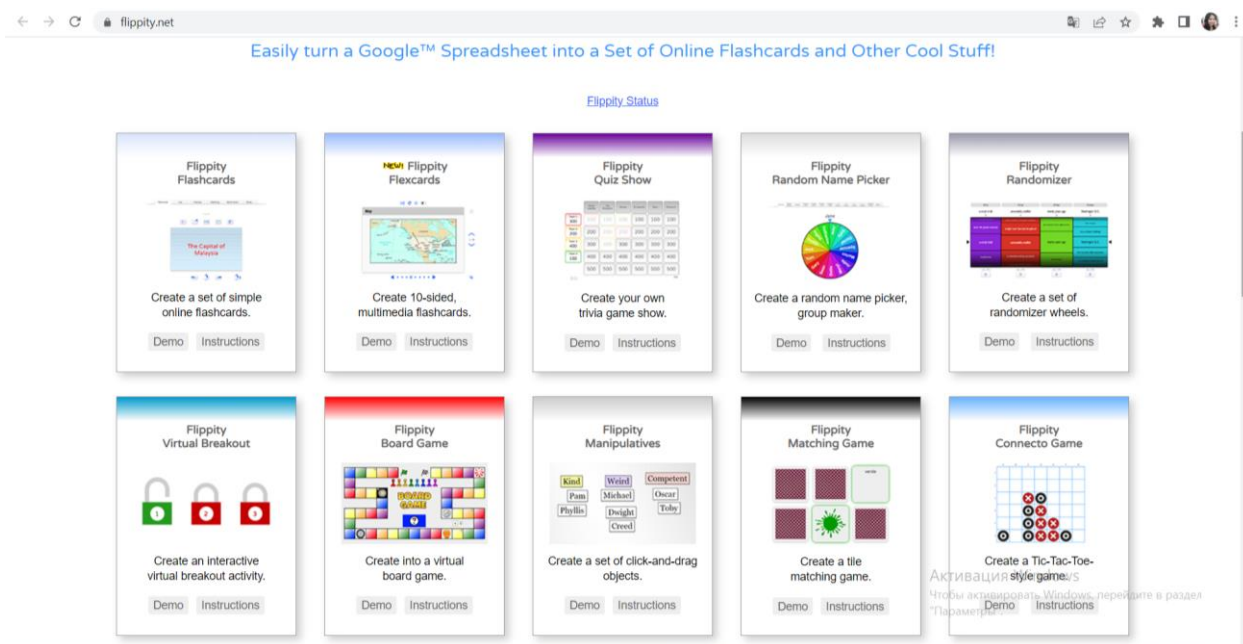


- **Socrative** - бұл тесттерді құруға, іздеуге және таратуға, сондай-ақ тестілеуден өткізуге арналған құрал. Сыртқы көрінісі мен пайдаланудың оңайлығымен ерекшеленеді. Веб-сервисте мұғалім ғана тіркеуден өтіп, жеке тапсырмалары жинақталған қосымша құрып, оған бірегей код алады. Сол арқылы оқушылар қосымшаны жүктеп тапсырмаларды орындай алады.



- **Flippity.net** – бұл мұғалімдерге арналған тегін ресурс, ол викториналар, флэш-карталар, презентациялар, есте сақтау ойындары, сөздерді іздеу және т.б. жасауға мүмкіндік береді. Оны мұғалім презентация құралы және жұмыс тапсырмасы ретінде пайдалана алады, сонымен қатар бұл студенттерді өз жобаларын жасауға тартудың тамаша тәсілі.





## Қорытынды

Қазіргі кезде білім беру саласында оқытуда ақпараттық технологияларсыз елестету мүмкін емес. Дамыған елдердегі білім беру жүйесінде ерекше маңызды болып табылатын мәселердің бірі оқытуды ақпараттандыру, яғни оқу үдерісінде ақпараттық – технологияларды пайдалану болып табылады. Ақпараттық-коммуникациялық технологияны келешек ұрпақ жан – жақты білім алуына, іскер әрі талантты, шығармашылығы мол, дамуына жол ашатын педагогикалық үдеріс.

## Пайдаланылған әдебиеттер:

1. Қазіргі сандық әлемдегі информатика / Информатика в современном цифровом мире : оқу құралы / Т.Б. Нурпеисова, И.Н. Кайдаш. Алматы: Бастау, 2021. 412-413 б.
2. Білім беру жүйесіндегі инновациялық технологиялар. Оқу құралы / Н.Ж. Жанатбекова, Ф. Борибекова. 2021. - 248 б.
3. Информатиканы оқытуда электрондық ресурстарды пайдалану. Үсенов С. 2015. - 152 б.
4. Білімді ақпараттандыру және оқыту мәселелері / Е.Ы.Бидайбеков, В.В.Гриншун, Г.Б.Камалова және т.б. 2014. - 352 б.
5. Педагогиканы цифрлік дәуірде қайта зерделеу. XXI ғасырдағы оқыту дизайны /Rethinking pedagogy for a didital age. Designing for 21ST century learning. С.К.Омарова. /2019. - 328 б.
6. Жоғары мектептегі цифрлық дидактика /Цифровая дидактика в высшей школе. Ж.К. Сағалиевой. - Алматы : Бастау, 2021. - 464 с.
7. <https://www.socrative.com/>
8. <https://app.wizer.me/>
9. <https://www.flippity.net/>

## НАХОЖДЕНИЕ ПОХОЖИХ ОБЪЕКТОВ С ПОМОЩЬЮ MAPREDUCE

Касекеева Айслу Бисеновна<sup>1</sup>, Марат Марта Мараткызы<sup>2</sup>,  
Кашкимбаева Нуржамал Мейркуловна<sup>3</sup>

<sup>1,2</sup> ЕНУ им Л.Н.Гумилева, кафедра Информационные системы, г.Нур-султан

<sup>3</sup> КазАТУ имени С.Сейфуллина, кафедра Компьютерные науки, г.Нур-султан  
[aibike\\_7474@mail.ru](mailto:aibike_7474@mail.ru), [maratkyzy.m@gmail.com](mailto:maratkyzy.m@gmail.com), [knm25-1979@mail.ru](mailto:knm25-1979@mail.ru)

**Аннотация.** В современном мире компании собирают огромное количество информации. Источниками данных могут быть непрерывно поступающие данные с измерительных приборов, потоки сообщений из социальных сетей, метеоданные, потоки данных о местонахождении абонентов сотовой сети, устройства аудио- и видеозаписи, которые нуждаются в дальнейшей обработке. Каждый день генерируется 2,5 квинтиллиона байт данных. Этот огромный набор данных известен как большие данные. Данные слишком разнообразны, быстры, изменчивы и массивны. Существующей вычислительной инфраструктуре сложно обрабатывать большие данные. Чтобы преодолеть этот недостаток, Google представил фреймворк MapReduce. Основной целью работы было изучение возможностей использования, а также преимуществ и недостатков MapReduce для массово-параллельной реализации задачи корреляционного анализа больших объемов данных. Было реализовано решение аналитической задачи о нахождении похожих объектов на базе веб-сервиса Amazon EMR.

### Введение

Традиционные реляционные СУБД плохо справляются с большими объемами данных, которые к тому же часто не структурированы. Для обработки таких данных потребуются кластеры с тысячами узлов. В тоже время необходимо совершать расчеты быстро и масштабируемо. Лаборатория Google предложила концепцию параллельных вычислений MapReduce [1], которая успешно используется для поисковой системы Google. MapReduce – это технология для обработки и генерации большого количества данных, которая применима для многих задач, возникающих в реальной жизни. Главные достоинства MapReduce – параллелизм, динамическое распределение нагрузки по узлам кластера и обеспечение отказоустойчивости. Еще одним преимуществом MapReduce является масштабируемость до десятков тысяч узлов.

### MapReduce

Технология MapReduce для обработки данных, распределенных между узлами кластера, использует две основные операции Map и Reduce. В каждом узле вызывается функция Map, которая перебирает входные данные и создает промежуточный набор пар ключ/значение. Затем данные распределяются между узлами с помощью функции Partition для дальнейшей обработки. Распределение происходит по принципу: одинаковые промежуточные ключи — один и тот же узел. Основными преимуществами MapReduce являются параллелизм, динамическое распределение нагрузки по узлам кластера и отказоустойчивость. На рис. 1 показаны основные этапы программ MapReduce.

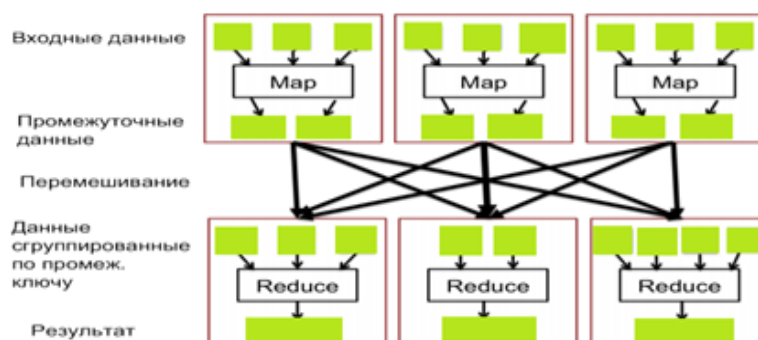


Рисунок 1. Архитектура MapReduce

## Реализация MapReduce

Реализации MapReduce от Google и Hadoop разворачиваются в среде из большого числа недорогих серверов [2]. Обычно это двухпроцессорные x86-машины с 4-8 Гб оперативной памяти с операционной системой Linux. В качестве соединения машин в кластере используется коммутируемый Gigabit Ethernet. Кластеры состоят из тысяч машин, поэтому существует вероятность отказа отдельных узлов. К каждой машине подключены недорогие IDE-диски для хранения данных. Для управления данными используется распределенная файловая система. Планировщик отображает набор задач, полученных от пользователя на набор узлов кластера. Hadoop состоит из двух уровней: уровень хранения данных Hadoop DFS (HDFS) и уровень обработки данных Hadoop MapReduce. HDFS распределенная файловая система с блочной структурой, управляемая узлом-мастером (master) [3].

## Исследование и построение решения задачи

Одной из важных задач статистики является нахождение взаимосвязей наблюдаемых процессов и явлений. Обычно связь бывает двух видов: функциональной и корреляционной (неполной). Например, функциональная связь достаточно часто проявляется в физике, химии, экономике. В этом случае величине факторного признака строго соответствует несколько значений функции. Корреляционная же связь не является строгой, она проявляется в среднем, для массовых наблюдений, когда заданным значениям зависимой переменной соответствует некоторый ряд вероятных значений независимой переменной [4]. Для определения степени связи и выявления неизвестных причинных связей выполняется корреляционный анализ. При этом для количественной оценки тесноты связи вычисляется коэффициент корреляции. Имеется несколько способов вычисления коэффициента корреляции. Одним из них является коэффициент линейной корреляции.

Коэффициент линейной корреляции (то есть коэффициент Пирсона) позволяет определить существование линейной зависимости между двумя величинами X и Y. Если связь между величинами имеет линейный характер, коэффициент Пирсона точно устанавливает тесноту этой связи. Общий вид формулы:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \times (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4),$$

где  $x_i$  – значения переменной X,

$y_i$  – значения переменной Y,

$\bar{x}$ ,  $\bar{y}$  – выборочные средние,

n – количество наблюдений.

Имеются следующие допущения:

– Исследуемые величины должны иметь нормальное распределение или быть близки к нормальному распределению.

– Одинаковое количество наблюдений исследуемых величин.

Часто в вычислениях используется аналог формулы (4), полученный с помощью преобразований:

$$r_{xy} = \frac{n \times \sum_{i=1}^n (x_i \times y_i) - (\sum_{i=1}^n x_i \times \sum_{i=1}^n y_i)}{\sqrt{[n \times \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2] \times [n \times \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2]}} \quad (5)$$

Коэффициент корреляции находится в диапазоне от -1 до 1. Если  $|r_{xy}|=1$ , то X и Y линейно зависимы. Если  $|r_{xy}|=0$  линейно независимы. Для оценки тесноты связи в корреляционном анализе применяется шкала английского статистика Чеддока, однако такая оценка является нестрогой.

Для оценки значимости полученного парного коэффициента корреляции можно применить z-преобразование Фишера [5]. Преобразование определяется как:

$$z(r_{xy}) = \frac{1}{2} \ln \left( \frac{1 + r_{xy}}{1 - r_{xy}} \right) \quad (6).$$

Так как величина  $z$  является нормально распределённой величиной, то проверка основной гипотезы о незначимости парного коэффициента корреляции сводится к проверке основной гипотезы о незначимости величины  $z$ :  $H_0: z=0$ . Конкурирующая гипотеза:  $H_1: z \neq 0$ .

Величина стандартной ошибки:

$$\omega_z = \frac{1}{\sqrt{n-3}} \quad (7)$$

Преобразование Фишера и обратное преобразование

$$\varphi(z) = \frac{e^{2z} - 1}{e^{2z} + 1} \quad (8)$$

может использоваться для построения доверительного интервала  $r_{xy}$ . Доверительный интервал строится следующим образом:

- Ошибка коэффициента корреляции (7) умножается на параметр  $t$ , который при доверительной вероятности  $P=95\%$  равен 1,96[5].
- Интервал строится как  $\varphi(z(r_{xy}) \pm t \times \omega_z)$ .

Если доверительный интервал содержит нуль, то с высокой вероятностью в генеральной совокупности может оказаться нулевая корреляция, то есть гипотеза  $H_0$  не опровергается. Если же нуль не попадает в интервал, то коэффициент корреляции является статистически значимым и по нему можно судить обо всей генеральной совокупности.

**Описание аналитической задачи.** На основе корреляционного анализа можно решить задачу нахождения похожих объектов. Одной из таких задач является нахождение похожих исполнителей музыки. Для измерения попарного сходства элементов в наборе данных используем коэффициент корреляции Пирсона рейтингов элементов, которые были оценены подмножеством пользователей. Этот подход используется также в рекомендательных системах.

**Построение решения задачи.** Для данной задачи используется набор данных Audioscrobbler. Данные хранятся в формате .txt. Audioscrobbler содержит два файла «Музыкальные предпочтения» и «Исполнители». Файл «Музыкальные предпочтения» содержит информацию о номере слушателя, о номере исполнителя, и том сколько раз каждый пользователь прослушал данного исполнителя. «Исполнители» состоит из номера исполнителя и наименования исполнителя. Предполагается, что чем больше пользователь слушал исполнителя музыки, тем выше оценил его. Данная задача также может быть использована и на других наборах данных, например для анализа различного рода услуг, товаров, фильмов.

Разделим задачу на четыре подзадачи для последовательной обработки с использованием технологии MapReduce.

Подзадача 1. Подсчет количества слушателей для каждого исполнителя. Для каждого исполнителя подсчитывается количество его слушателей. Под рейтингом далее будет пониматься подсчет количества слушателей для каждого исполнителя. Под значением рейтинга далее будет пониматься количество прослушиваний данного исполнителя данным пользователем.

Подзадача 2. Подсчет предпочтений пользователей. Для каждого пользователя формируется список, состоящий из количества прослушанных исполнителей, общего количества его прослушанных песен и последовательности исполнителей с количеством прослушиваний данным пользователем.

Подзадача 3. Вычисление коэффициента корреляции Пирсона. Из результатов, полученных на шаге 1 и 2, производится случайная выборка. Для каждой пары исполнителей вычисляется коэффициент корреляции Пирсона.

Подзадача 4. Формирование окончательного результата.

Описание практической части

Для проведения вычислений использовался веб-сервис Amazon Elastic MapReduce (Amazon EMR) [6], который позволяет аналитикам, разработчикам, исследователям и предприятиям легко и экономически эффективно обрабатывать огромные объемы данных. Amazon EMR основан на реализации Hadoop MapReduce, который автоматически разворачивается в облачном сервисе Amazon Elastic Compute Cloud (Amazon EC2) [7], и на использовании хранилища данных Amazon S3[8] для чтения данных и сохранения результата. Amazon EMR удобен тем, что можно разрабатывать приложения на любом знакомом языке: Java, Ruby, Perl, Python, PHP, R или C++.

В качестве языка программирования использовался Python, который хорошо работает с текстовыми данными. На этом этапе отбираются из результатов подзадачи 3 пары, у которых больше 20 слушателей. Для каждого исполнителя выбираются k наиболее «родственных» исполнителей (на основании значения коэффициента корреляции), также дополнительно из файла «Исполнители» выводятся наименования исполнителей.

**Общая схема работы.** Задача нахождения похожих исполнителей состоит из четырех MapReduce задач. Они запускались последовательно, промежуточные результаты загружались в Amazon S3 либо в HDFS, так что трафик передачи по сети использовался

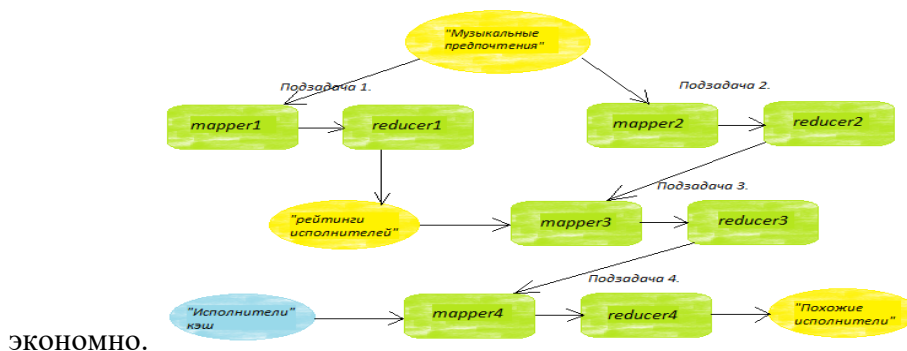


Рисунок 2. Обзор схемы выполнения задачи

Подсчет количества слушателей для каждого исполнителя. Входные данные содержат 147160 исполнителей и максимальное количество исполнителей, прослушиваемых одним пользователем 3331.

На этапе map для каждой строки исходного набора данных «Музыкальные предпочтения» формируются пары <ключ, значение>: <<номер пользователя, номер исполнителя>, ln значение рейтинга>. Данные берутся в логарифмическом масштабе.

Разделение данных partitioner происходит с помощью библиотеки org.apache.hadoop.mapred.lib.KeyFieldBasedPartitioner по номеру пользователя. То есть данные группируются по номеру пользователя, а затем сортируются по всему ключу.

Результатом reducer является множество пар <номер пользователя, количество прослушанных исполнителей – общее количество прослушиваний – <последовательность номер исполнителя, значение рейтинга >>. Результат также отсортирован.

Вычисление коэффициента корреляции Пирсона

На этапе map из результатов полученных на шаге 1 и шаге 2 производится случайная выборка для построения промежуточного результата: <номер исполнителя1, номер исполнителя2; значение рейтинга исполнителя1 у данного пользователя, значение рейтинга исполнителя2 у данного пользователя>

На этапе reduce подсчитываем количество пользователей n для каждой пары исполнителей, которые слушали обоих исполнителей, и высчитываем коэффициент Пирсона, с учетом преобразований Фишера для этих пар. То есть если X-это исполнитель 1, а Y- это исполнитель 2, то значениями величины X и Y, соответственно  $x_i$  и  $y_i$ , будут значения рейтингов исполнителей для каждой пары <X,Y>. На выходе будет набор данных <номер исполнителя1, коэффициент корреляции, номер исполнителя2, количество

слушатели обоих исполнителей>, <номер исполнителя2, коэффициент корреляции, номер исполнителя1, количество слушатели обоих исполнителей> и так далее.

**Формирование конечного результата.** На этапе map отбираются из результатов подзадачи 3 пары, у которых больше 20 слушателей. На этапе reduce для каждого исполнителя выбираются k наиболее «родственных» исполнителей (на основании значения коэффициента корреляции), а также дополнительно из файла «Исполнители» выводятся данные исполнителей. В результате мы видим парную корреляцию между исполнителями, то есть похожих исполнителей. На рисунке 11 образец из семи исполнителей. Соответственно, первые две колонки номера исполнителей, третья коэффициент корреляции Пирсона, четвертая общее количество прослушиваний и пятая колонка это наименование исполнителя.

1000143	1000143	1	0	Toby Keith
1000143	1000418	0.1574	120	Mark Chesnutt
1000143	1003557	0.2967	809	Garth Brooks
1000143	1012511	0.1004	237	Kenny Rogers
1000418	1000143	0.1574	120	Toby Keith
1000418	1000418	1	0	Mark Chesnutt
1000418	1003557	0.0835	114	Garth Brooks
1000418	1012511	0.0734	50	Kenny Rogers
1001820	1001820	1	0	2Pac
1001820	1003557	0.0764	878	Garth Brooks
1001820	1011819	0.1655	2428	A Tribe Called Quest
1001820	700	0.1205	2250	De La Soul
1003557	1000143	0.2967	809	Toby Keith
1003557	1000418	0.0835	114	Mark Chesnutt
1003557	1003557	1	0	Garth Brooks
1003557	1012511	0.155360	Kenny Rogers	
1011819	1001820	0.1655	2428	2Pac
1011819	1003557	0.0387	288	Garth Brooks
1011819	1011819	1	0	A Tribe Called Quest
1011819	700	0.3527	3050	De La Soul
1012511	1000143	0.1004	237	Toby Keith
1012511	1000418	0.0734	50	Mark Chesnutt
1012511	1003557	0.155360	Garth Brooks	
1012511	1012511	1	0	Kenny Rogers
700	1000143	0.0713	114	Toby Keith
700	1001820	0.1205	2250	2Pac
700	1011819	0.3527	3050	A Tribe Called Quest
700	700	1	0	De La Soul

Рисунок 3. Обзор схемы выполнения задачи

### Заключение

MapReduce является несомненно востребованной технологией на быстроразвивающемся рынке. Применение MapReduce позволяет аналитикам быстро распараллелить аналитическую задачу и реализовать сложную аналитику. Поддержка технологии MapReduce означает горизонтальную масштабируемость аналитических систем.

Недостатком MapReduce является неэффективность при маленьких кластерах с небольшими данными.

В рамках данной работы было выполнено:

- Исследование существующих подходов к аналитической обработке данных.
- Анализ общих возможностей MapReduce
- Было предложено решение для задачи о нахождении похожих объектов, в основе которой лежит корреляционный анализ данных, с использованием технологии MapReduce.

### Список литературы

- [1] Dean J., Ghemawat S. MapReduce: Simplified Data Processing on Large Clusters // In Proceedings of the USENIX Symposium on Operating Systems Design & Implementation (OSDI). San Francisco, USA, 2004. P. 137-147.
- [2] Chu C.-T., Kim S. K., Lin Y.-A., Yu Y.Y., Bradski G. R., Ng A. Y., Olukotun K. Map - Reduce for machine learning on multicore.// Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems. Vancouver, Canada: MIT Press 2007. P. 281-288.
- [3] Кузнецов С.Д. MapReduce: внутри, снаружи или сбоку от параллельных СУБД? [HTML] ([http://citforum.ru/database/articles/dw\\_appliance\\_and\\_mr/](http://citforum.ru/database/articles/dw_appliance_and_mr/)).
- [4] Miner D., Shook A. MapReduce Design Patterns. Sebastopol: O'Reilly Media, 2012. 227 p.
- [5] Харченко Л.П., Долженкова В.Г., Ионин В.Г. Статистика. М.: ИНФРА-М, 2001. 384 с.

- [6] G.Loftus. Confidence intervals around Pearson  $r$ 's [PDF] ([http://faculty.washington.edu/gloftus/P317-318/Useful\\_Information/r\\_to\\_z/PearsonrCIs.pdf](http://faculty.washington.edu/gloftus/P317-318/Useful_Information/r_to_z/PearsonrCIs.pdf)).
- [7] Amazon EMR.[HTML] (<http://aws.amazon.com/elasticmapreduce/>).
- [8] Amazon EC2. [HTML] (<http://aws.amazon.com/ec2/>)

# СТРУКТУРА БАЗ ЗНАНИЙ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ В ОБЛАСТИ ВЕТЕРИНАРИИ

*Шопагулов Олжас Алматович*

*КАТУ им. С. Сейфуллина, [shopagulov@gmail.com](mailto:shopagulov@gmail.com), г. Астана*

**Абстракт.** Формирование компьютерных технологий, интернета и мобильных приложений раскрыли абсолютно новые возможности в использовании компьютеров и современных гаджетов в различных отраслях жизнедеятельности человека. В эту область входит и ветеринарная отрасль. На сегодняшний день учеными из ближнего и дальнего зарубежья разрабатываются и вводятся всевозможные программы и приложения, ориентированные на профессию ветеринарного врача. В данной статье представлена информация по разработанной интеллектуальной системе в области ветеринарии. Данное программное обеспечение создано для диагностики заболеваний крупного рогатого скота и, в результате, представлению методов лечения выявленного заболевания. Несомненно, фактором, приводящим к гибели скота, является распространение болезней, которые могут ухудшить здоровье поголовья. Таким образом, экспертная система для диагностики заболеваний коров необходима для упрощения работы ветеринаров и в последующем уменьшении уровня заболеваемости коров.

**Ключевые слова:** интеллектуальная система, ветеринария, базы знаний, данные, крупный рогатый скот.

**Введение.** В настоящее время в Казахстане центром технологической компетенции в области цифровизации АПК Министерством сельского хозяйства РК определен «Казахский Агротехнический Университет им. С. Сейфуллина» [1]. Формирование компьютерных технологий, интернета и мобильных приложений раскрыли абсолютно новые возможности в использовании компьютеров и современных гаджетов в различных отраслях жизнедеятельности человека. В эту область входит и ветеринарная отрасль. На сегодняшний день учеными из ближнего и дальнего зарубежья разрабатываются и вводятся всевозможные программы и приложения, ориентированные на профессию ветеринарного врача [2]. Для ветеринара, работающего с сельскохозяйственными животными, необходим многофункциональный инструмент, который посредством имеющихся у системы функций улучшит эффективность труда специалиста и повысит его производительность, посредством правильно принятых решений при диагностике, лечении, профилактике и ликвидации той или иной болезни сельскохозяйственных животных [3].

Для ветеринара, работающего с сельскохозяйственными животными, необходим многофункциональный инструмент, который посредством имеющихся у системы функций улучшит эффективность труда специалиста и повысит его производительность, посредством правильно принятых решений при диагностике, лечении, профилактике и ликвидации той или иной болезни сельскохозяйственных животных [5].

**Литературный обзор.** В числе зарубежных ученых, занимающихся разработкой экспертных систем для решения задач ветеринарии в сфере сельского хозяйства лежат разработки Chutchada Nusai [4], однако в числе Казахстанских ученых, занимавшихся данной проблематикой подобных работ нет. В области медицины, к примеру, в числе зарубежных разработок присутствует «Экспертная система для регуляции процесса питания пожилых людей», страна производитель - Румыния. Система основана на определенных правилах для употребления пищевых продуктов и блюд с использованием генетических алгоритмов для расчета пригодности возможных решений с использованием персонализированных целевых значений различных питательных веществ. База знаний, состоит из заранее определенных продуктов, сгруппированных в соответствии с пропорциями углеводов, белков и жиров. Данная система использует процесс



рассуждений, основанный на правилах SWRL (язык правил семантической группы) для выработки рекомендаций с помощью мер семантического сходства [5].

**Материалы и методы исследования.** При разработке ЭС использовались продукционные модели представления знаний, обладающие различными свойствами, необходимыми для представления знаний. [6,7,8].

Математический аппарат и аппарат теории нечеткого композиционного вывода использовались при построении теории нечетких множеств и нечеткой логики [9].

Статистические же методы использовались как системы обработки статистической экспертной информации в общей теории автоматизированных систем.

Системный, аналитический и логический подходы использовались при обработке экспертных слабоструктурированных данных.

Метод сравнения получил применение при сопоставлении имеющихся интеллектуальных систем в области ветеринарии при обработке полученных экспериментальных данных.

**Обсуждение.** Источник знаний в определенной степени определяет точность, надежность, практичность и эффективность приобретения знаний. Принимая во внимание огромный объем знаний, необходимых для создания экспертной системы диагностики болезней животных, выбраны 4 основные области: книги, эксперты в предметной области, другие материалы, литература и информация из интернета.

Книги в основном включают монографии по болезням животных: такие как «Болезни крупного рогатого скота», «Болезни молочных коров» и т. д. [10].

Эксперты в данной области – это специалисты-ветеринары Казахского Агротехнического Университета им. С. Сейфуллина в различных дисциплинах, таких как внутренняя медицина, хирургия, акушерство, инфекционные заболевания, паразитарные заболевания, нарушения обмена веществ и отравления. и т. д.

Другие материалы и документы в основном включают журналы и технические отчеты. Интернет - это в основном публичные и частные сайты некоторых подразделений ветеринарии и информационных систем.

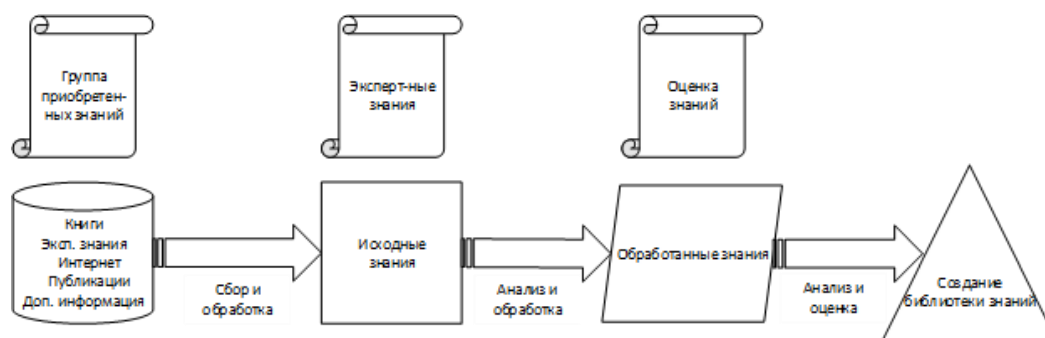


Рисунок 1 - Процесс получения, анализа и оценки знаний

Задачи экспертной системы диагностики болезней КРС ориентированы на проблемы, связанные с диагностикой болезней - это проблема, которую необходимо решить. Взяв за основу болезни, проведено исследование на основе экспертных знаний и обнаружено, что каждая болезнь может иметь один или несколько фенотипов, и каждый фенотип состоит из общей информации и симптомов, а общая информация и симптомы имеют свои собственные характеристики и значения. Исходя из этого, существует возможность построить модель отношений для болезней и симптомов. Все болезни могут использоваться совместно с моделью представления элементов знаний в качестве стандарта для классификации и систематизации, структура показана на рисунке 2.

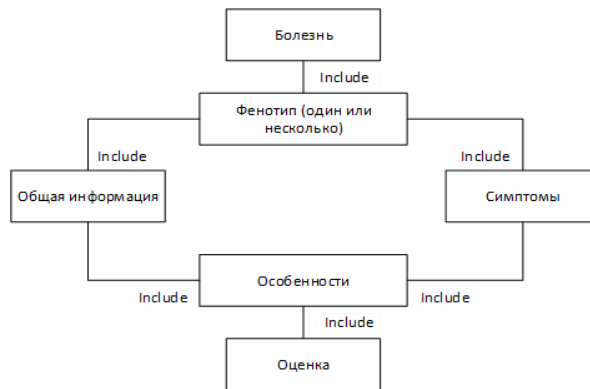


Рисунок 2 - Модель представления элементов знаний

Алгоритм принятия решения врачей-ветеринаров имеет особое значение в разработке системы, в соответствии с этим создан алгоритм мышления ветеринаров (Рисунок 3).



Рисунок 3 – Алгоритм мышления ветеринаров

Чтобы реализовать целенаправленный и интеллектуальный опрос системы, уменьшить количество опросов, улучшить работу системы и сделать процесс диагностики заболеваний в большей степени в соответствии с диагностическим мышлением эксперта, диагностические знания были классифицированы. Классификация знаний о диагностике болезней животных может быть выражена в виде структуры, представленной на рисунке 4.

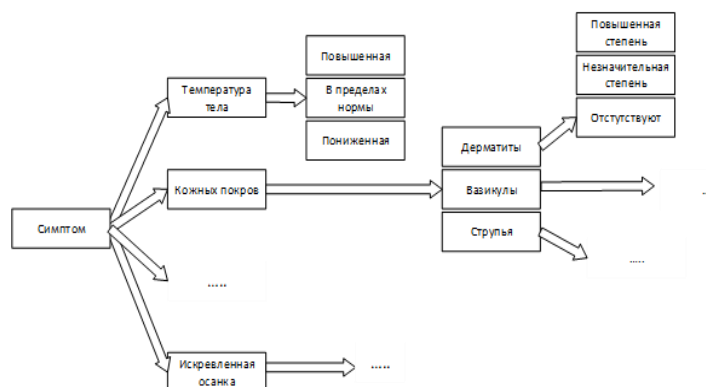


Рисунок 4 - Пример группировки уровней знаний при диагностике болезней животных

Базовая структура продукционных правил включает две части: преамбула и заключение: преамбула (или часть «ЕСЛИ») описывает состояние, а заключение (или часть «ТО») описывает некоторые действия, которые происходят при условии осуществления условий преамбулы. В знаковом виде продукционные правила можно представить, как:

$$\text{ЕСЛИ } \{A\}, \text{ ТО } \{B\}, \quad (1)$$

где  $\{A\}$  обозначает набор из нескольких условных факторов,  $\{B\}$  обозначает фактор вывода. В наборе условных факторов может использоваться отношение «и» либо отношение «или». Примерами могут быть факторы:

$$\text{ЕСЛИ } \{A_1\} \wedge \{A_2\} \wedge \{A_3\} \wedge \dots \wedge \{A_n\}, \text{ ТО } \{B\}, \quad (2)$$

указывает, что если все  $\{A_1\} \dots \{A_n\}$  истинны, тогда  $\{B\}$  истинно.

Продукционные правила могут быть способом представлением знаний, которые можно разложить на причинно-следственные связи [56]. Диагностика болезней животных - это процесс решения проблем, главный принцип которых заключается в использовании наблюдаемых симптомов больных животных для определения заболевания. Связь между симптомами и заболеванием может быть выражена как причинно-следственная связь, данный метод хорошо отражает знания о диагностике болезней животных.

$$\text{ЕСЛИ (Симптомы), ТО (Болезнь)} \quad (3)$$

Объем знаний о диагностике болезней животных довольно велик. Некоторые заболевания могут иметь несколько десятков симптомов, а некоторые заболевания зачастую имеют общие симптомы.

**Заключение.** В соответствии с характеристиками и потребностями процесса диагностики болезней животных принимается стратегия реализации, состоящая из комбинирования различных функциональных блоков системы. Основная функция экспертной системы диагностики болезней животных состоит в том, чтобы диагностировать болезни и предлагать меры профилактики и контроля, поэтому построение системы систематизированных баз данных ведется вокруг этих двух аспектов. В экспертной системе диагностики болезней КРС хранение и обработка всей информации и знаний осуществляется с помощью базы данных правил, базы данных фактов, базы данных интерпретации, базы данных профилактических и контрольных мер, динамической базы данных и мультимедийной базы данных. Создание базы знаний описанным выше способом может решить следующие проблемы в процессе ввода симптомов во время диагностики заболевания: 1) Вся информация, выбранная пользователем, предоставляется системой, и нет проблемы, что вводимая информация не совпадает, 2) в базе правил принято, что каждая часть входной информации соответствует по крайней мере одному правилу, то есть должно быть одно или несколько заболеваний с информацией, введенной пользователем в качестве условия диагностики, 3) когда выбраны несколько болезней, их активные состояния различаются.

#### Список литературы.

1. Постановление Правительства Республики Казахстан О присвоении статуса исследовательского университета некоммерческому акционерному обществу "Казахский агротехнический университет имени С. Сейфуллина": утв. 1 сентября 2020 года, № 545.

2. Ruben PurroyVásquez, Alberto AlfonsoAguilar-Lasserre. Expert system based on a fuzzy logic model for the analysis of the sustainable livestock production dynamic system//Computers and Electronics in Agriculture. - 2019.- Vol.161.-P.104-120

3. Алимжанова Л. Отчет о научно-исследовательской работе "Тренсферт и адаптация инновационных технологий для оптимизации производственных процессов на молочных фермах Северного Казахстана": дис. д. с.-х. н. профессор тех. наук: О.0907. - Нур-Султан:НАО КазАТУ им. С. Сейфуллина, 2020.

4. Beverly G. Hope, Rosewary H. Wild. « An Expert Support System for Service Quality Improvement»//Proceedings of the Twenty-Seventh Annual Hawaii International Conference on System Science. – Hawaii, 1994.
5. Tudor Cioara, Ionut Anghel, Ioan Salomie, Lina Barakat, Simon Miles, Dianne Reidlinger, Adel Taweel, Ciprian Dobre, Florin Pop Expert system for nutrition care process of older adults // Future Generation Computer Systems. -2018. - №80. - P. 368–383.
6. Sachin Raju. Knowledge Representation Models in Artificial Intelligence.  
<https://www.fingent.com/blog/classifying-knowledge-representation-in-artificial-intelligence/>  
07.05.2019.
7. Morrow. C. T. R.E. Graves, and P. H. Heinemann. Expert systems implementation using “Rulemaster” at Penn State//Am. SOC. Agric.-1987.- P. 87.
8. Byrd, TA., K.L. Cossick and R.W. Zmud. A synthesis of Research on Requirements Analysis and Knowledge Acquisition Techniques//MIS Quarterly. - 1992.-P. 117-138.
9. Николай Паклин. Нечеткая логика — математические основы.  
<https://loginom.ru/blog/fuzzy-logic> 03.06.2020.
10. Лимаренко А.А., Бараников А.И., Лимаренко Ан.А. Болезни крупного рогатого скота. – СПб.: Издательство «Лань», 2010.

# ПРИМЕНЕНИЕ ORM-ТЕХНОЛОГИИ ПРИ РАЗРАБОТКЕ ОБРАЗОВАТЕЛЬНОГО ОНЛАЙН ПОРТАЛА С ЭЛЕМЕНТАМИ СОЦИАЛЬНОЙ СЕТИ

Л.А. Танченко<sup>1</sup>, Д.В. Одноуров<sup>1</sup>

*Евразийский национальный университет им. Л. Н. Гумилева, г. Астана, Казахстан  
ltanchenko@mail.ru*

При разработке большинства информационных систем, WEB-сервисов, сайтов и других информационных продуктов, а также непосредственно при разработке образовательного онлайн портала активно используются технология базы данных. База данных (БД) — организованная в соответствии с определёнными правилами и поддерживаемая в памяти компьютера совокупность данных, характеризующая актуальное состояние некоторой предметной области и используемая для удовлетворения информационных потребностей пользователей [1].

Однако для подключения к базе данных и обмена данными с ней необходим специальный посредник. В этой статье рассматривается применение в качестве посредника для подключения к базе данных при разработке онлайн-портала – ORM-технология Entity Framework. Перед тем, как перейти к понятию ORM-технологий и их актуальности в разработке, сначала рассмотрим обычный способ подключения к базе данных без использования ORM-технологии, на примере технологии ADO.NET.

ADO.NET представляет собой технологию работы с данными, разработанную компанией Microsoft, основанную на платформе .NET Framework. Эта технология обеспечивает простой интерфейс подключения к СУБД [2].

Таким образом, для подключения приложения к базе данных применяются объекты из ветки Connection, которые в свою очередь после завершения транзакции передают полученные данные в объект DataSet. Данные из базы данных хранятся непосредственно в объекте DataTable, который представляет собой двумерную матрицу, в которой хранятся полученные данные в каждой строке. При таком подходе возникают следующие минусы:

1. Для обработки данных после получения их из БД – необходимо знать структуру базы данных, так как в полученных матрицах не отмечаются ни названия полей, ни типы хранимых данных.

2. Ограничен функционал при работе с полученными данными. После получения данных из БД они хранятся в виде матрицы в специальном объекте, что в свою очередь ограничивает набор операций, которые можно совершить с ними. Такие операции как сортировка, фильтрация, пагинация и другие выполнить крайне сложно либо не возможно, поэтому все эти настройки необходимо включать в изначальный SQL запрос.

3. Для работы с БД необходимо знание SQL, а также структуры БД, что в свою очередь затрудняет и замедляет разработку приложения.

4. Минимальная защита от SQL-инъекций (SQL-injection).

Для решения этих проблем были разработаны ORM-технологии. ORM (англ. Object-Relational Mapping) – это технологии, главная задача которых – связать таблицы базы данных с методологией объектно-ориентированного программирования, таким образом решая перечисленные выше проблемы.

Entity Framework – ORM-технология, разработанная компанией Microsoft для языка программирования C#, а также фреймворков: .Net Framework, ASP.Net Core и других, основанных на языке C#. Эта технология позволяет получать данные из базы данных и связывать их с соответствующими классами. Пример такого отношения показан на рисунке 1 [3].

Одно из важнейших преимуществ использования ORM-технологий, и непосредственно EntityFramework заключается в том, что все полученные данные из таблицы представляются в виде объектов (классов) C#. Таким образом работа с данными

из базы данных сводится к работе с простыми классами. Такие классы называются моделями.

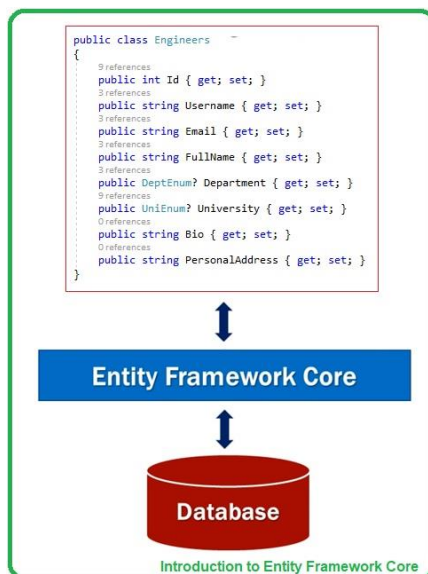


Рисунок 1. Представление класса на языке C# в таблице базы данных

Также ORM-технология Entity Framework позволяет не только работать с уже спроектированной базой данных, но и производить полный процесс моделирования базы данных непосредственно в приложении. Для этого в приложении описывается набор классов, которые будут отображать таблицы в базе данных, и в ряде настроек указываются дополнительные характеристики полей таблиц, такие как индексы, ключевые поля, значения по умолчанию и т.д [4].

Также для доступа к базе данных нет необходимости в написании сложных SQL-запросов, для этого необходимо просто указать какие поля необходимо загрузить через свойства моделей. После выполнения транзакции полученные данные получают обширные возможности для дополнительных манипуляций, такие как фильтрация, сортировка, пагинация запись, изменение, удаление и другие, без необходимости в повторных запросах.

При разработке образовательного онлайн портала с элементами социальной сети ORM-технология EntityFramework активно используется как для проектирования базы данных, так и для последующих операций с ней. Использование этой технологии значительно упрощает и ускоряет разработку приложения, а также повышает степень защищенности от SQL-инъекций [5].

Рассмотрим примеры использования ORM-технологии EntityFramework при разработке образовательного онлайн портала с элементами социальной сети. Во первых, необходимо подключить эту библиотеку при помощи менеджера пакетов NuGet и выбрать необходимую версию библиотеки для выбранной СУБД, в моем случае – MS SqlServer 2012. Версия библиотеки выбирается на основе версии проекта ASP .NET Core. В моем случае – используется ASP .NET Core 6. Установка этой библиотеки показана на рисунке 2 [6].

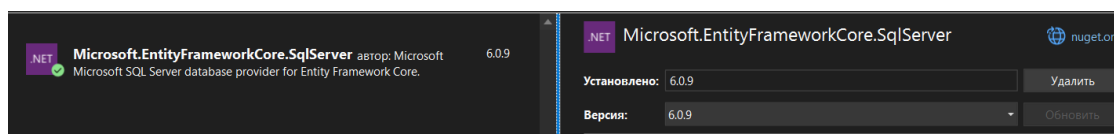


Рисунок 2. Установка библиотеки EntityFramework

После установки необходимо настроить подключение к СУБД. Для этого в конфигурации проекта задается строка подключения. Для СУБД MS SqlServer она имеет вид: Server=[имя\_сервера];Database=[название\_базы\_данных];Trusted\_Connection=True; После этого создается класс – контекст подключения. Именно через этот контекст и ведется работа с базой данных. Контекст имеет следующий код, показанный в листинге 1.

Листинг 1. Код класса ApplicationDbContext

```
public class ApplicationDbContext : DbContext
{
    public DbSet<User> Users { get; set; }
    public ApplicationDbContext(DbContextOptions options) : base(options)
    {
        Database.EnsureCreated();
    }
}
```

Этот класс наследуется от класса DbContext. Поле Users характеризует таблицу в БД, которая строится на основе модели User. Так как разработка образовательного онлайн портала с элементами социальной сети находится только на начальной стадии, рассмотрим применение этой ORM-технологии только с одной таблицей Users. После создания контекста подключения – можно отправлять запросы к БД. И так как все это основано на ORM-технологии EntityFramework – нет необходимости вручную открывать соединение с БД для получения данных – EntityFramework делает все промежуточные этапы автоматически. Рассмотрим простейшие операции CRUD, которые будут применяться в онлайн портале.

Операция Create (создание записи). Для создания записи применяется метод ApplicationDbContext.Model.Add(model). Этот метод генерирует SQL запрос типа CREATE и отправляет его на сервер базы данных. Применение этого метода приведено в листинге 2.

Листинг 2. Создание записи

```
public async Task<IActionResult> Post(User user)
{
    if (user == null)
    {
        return BadRequest();
    }
    if (user.Age > 120)
    {
        ModelState.AddModelError("Age", "age must be <= 120");
    }
    if (!ModelState.IsValid)
    {
        return BadRequest(ModelState);
    }
    await db.Users.AddAsync(user);
    await db.SaveChangesAsync();
    return Ok(User);
}
```

В приведенном выше листинге перед тем, как добавить запись – проверяется ее наличие в БД и корректность введенных данных. Операция Read – выборка одной или нескольких записей из таблицы. Чтение всех записей в таблице реализуется на прямую через свойство ApplicationDbContext.Model. Выборка одной записи реализуется посредством метода ApplicationDbContext.Model.FirstOrDefault(searchDelegate). Здесь searchDelegate – это функция поиска, в результате которой возвращается объект по заданным параметрам поиска. Пример использования этих методов показан в листинге 3.

Листинг 3. Выборка из базы данных.

//получение всех записей

```
public async Task<ActionResult<IEnumerable<User>>> Get()
```

```
{  
    return await db.Users.ToListAsync();  
}
```

```
[HttpGet("{id}")]
```

//получение одной записи

```
public async Task<ActionResult<User>> Get(int id)
```

```
{  
    if (id == null)  
    {  
        return BadRequest();  
    }  
    User user = await db.Users.FirstOrDefaultAsync(x => x.UserId == id);  
    if (user == null)  
    {  
        return NotFound();  
    }  
    return new ObjectResult(user);  
}
```

Операции Update и Delete – обновление записи и удаление записи. Для этих операций предусмотрены методы `ApplicationContext.Model.Update(model)` и `ApplicationContext.Model.Remove(model)` соответственно. Эти методы автоматически ищут в таблице записи, которые совпадают с переданным в параметр метода объектом, так что нет необходимости осуществлять поиск вручную. Однако перед выполнением этих операций желательно убедиться в наличии искомым записей, чтобы избежать непредвиденных ситуаций. Пример использования этих методов показан в листинге 4 и 5 соответственно [7].

Листинг 4. Операция обновления данных.

```
public async Task<IActionResult> Put(User user)
```

```
{  
    if (user == null)  
    {  
        return BadRequest();  
    }  
    if (!db.Users.Any(x => x.UserId == user.UserId))  
    {  
        return NotFound();  
    }  
    db.Update(user);  
    await db.SaveChangesAsync();  
    return Ok(user);  
}
```

Листинг 5. Операция удаления данных.

```
public async Task<IActionResult> Delete(int id)
```

```
{  
    if (id == null)  
    {  
        return BadRequest();  
    }  
    User user = await db.Users.FirstOrDefaultAsync(x => x.UserId == id);
```



```

if (user==null)
{
    return NotFound();
}
db.Users.Remove(user);
await db.SaveChangesAsync();
return Ok();
}

```

Следует также отметить, что после выполнения каких-либо операций с БД – необходимо вручную сохранить все изменения при помощи метода `ApplicationContext.SaveChanges()` либо `ApplicationContext.SaveChangesAsync()` [8].

В будущем технология EntityFramework при разработке образовательного онлайн портала с элементами социальной сети будет применяться в следующих целях:

- Регистрация, авторизация пользователя;
- Создание статей;
- Поиск\фильтрация по статьям, друзьям, группам и др;
- Контроль ролей пользователей.

#### Список литературы

[1] Когаловский М. Р. Энциклопедия технологий баз данных. — М.: Финансы и статистика, 2012. — 800 с

[2] Сериков А. Разработка веб-приложений. Современные веб-технологии [онлайн-ресурс] URL: [https://intuit.ru/studies/professional\\_retraining/942/courses/467/info](https://intuit.ru/studies/professional_retraining/942/courses/467/info)

[3] Столбовский Д. Разработка Web-приложений ASP .NET с использованием Visual Studio .NET [онлайн-ресурс] URL: <https://intuit.ru/studies/courses/1139/250/info>

[4] Работа с Entity Framework 6 [онлайн ресурс] URL: <https://professorweb.ru/my/entity-framework/6/level1/>

[5] Руководство по ASP.NET Core MVC [онлайн ресурс] URL: <https://metanit.com/sharp/aspnetmvc/>

[6] Руководство по Entity Framework Core 6 [онлайн ресурс] URL: <https://metanit.com/sharp/efcore/>

[7] Несвижский А. Рябов В. Академия Microsoft. Современные веб-технологии [онлайн ресурс] URL: <https://intuit.ru/studies/courses/611/467/info>

[8] Разработка веб-приложений. Основы ASP.NET 2.0 [онлайн ресурс] URL: [https://intuit.ru/studies/professional\\_retraining/942/courses/69/info](https://intuit.ru/studies/professional_retraining/942/courses/69/info)

## ЖОҒАРЫ БІЛІМ БЕРУ ЖҮЙЕСІНЕ ЗАТТАР ИНТЕРНЕТ (IoT) ТЕХНОЛОГИЯСЫН ЕНГІЗУ МҮМКІНДІКТЕРІ

*Байшоланова К.С. - э.ғ.д., профессор м.а. Әл-Фараби атындағы Қазақ ұлттық университеті*

*Бақытжанұлы А. - Әл-Фараби атындағы Қазақ ұлттық университеті, 2-курс магистранты*

**Аннотация:** Алдағы жылдары технология оқуға көптеген жолдармен әсер етеді. Интернет заттары (IoT) ақпараттық-коммуникациялық технологиялары қоғамның дамуында өзінің маңызды рөлін растауды жалғастыруда. IoT қолдауымен білім беру мекемелері оқу тәжірибесін ұсыну, тиімділікті арттыру және қажетті кезде білім алушылардың үлгерімі туралы тиімді түсінік алу арқылы оқу нәтижелерін жақсарты алады. IoT технологиясының жоғары білімге, әсіресе университеттерде әсері туралы зерттеулер келтірілген. IoT университеттердің жұмыс әдістерін түбегейлі өзгерте алады, білім алушылардың көптеген пәндер бойынша және кез-келген деңгейде оқуын жақсарты алады. Ғалымдар, зерттеушілер, оқытушылар мен студенттер IoT жүйелерін, құрылғыларды, қосымшалар мен қызметтерді дамыту мен іске асыруды басқаратын ерекше орынға ие.

Интернет заттары (IoT) - бұл біздің күнделікті өміріміздің көптеген аспектілерін өзгерту процесі. IoT технологиялары алдыңғы инновациялардан ерекшеленеді, өйткені олар кең таралған, олар зияткерлік және автономды орта жасайды. Заттар интернетіндегі жетістіктер адам қызметінің барлық салаларындағы қазіргі заманғы негізгі стратегиялық бағыт болып табылады. Жалпы сенсорлар мен физикалық әлем мен виртуалды әлем арасындағы алшақтықты жою мүмкіндігі оқытудың жаңа моделінің тұжырымдамалық негізі ретінде қабылданады. Физикалық әлем миллиардтаған нысандарды-құрылғыларды қолданыстағы интернет инфрақұрылымына қосу үшін кез-келген объектіге сенсорларды салу идеясының арқасында интернетте пайда болады.

IoT технологиясы тез дамып келеді және бүкіл әлемде қолданысқа жеңіл қарапайым жүйеге айналууда. IoT көптеген салаларда, соның ішінде оқу орындарында, әсіресе университеттерде өзгеріп келеді. Енді университеттерде IoT үшін техникалық әзірлемелер мен инновациялық модельдерді басқару мүмкіндігі бар.

IoT - бұл нақты протоколдарға сәйкес ақпарат алу құрылғылары арқылы ақпарат алмасу мақсатында ішкі және сыртқы ортамен өзара әрекеттесу үшін интернет инфрақұрылымымен құрылғыларды, объектілерді және заттарды біріктіретін ғаламдық физикалық желі. Бірқатар ғылыми-зерттеу ұйымдары мен талдаушылар IoT технологиясының болашағы мен оның интернетке тигізетін әсерін болжайды: әлемдегі сандық сенсорлары бар және кез-келген желіні пайдалану арқылы өзара байланысып миллиардтаған физикалық құрылғыларды басқарады. IoT қосымшалары денсаулық сақтау, қауіпсіздік және сауда сияқты салаларда қолданылады. Сонымен қатар мектептер мен университеттерде интеграциялау жүзеге асуда. Университеттердің болашағы қол жетімді технологияларды қолдануда емес, университеттер болашақ қызметкердің, болашақ экономиканың өзгертін қажеттіліктеріне қалай бейімделетіні туралы негізделген. IoT технологиялық әзірлеушілерге энергияны аз тұтынатын және кез-келген құрылғыға біріктіруге болатын шағын және қол жетімді сымсыз жүйелерді құруға мүмкіндік береді. Мысалы, Digital Campus System жүйесі студенттерге ақпараттың барлық түрлерін алудың маңызды платформасы болып табылады. Жоғары оқу орындарына, әсіресе университеттерге, академиялық және ғылыми қызметкерлердің сандық ортада тиімді жұмыс істеуі үшін олардың мазмұны мен қызметін цифрландыруға және әдістерін бейімдеуге қажеттілік артып келеді. Технологияны толығымен біріктіретін жақсы жобаланған физикалық кампус студенттердің тәжірибесін жақсарту, оқыту және зерттеу үшін тиісті параметрлер мен мүмкіндіктер беру арқылы сандық университет брендин

құрудың негізі болып табылады. Цифрлық университет оқытуды қамтамасыз ететін және бірлескен зерттеулерге ықпал ететін технологияға ие болуы керек. Қазіргі заманғы цифрлық қауіптердің барлығы университеттерге бәсекелес бола алады, алайда олардың нарықтық жағдайларға жаңашылдық немесе шабуыл жасай алатындығына кепілдік беретін стратегияларды енгізу үшін икемділік, платформалар немесе тиісті құрылғылар бар[1].

Сандық кампуста технологиялар операциялық шығындарды азайтып, қауіпсіздікті арттырады және зерттеушілерге, ғалымдарға, студенттерге және қызметкерлерге құралдар ұсына алады. Бұл артықшылықтар университеттің қызметі мен дамуы, студенттер мен зерттеушілердің тәжірибесі үшін нақты құндылықты қамтамасыз етеді. IoT дамуымен жоғары білім беруде болжамды қолдану үш аспектіден тұрады: студенттерді прогрессивті бағалау, қолданыстағы оқыту платформаларын біріктіру және аралық оқыту бағдарламалық жасақтамасын дамыту болып табылады. Бұл өзгеріс студенттерге ыңғайлылықты арттырады және оқытушылар үшін оқыту процесін тиімді етеді. Оқыту нақты студенттің жеке қажеттіліктеріне көбірек қатыса алады және бағдарлай алады. Сонымен қатар, IoT нақты уақыт режимінде оқушылардың үлгерімі туралы тиімді түсінік бере отырып, оқу сапасын жақсартуға мүмкіндік алады. Қазіргі уақытта студенттер, әсіресе университет, дәстүрлі кітаптардан планшеттер мен ноутбуктер сияқты жаңа технологияларға біртіндеп көшуде. Электрондық оқытудың жетілдірілген қосымшалары студенттерге өз қарқынымен оқуға мүмкіндік береді және білім алушылар тәжірибеге ие болады. Бұл үлгерім деңгейін арттырады, ал оқытушылар жеке жаттығулар мен білім алушылардың тұрақты бағасын жүргізе алады. Сонымен қатар, IoT технологиясының арқасында профессорлар студенттердің үлгерімі туралы мәліметтерді жинай алады, содан кейін олардың әрқайсысына қаншалықты назар аударуды қажет ететінін анықтай алады. Бұл деректерді талдау білім алушыларға болашақ сабақтардың жоспарлары мен әдістерін дәл өзгертуге көмектеседі. Сонымен қатар, қосылған құрылғылар білім алушыларға динамикалық интерактивті сабақтар өткізуге мүмкіндік береді.

Курстардың жүйелі тәсілі мен мазмұны біртіндеп дамып, жетілдірілуі керек. IoT виртуалды әлем мен физикалық әлемнің біртұтастығын қамтамасыз ететіндіктен, болашақта көптеген жаңа оқыту әдістері мен бағыттар пайда болады [2]. Сонымен қатар, IoT болашақ экономикасы жоғары білім беру саласындағы сарапшылар мен көшбасшылармен және студенттерді оқыту арқылы құрылуы мүмкін [3]. Жоғары білім беру жүйелеріндегі даму визуализациялайтын, жетілдіретін және жаңа технологиялық инновацияларды басқаратын болады. Сондықтан IoT қолдауымен экономиканың болашағын қалыптастыру және құру үшін жоғары білім беру секторы бизнеспен және өнеркәсіптік сектормен жұмыс істеуі керек. Сонымен қатар, жоғары білім беру секторы, әсіресе университеттер, IoT технологиясының болашағын бизнеске арналған курстарды дамыту және студенттер мен зерттеушілерге пәнаралық негізде IoT технологияларын қолданатын жаңа бизнес-идеяларды құруға көмектесу арқылы жүзеге асыра алады. Осылайша, технологияның, яғни заттар интернетінің дамуымен университеттер көптеген мәселелерді шеше алады, мысалы; негізгі ресурстарды бақылау, ақпаратқа қол жетімділікті дамыту, саналы жоспарларды құру және қауіпсіз кампустарды дамыту. Алдағы бес жылда IoT технологиясы оқу мүмкіндіктерін әр түрлі жолмен кеңейтеді. Оқу тәжірибесі әлі де виртуалды болады, студенттер білімді жаңа тәсілдермен қолданады. Оқыту бүкіл әлем бойынша жаңа идеялар мен шешімдер әкелетін оқытушылар мен студенттер үшін керемет тәжірибе болады. IoT жүйелері студенттер мен қызметкерлерді ынталандыру, сонымен қатар оқу жылдамдығын арттыру арқылы жоғары білімге үлкен құндылықтар енгізудің үлкен әлеуетіне ие.

## Пайдаланылган әдибеттер

1. Aldowah, Hanan & Rehman, Shafiq & Ghazal, Samar & Umar, Irfan. (2017). Internet of Things in Higher Education: A Study on Future Learning. Journal of Physics Conference Series. 892. 012017. DOI 10.1088/1742-6596/892/1/012017.
2. High, P.G., High, P. (2015). Gartner: Top 10 Strategic Technology Trends For 2016. Retrieved from <http://www.forbes.com/sites/peterhigh/2015/10/06/gartner-top-10-strategic-technologytrends-for-2016>. 2015.
3. Ma, Li & Zhuo, Yu & Liao, Kuo & Liu, Shen & Qiao, Jie & Han, Zhenyu & Wang, Jin. (2014). Development and Research of Digital Campus System Based on Android. International Journal of Smart Home. 8. 25-36. DOI 10.14257/ijsh.2014.8.4.03

# USE OF METHODS TO DETERMINE DAMAGE TO AGRICULTURAL CROPS (AKMOLA REGION, SHORTANDY DISTRICT)

Moldir Yessenova, Gulzira Abdikerimova, Mukhammed Bolsynbek

*L.N. Gumilyov Eurasian National University  
Satpayev str., Astana, Republic of Kazakhstan*

*E-mail: [moldir\\_11.92@mail.ru](mailto:moldir_11.92@mail.ru), [gulzira1981@mail.ru](mailto:gulzira1981@mail.ru), [mbolsynbek@bk.ru](mailto:mbolsynbek@bk.ru)*

## Abstract

This article discusses methods for analyzing texture images to identify negative factors affecting the growth of wheat on a land plot under the cadastral number 01-012-025-040:42, owned by the A.I. Baryaev Research Institute. To solve the tasks set, the spectral brightness coefficient (SBC), NDVI, orthogonal transformation methods, statistical and structural features, Laws texture masks, K-means method, clustering are used.

## Introduction

Currently, image recognition is one of the main problems in many fields. Problems encountered in everyday life are solved, such as searching for a car registration number, improving road signs, detecting pathology in medicine, and recognizing a person's image. In the world in various fields, that is, genetics, medical research institutes, the Institute of Space Research, the Institute of Automation and Electrometry, the I.S. Vavilov State Optical Institute, etc. texture images are studied in scientific institutions.

One of the tasks identified in the framework of the State Program for the Development of the Agro-Industrial Complex (AIC) of the Republic of Kazakhstan for 2017-2021 is to improve the quality of public services and ensure the introduction of digital technologies in AIC one.

The main goal of this research work is to use methods for identifying crops and pests using information from aerial photography, the growing season and the spectral brightness factor.

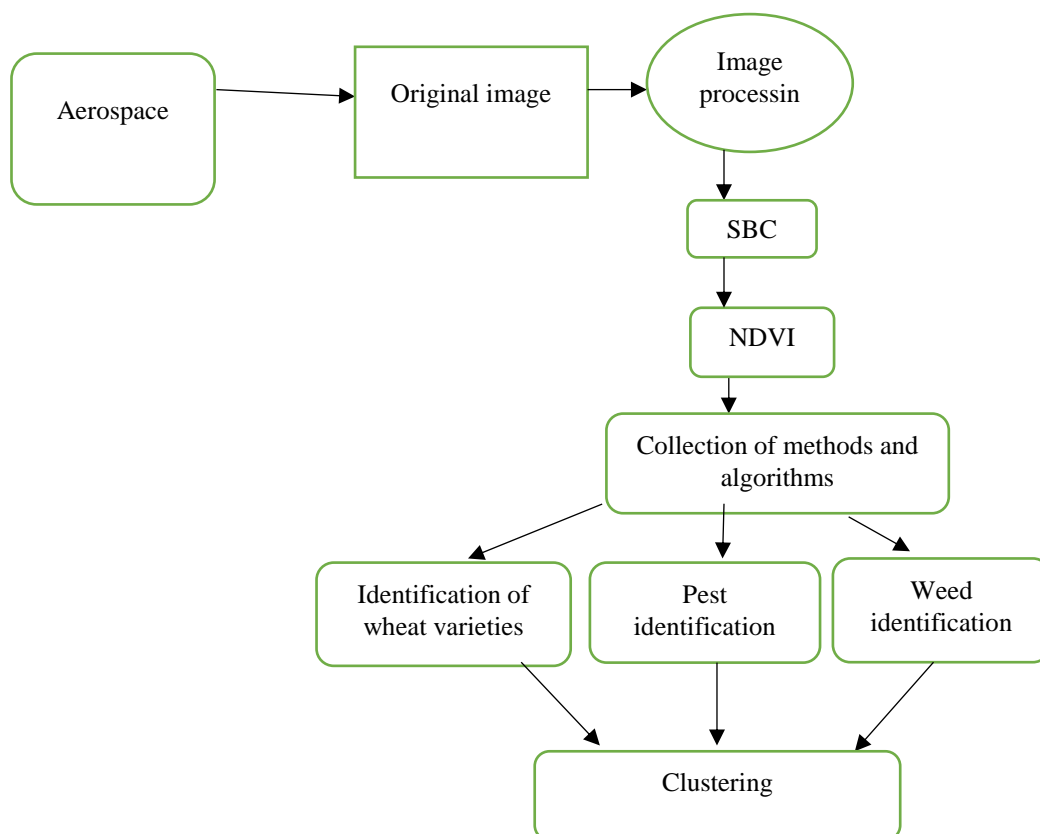


Figure 1. Method for processing aerospace images

[1] Theoretical and practical results of using an orthogonal transformation based on Walsh functions for data compression when transmitting aerial photographs over a communication channel are presented.

In the work [2], orthogonal transformations have shown their effectiveness in compressing documents for storage and transmission in information systems. The authors note that this makes it possible to eliminate the information redundancy of images, depending on the purpose of the system and the characteristic indicators of the images being restored.

The work [12] presents an automatic algorithm for clustering statistical textural features of an image based on histograms, which includes an assessment of the quality of the obtained distribution of feature vectors over clusters. Algorithms were used to classify forest aerospace images, but the article does not consider the classification of crops.

A feature of our work is the application of orthogonal transformations, texture labels and Lava texture masking methods to aerospace images. In particular, the effectiveness of using these methods in determining the factors that negatively affect the growth of agricultural crops, i.e. weeds.

### 3 Textural feature methods

To perform spectral analysis, it is necessary to first decompose the signal or image into frequencies. For this, various sets of basis functions are used [3]. The corresponding algorithms are called transformations: cosine, Hadamard, Haar, oblique, etc. Note that the Haar and Daubechies transformations are the simplest wavelet transforms. The program is implemented in the MATLAB environment and allows performing spectral transformations of six types: 1) cosine, 2) Hadamard of order  $n = 2$ , 3) Hadamard of order  $n = p + 1$ ,  $p = 3 \pmod{4}$  - a prime number, i.e. based on the Legendre symbol, 4) Haar, 5) oblique, 6) Daubechies-4.

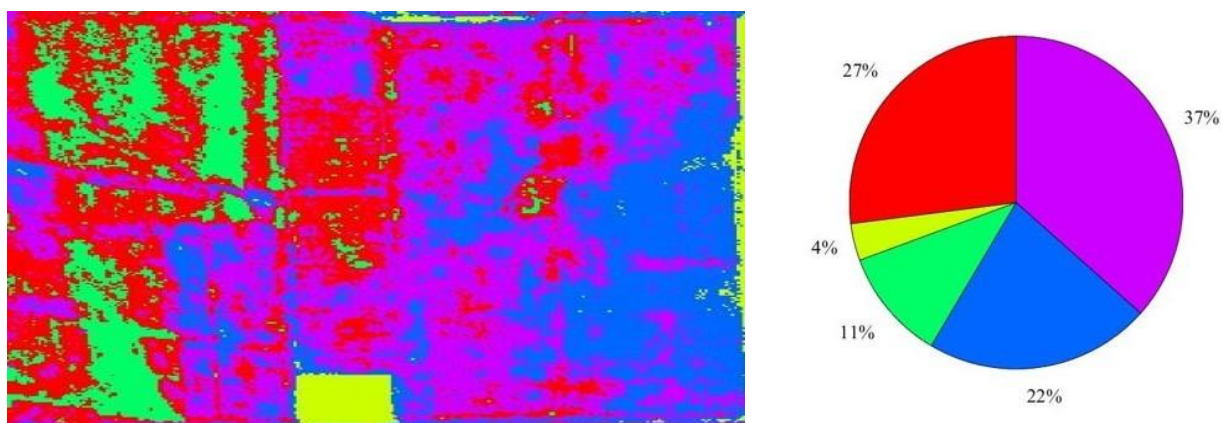


Figure 2. The result of clustering by the lower frequency of the spectrum, where the Haar transform is applied

According to the agronomists of the A.I. Barayev Research Institute, wheat was grown from 05/12/2021 to 05/25/2021. 06/17/2021 weeds are hybridized and removed from 07/18/2021. The table below shows the percentage of weeds in different growing seasons.

Table 1. Identification of weeds by orthogonal transformations in different growing seasons

data \ methods	10.06.2021	01.07.2021	21.07.2021	06.08.2021
Haar	27%	15%	9 %	4%
Dobeshi-4	28%	16%	10 %	5%
Cosine	28%	16%	10 %	5%
Hadamard of the order $n = p + 1$ , $p = 3 \pmod{4}$ - prime number, i. e. based on Legendre's symbol	27%	15%	9 %	4%

Hadamard of order	28%	16%	10%	5%
Slant	27%	15%	9%	4%

The main tasks of the analysis of texture areas include: selection of features that describe texture differences, formation, selection of homogeneous areas, classification of areas with several textures, identification of a particular object by texture. Segmentation methods based on the analysis of large areas of agricultural grain fields on aerial photographs, depending on the characteristics of the texture areas of the images used, can be divided into statistical, structural, fractal, spectral and combined [4]. With the help of the complex use of statistical and structural signs of the second order from the methods of textural signs, the difference between agricultural crops and weeds is determined. Figure 3 shows informative textural features.

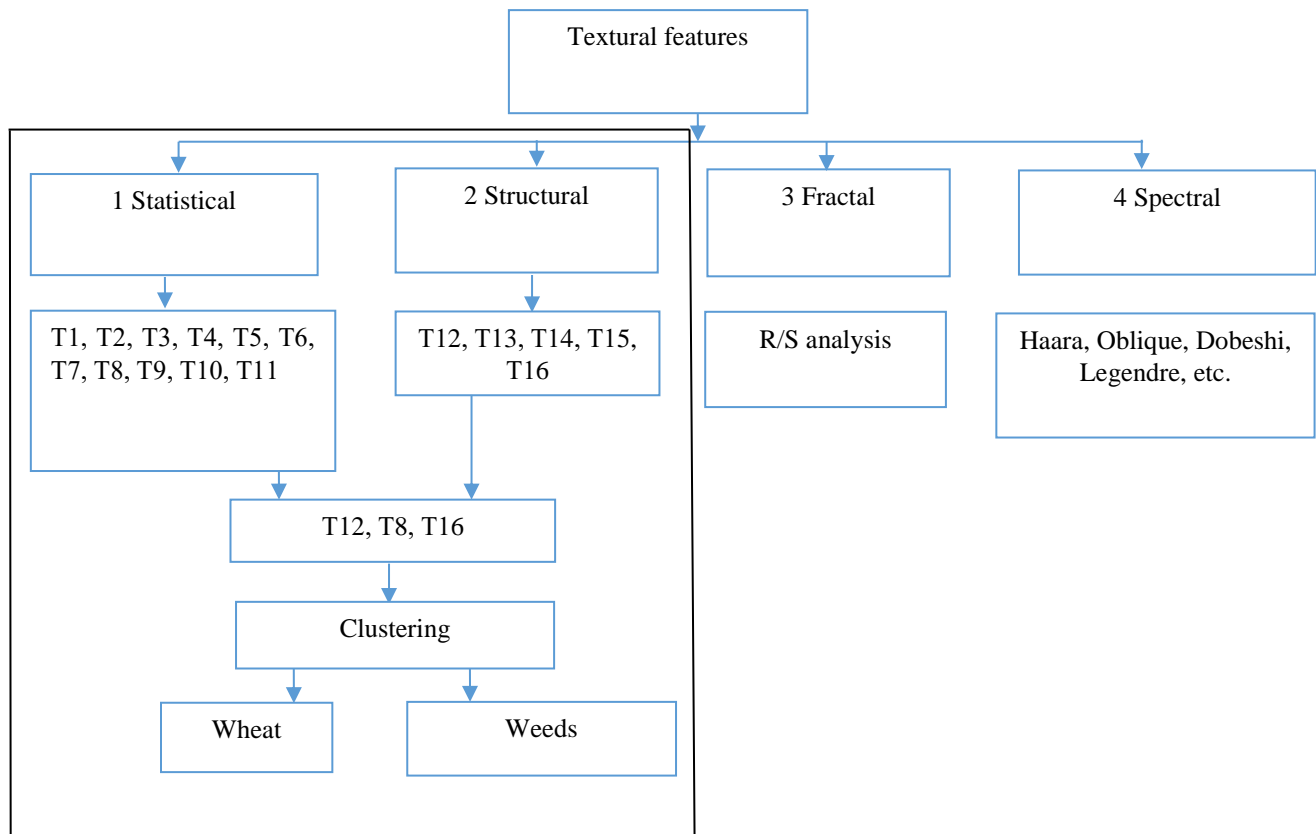


Figure 3. Informative texture symbols

In the scientific literature, with the help of textural features, the problems of many scientific fields are solved. Agricultural cereals - informative textural features that determine the foci of weeds in wheat:

8 The overall average value is determined from the histogram of the sum of brightness values for pairs of video elements closely related to the adjacency matrix:

$$T8 = \sum_{i=2}^{2N} i \cdot p_{x+y}(i);$$

here  $p_{x+y}(k) = \sum_{i=1}^N \sum_{j=1}^N i+j=k p(i,j)$ ,  $k = 2,3,\dots,2N$  - distribution along the tangent diagonal.

12 Inverse torque is a measure of local similarity that determines the clarity of short series:

$$T_{12} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{p(i, j)}{j^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j)};$$

16. Share of the image in the series:

$$T_{16} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} jp(i, j)}$$

Considered textural features are extracted [5]. In this work, out of 14 statistical and structural textural features, 3 informative textural features, i.e., T12, T8, T16, identifying wheat and its weeds, are identified. In Figure 4, clustering is carried out in a standard way, that is, the size of the current window is 3x3.

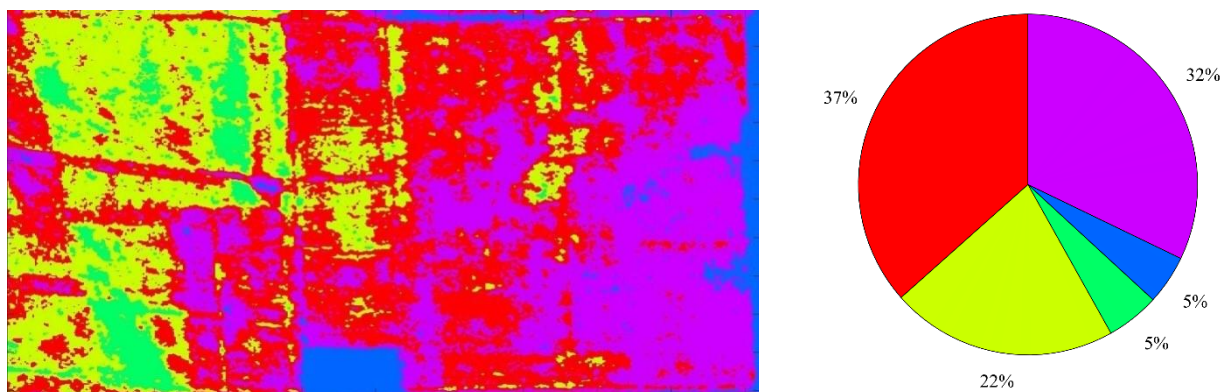


Figure 4. The result of clustering by textural features T12, T8, T1

Table 2. Results of the value of clustering by textural features

data	09.06.2021	01.07.2021	21.07.2021	06.08.2021
methods				
Textural features (T12, T8, T16)	22 %	12 %	7 %	2 %

#### 4 Conclusion

The studies were carried out on wheat lands in the Akmola region of Northern Kazakhstan. The data provided by the A.I. Barayev Research and Production Center were analyzed: the timing of wheat sowing, weed poisoning, the timing of wheat harvesting. The original images used in this work were obtained from Planet.com, where daily information about the Earth is available.

1. The possibility of using orthogonal transformation methods for processing aerial photographs is determined.;
2. By processing aerospace images, informative textural features were revealed to identify weed foci that negatively affect the growth of an agricultural crop - wheat. (T15, T9, T18);
3. The area of damage to agricultural cereals was determined using the Laws texture mask method.

All published results are consistent with data provided by the A.I. Barayev Research Institute. In the future, using the most effective of the mentioned methods in machine learning, an automated application will be developed that will develop a system of actions to optimize various resources, increase productivity and eliminate factors that negatively affect high performance, without harming the environment.



## References

- [1] Kostrov B. V. et al. Theoretical aspects of aerospace image processing in quasi two-dimensional spectral space //MATEC Web of Conferences. – EDP Sciences, 2016. – Т. 75. – С. 03006.
- [2] Abdikerimova G. B. et al. The analysis of textural images on the basis of orthogonal transformations //Journal of Theoretical and Applied Information Technology. – 2019. – Т. 97. – №. 1. – С. 15-22.
- [3] Sidorova, V. S. (2012). Hierarchical cluster algorithm for remote sensing data of earth. Pattern Recognition and Image Analysis, 22(2), 373-379.
- [4] Yessenova M. B. et al. The applicability of informative textural features for the detection of factors negatively influencing the growth of wheat on aerial images // Eastern-European Journal of Enterprise Technologies. – 2022. – Т. 4. – №. 2. – С. 51-58.
- [5] Харалик Р. М. Статистический и структурный подходы к описанию текстур //тиизэр. – 1979. – Т. 67. – №. 5. – С. 98.

# ИСПОЛЬЗОВАНИЕ МЕТОДА БЛИЖАЙШЕГО СОСЕДА ДЛЯ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ

Н.Ш.Максимова, Ж.Ж.Хамитова

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

**Аннотация.** Успешная реализация алгоритмов классификации изображений имеет большую научную и практическую значимость. Классификация изображений является фундаментальной задачей компьютерного зрения, и открывает новые прикладные возможности такие как визуальные поисковые системы, классификация видео, системы с дополненной реальностью, и многие другие возможности в индустрии развлечений и сфере безопасности. С точки зрения науки, классификация изображений продвигает исследовательскую работу в таких направлениях как мультимодальное обучение, когда результаты в классификации изображений используются для автоматического описания изображений естественным языком.

**Ключевые слова:** метод ближайшего соседа, классификация изображений, выборка, NN-классификатор

**Введение.** В эпоху цифровых технологий производится огромное количество неструктурированных данных. К примеру, ежедневно около 70 млн. фотографий загружаются на сервис Instagram [6]. 500 млн. твитов в день, и около 200 млрд. в год появляются в Twitter [6]. 300 часов видео загружаются на Youtube каждую минуту, более 1 млрд. пользователей просматривают сотни миллионов часов видео в день [6]. В поисковой системе Google производится более 3.5 млрд. запросов в день, и 1.2 трлн. запросов в год [6].

До недавнего времени эти данные просто накапливались. Мир находился в фазе сбора данных. С появлением технологий анализа больших данных мы плавно переходим в фазу использования этих данных для предсказательной аналитики, и выявления шаблонов.

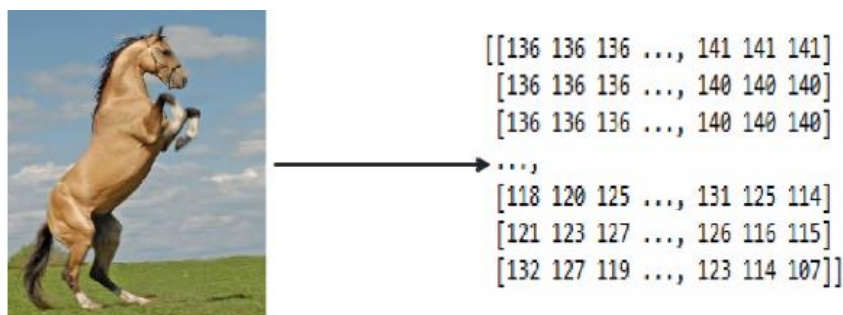


Рис. 1 - Представление изображения в виде матрицы

Классификация изображения – это задача присвоения входному изображению метки из набора существующих категорий. Эта одна из центральных задач в Computer Vision имеющая большое количество приложений. К примеру, изображение на рисунке 1 имеет размер 494 на 345 пикселей, и 3 цветовых канала (Red, Green, Blue или RGB). Соответственно, изображение представляет собой

$494 \times 345 \times 3$  или же 511290 чисел.

Каждое число является целочисленным, принимающим значения от 0 (черный цвет) до 255 (белый цвет). Задача классификации – получить из полумиллиона чисел единственную метку, а именно «лошадь».

Данная задача примитивна для человека, но представляется сложной для алгоритмов, используемых в Computer Vision. Вот некоторые проблемы с которыми сталкиваются алгоритмы [7]:

- Разнообразие точек зрения - один и тот же объект может быть сфотографирован с разных ракурсов
- Разнообразие масштаба – объекты в реальном мире имеют вариации в размерах, а также разные масштабы на изображениях
- Деформация – многие объекты могут быть деформированы, так как не имеют строго фиксированной формы
- Окклюзия – ситуация, когда только небольшая часть объекта видима.
- Условия иллюминации
- Фоновый беспорядок – ситуация, когда объекты интереса смешиваются с окружающей средой
- Межклассовая вариация – объекты внутри класса могут быть разнообразными.

В таблице 1 приведены примеры изображений, которые могут вызвать трудности в задаче классификации изображений



Хорошая модель классификации должна быть инвариантна к такому многообразию представлений изображения. Для алгоритма классификации используется подход основанный на анализе данных, т.е. алгоритму предоставляется множество примеров изображения из каждого класса (*тестовые данные или тестовая выборка*) и алгоритм учится различать эти классы. Пример тестовой выборки для 4-х классов приведен на рисунке 2.



**Рис. 2 – Пример тестовой выборки для 4-х классов изображений**

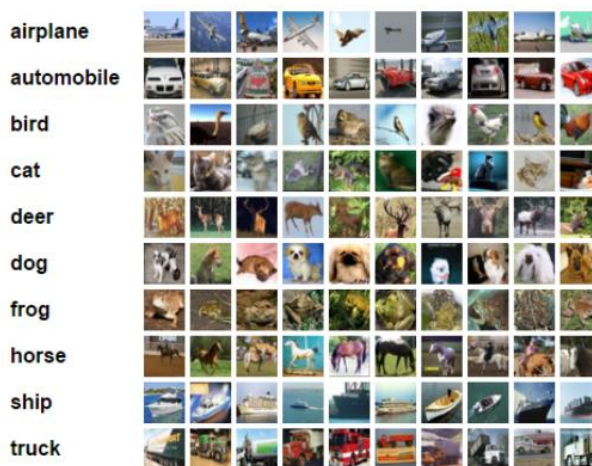
Полная процедура классификации выглядит так:

- Вход: множество из  $N$  изображений маркированных одним из  $K$  классов. Это множество называется *обучающей выборкой*.

- **Обучение:** задачей является изучение представления каждого из классов при помощи обучающей выборки. Это называется *обучением классификатора* или *обучением модели*.

- **Оценка:** оценка качества классификатора при помощи нового набора изображений, который не был использован при обучении. Предсказание классификатора сравнивается с истинной меткой каждого нового изображения и получается численная характеристика качества модели (точность классификации).

Для понимания задачи классификации рассмотрим классификатор на основе задачи поиска ближайшего соседа (Nearest Neighbor или NN классификатор). В качестве данных используем базу CIFAR-10 [8]. Данная база содержит 60000 изображений размером 32 на 32 пикселя. Каждое изображение имеет одну из 10 меток. Изображения разделены на два множества: 50000 для обучения (обучающая выборка), и 10000 для тестирования (тестовая выборка). На рисунке 3 приведен пример изображений для каждого из 10 классов [8].



**Рис. 3 – Пример изображений из базы CIFAR-10. 10 классов или категорий: самолет, автомобиль, птица, кошка, олень, собака, лягушка, лошадь, корабль, грузовик.**

Предположим необходимо используя 50000 данных из обучающей выборки (5000 изображений на каждый класс) классифицировать оставшиеся 10000 изображений из тестовой выборки. При использовании NN классификатора для каждого примера из тестовой выборки производится его сравнение со всеми примерами из обучающей выборки, и в качестве результата классификации выдается метка ближайшего примера. В качестве метрики для сравнения можно использовать L1-метрику, т. е. сравнить все пиксели двух изображений и просуммировать значения отличающихся пикселей:

$$d1(I1, I2) = \sum p |I_{p1} - I_{p2}|,$$

где  $I1, I2$  – изображения представленные в виде вектора,  $p$  – пиксель.

Еще одной метрикой может быть L2-метрика, которая измеряет Евклидово расстояние между векторами:

$$d2(I1, I2) = \sqrt{\sum p (I_{p1} - I_{p2})^2} \text{---} \sqrt{k\text{-NN}}$$

( $k$ - Nearest Neighbor классификатор) является обобщением метода ближайшего соседа, но вместо одного

ближайшего примера находит  $k$  похожих примеров. Этот метод часто работает немного лучше.

### **Проверочная выборка для настройки гиперпараметров модели**

Какое число  $k$  необходимо выбрать для алгоритма  $k$ -NN? Как метрика лучше L1 или L2? Эти параметры называются *гиперпараметрами* и часто их значения получают эмпирическим путем из обучающей выборки. Нельзя использовать тестовую выборку для

поиска значений гиперпараметров. Иначе, модель может быть *переобучена*, то есть модель хорошо объясняет примеры из выборки, но плохо работает на абсолютно новых примерах. Таким образом, использование тестовой выборки для поиска параметров модели отрицательно влияет на ее *обобщающую способность*. Для подбора гиперпараметров без использования тестовой выборки существует так называемая *проверочная выборка* (validation set).

Проверочная выборка является небольшим подмножеством обучающей выборки. Например, из 50000 обучающих примеров в CIFAR10 можно использовать 45000 для обучения, а остальные 5000 для проверки.

Существует также понятие *перекрестной проверки* (cross-validation). Вместо выбора определенного числа примеров, обучающая выборка делится на  $n$  равные части. Затем  $n-1$  частей используется для обучения модели, а одна часть для проверки. Данная процедура повторяется  $n$  раз. Производительность модели оценивается как среднее  $n$  проверок. На практике перекрестная проверка может быть ресурсоемкой, поэтому использование простой проверочной выборки в большинстве случаев предпочтительней. Для определения размера проверочной выборки нет определенного аналитического метода. Исследователи используют 50% - 90% обучающей выборки для обучения, а оставшуюся часть для проверки. На размер выборки влияют много факторов. Например, чем больше число гиперпараметров, тем предпочтительней большая проверочная выборка.

#### **Выводы по NN классификатору**

NN классификатор обладает рядом преимуществ и недостатков. Он очень прост в реализации и

понимании. Практически отсутствует сам процесс обучения модели, так как необходимо только хранить обучающие данные. Тем не менее, тестирование новых данных очень ресурсоемкий процесс, так как необходимо сравнение нового примера со всей обучающей выборкой. На практике более желательным является уменьшение времени тестирования, потому что это подходит к реальным сценариям использования. NN классификатор подходит к задачам в которых данные не многомерные. Более того, NN классификатор работает только на уровне пикселей, не принимая во внимание семантику изображения. На рисунке 4 приведен такой пример. Изображения б) и в) находятся на одинаковом Евклидовом расстоянии от оригинала, соответственно они будут классифицированы одинаково.



**Рис. 4 – Пример отсутствия семантики в алгоритме k-NN**

#### Список литературы

- [1]. Вьюгин В. В. Математические основы теории машинного обучения и прогнозирования. - Москва: МЦМНО, 2013. – С.150.
- [2]. Karpathy A., Joulin A., Li F.-F. Deep Fragment Embeddings for Bidirectional Image Sentence Mapping // Advances in Neural Information Processing Systems 27 (NIPS 2014). - 2014.
- [3]. Krizhevsky A. Learning Multiple Layers of Features from Tiny Images: MSc thesis.

- 2009.

[4]. Karpathy A., Joulin A., Li F.-F. Deep Fragment Embeddings for Bidirectional Image Sentence Mapping // Advances in Neural Information Processing Systems 27 (NIPS 2014). - 2014.

[5]. Compositional Semantics for Finding and Describing Images with Sentences // Transactions of the Association for Computational Linguistics. - 2014.

[6]. Google Search Statistics - Internet Live Stats: [Электронный ресурс]. URL: <http://www.internetlivestats.com/google-search-statistics/>.

[7]. Mitchell T. M. Machine Learning. – М.: McGraw-Hill, 1997. – С.10.

[8]. Samuel A. L. Some Studies in Machine Learning Using the Game of Checkers - М.: IBM Journal, 1959. - С. 210-229.

# СВЕРТОЧНАЯ НЕЙРОННАЯ СЕТЬ ПРИМЕНИТЕЛЬНО К КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ

Н.Ш.Максутова, Ж.Ж.Хамитова

*Евразийский национальный университет им. Л.Н.Гумилева, Астана, Казахстан*

## **Аннотация.**

В данной статье была изучена одна из моделей машинного обучения с учителем, а также исследованы вопросы ее практического применения к классификации задач. Рассматривался метод сверточной нейронной сети. Данная модель была применена к базам рукописных изображений MNIST и цветных натуральных изображений CIFAR-10. Модель показала высокую точность классификации на рукописных изображениях MNIST. Для сверточной нейронной сети был зафиксирован высокий результат, 99.06%. Сверточная сеть является многослойной нейронной сетью со специальными слоями свертки и подвыборки. Для базы MNIST экспериментальные исследования были проведены на процессоре и графическом процессоре (GPU). При анализе было обнаружено, что сверточная сеть использует около 70000 секунд. На GPU эти цифры уменьшаются до 40 раз. Высокие результаты классификации рукописных цифр объясняются простотой исходных данных. Цифры белого цвета написаны на черном фоне и не имеют множества вариаций как натуральные изображения, поэтому даже простейшие модели достигают высоких показателей.

**Ключевые слова:** сверточная нейронная сеть, классификация изображений, выборка, NN-классификатор

## *Введение.*

В современном мире большое значение имеет задача классификации изображений, это работа с анализом больших данных. Машинное обучение (machine learning) является основой интеллектуального анализа данных. Согласно Самуэлю [1], машинное обучение – это процесс, в результате которого машина (компьютер) способна показывать поведение, которое не было в нее явно заложено. Митчелл [2] дает более конкретное определение: «говорят, что компьютерная программа обучается на основе опыта  $E$  по отношению к некоторому классу задач  $T$  и меры качества  $P$ , если качество решения задачи из  $T$ , измеренное на основе  $P$ , улучшается с приобретением опыта  $E$ ». Машинное обучение использует методы теории вероятностей и математической статистики. Существует несколько видов машинного обучения:

- обучение с учителем или supervised learning;
- обучение без учителя или unsupervised learning;
- обучение с подкреплением или reinforcement learning;

В данной статье рассматривается задача машинного обучения с учителем применительно к классификации изображений, используя модель сверточных нейронных сетей.

Сверточная нейронная сеть или Convolutional Neural Network (CNN) – это нейронная сеть в которую заложено предположение о том, что входные данные имеют сеточную топологию. К примерам таких данных относятся временные ряды (одномерные сетки), изображения (двухмерная сеть пикселей), и др. В основе этих сетей лежит так называемая *операция свертки* (convolution operation).

Операция свертки позволяет улучшить модель машинного обучения благодаря трем важным компонентам: разреженные взаимодействия (sparse interactions), деление параметров (parameter sharing), и эквивариантное представление (equivariant representation). Традиционная нейронная сеть использует матричное перемножение для описания взаимодействий между входами и выходами нейронов. В таком случае нейронная сеть является полносвязной, т.е. в каждый нейрон слоя  $l$  входят данные от всех нейронов слоя  $l-1$ . К примеру, в базе CIFAR-10 изображения имеют размер  $32 \times 32 \times 3 = 3072$ ,

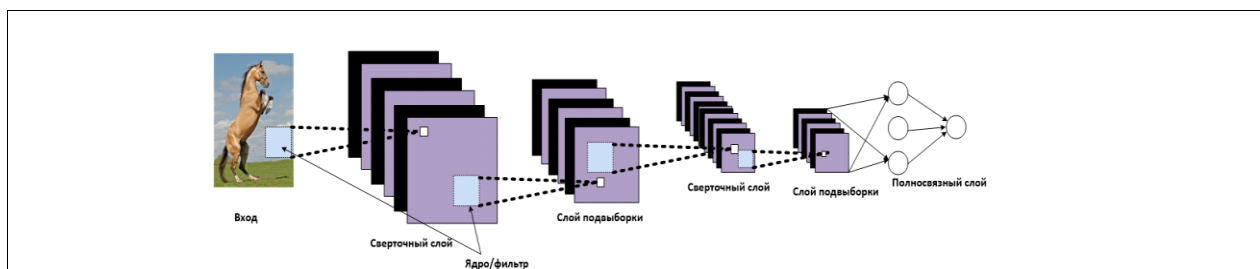
поэтому один полносвязный нейрон в первом скрытом слое имеет 3072 параметра. Для более реалистичных размеров изображений число параметров существенно больше. Полносвязность сети требует больших вычислительных ресурсов и увеличивает число параметров, что в свою очередь ведет к переобучению модели [3]. В сверточной сети, нейроны разрежены. Это достигается за счет того, что ядро или фильтр меньше по размерам чем вход. Изображение может иметь миллионы пикселей, но ядро может быть небольшой матрицей из сотен пикселей, которая определяет границы или другие признаки в изображении. Разреженность также уменьшает число параметров, которые нужно хранить, и предотвращает переобучение модели [4]. Также уменьшается время обучения модели. Если имеется  $m$  входов и  $n$  выходов, то матричное перемножение для  $m \times n$  параметров требует  $O(m \times n)$  времени для каждого примера (изображения в обучающей выборке), в то время как в CNN нужно  $O(k \times n)$ , где  $k$  – число связей выходного нейрона. Во многих практических приложениях  $k \ll n$ .

В CNN нейроны и данные организованы в трех плоскостях: ширина, высота и глубина. Например, для изображения в CIFAR-10 ширина 32, высота 32, и глубина 3. Нейроны в сети будут соединены с небольшим регионом из предыдущего слоя вместо того, чтобы соединяться со всеми нейронами как в полносвязной нейронной сети. В архитектуре сверточной нейронной сети используется 3 типа слоя:

- сверточный слой (convolutional layer или CONV);
- слой субдискретизации или подвыборки (pooling layer или POOL);
- полносвязный слой (fully-connected layer или FC).

Эти слои накладываются друг на друга и чередуются определенное количество раз для построения полной сети. К примеру, простая сверточная сеть для CIFAR-10 может иметь следующую архитектуру: INPUT → CONV → RELU → POOL → FC.

Таким образом, от слоя к слою первоначальное изображение трансформируется в баллы классов. При этом не все слои имеют параметры.

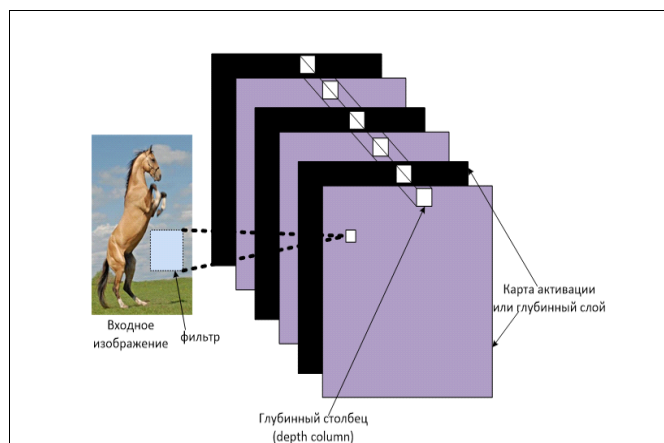


**Рис. 1 – Пример сверточной сети без нелинейной функции активации**

### Сверточный слой

Параметры сверточного слоя – это множество фильтров/ядер, которые можно выучить по мере тренировки нейронной сети. Каждый фильтр или ядро является небольшой по ширине и высоте областью, которая покрывает входной слой по всей глубине (рисунок 2). В алгоритме прямого распространения каждый фильтр сдвигается (выполняет операцию свертки) по ширине и высоте входного пространства, образуя 2-х мерную карту активации для данного фильтра. Во время сдвига фильтра вычисляется скалярное произведение между элементами фильтра и входными данными. Накладывая карты активации для всех фильтров по направлению глубины входного пространства, получаем выходное пространство. Каждый элемент выходного пространства можно интерпретировать как выход нейрона, который использует только небольшую область входа и разделяет параметры с нейронами в одной карте активации.





**Рис. 2 – Детали сверточного слоя**

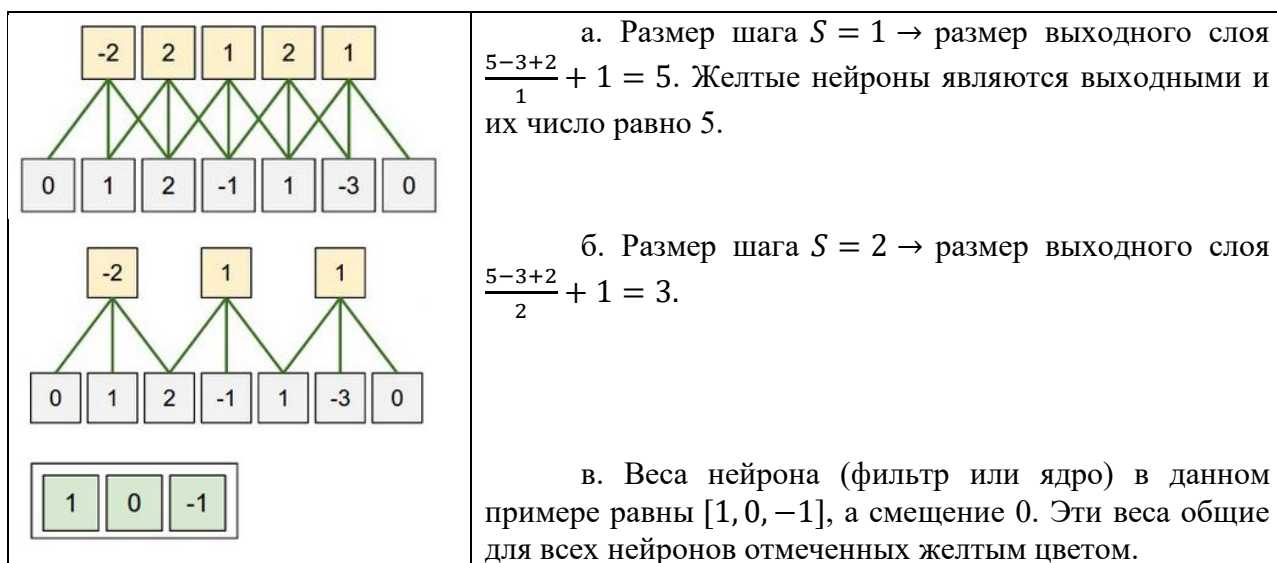
В сверточном слое, каждый нейрон соединяется с локальной областью, размер которого определяется гиперпараметром называемым рецептивным полем (receptive field) нейрона. Рецептивное поле ограничивает размер области входного пространства только по ширине и высоте. Например, в CIFAR-10 изображение имеет размер  $32 \times 32 \times 3$ . Если размер рецептивного поля составляет  $5 \times 5$ , то каждый нейрон в сверточном слое будет иметь веса из входного пространства размером  $5 \times 5 \times 3$ , т.е. 75 весов или параметров.

Размер выходного пространства сверточного слоя определяется 3 гиперпараметрами: глубина (depth), шаг (stride), и заполнение нулями (zero-padding).

Итак, размер выходного пространства можно задать как значение функции от размера входного пространства ( $W$ ), рецептивного поля нейрона сверточного слоя ( $F$ ), шага ( $S$ ), и количества заполнения нулями ( $P$ ):

$$\frac{W - F + 2P}{S} + 1.$$

Данное число определяет количество нейронов выходного слоя, и должно быть целочисленным. Если оно не целое, то вероятнее всего неправильно подобран размер шага. На рисунке 3 показан пример сверточного слоя и расчет размера его выхода.



**Рис. 3 – Иллюстрация архитектуры сверточного слоя. Данный пример адаптирован из [26]. Имеется только одна размерность (горизонтальная), один нейрон с размером рецептивного поля  $F=3$ , размер входного слоя  $W=5$ , и заполнение нулями  $P=1$ .**

Из рисунка 2 видно, что используется заполнение нулями с краев входных данных. При размере шага размер заполнения гарантирует, что размеры входного и выходного пространств одинаковы. Подбор параметров зачастую осуществляется эмпирически.

### Слой подвыборки

Слой подвыборки или субдискретизации (pooling layer) вставляется между сверточными слоями в архитектуре CNN. Задачей данного слоя является уменьшение пространственных размеров данных с целью уменьшения параметров и вычислений в системе и, соответственно, контроля переобучения. Слой подвыборки применяется к каждому глубинному слою (карте активации), и уменьшает ее размер используя различные операции. Наиболее популярной операцией является взятие максимума или max pooling. Например, сетка размером  $2 \times 2$  может накладываться на глубинный слой с шагом 2, и уменьшать размерность входа на 2 единицы по ширине и высоте. Это уменьшает размерность на 75%. В общем виде на вход в слой подвыборки поступают входные данные размерностью  $W_1 \times H_1 \times D_1$ , и задаются два гиперпараметра: размер поля  $F$  и шаг сдвига  $S$ . На выходе из слоя получается массив данных размерностью  $W_2 \times H_2 \times D_2$ , где:

$$W_2 = \frac{W_1 - F}{S} + 1, \quad H_2 = \frac{H_1 - F}{S} + 1, \quad D_2 = D_1.$$

При этом в pooling слое не вводятся новые параметры, и не используется заполнение нулями. На практике встречаются два варианта слоя подвыборки:

$F=3, S=2$  – перекрестный слой подвыборки (overlapping);

$F=2, S=2$  – неперекрестный слой подвыборки.

### Архитектура сверточных сетей

Все сверточные сети имеют следующую архитектуру:

$$INPUT \rightarrow ([CONV \rightarrow RELU]_N^* \rightarrow POOL?)_M^* \rightarrow [FC \rightarrow RELU]_K^* \rightarrow FC,,$$

где  $_N^*$ ,  $_M^*$ ,  $_K^*$  означает, что последовательность операций повторяется  $N \geq 0$ ,  $M \geq 0$ ,  $K \geq 0$  раз соответственно. При построении сверточной сети в [26] советуют следовать следующим положениям:

- размерность изображения должна быть делима на 2 большее количество раз;
- размер фильтра в сверточном слое должен быть небольшим, не более  $5 \times 5$ ;
- использовать max pooling с размером поля  $2 \times 2$ .

Ограничением при разработке архитектуры сети является размер оперативной памяти. Например, при размере изображения  $32 \times 32 \times 3$  и использовании 3-х сверточных слоев с размером рецептивного поля  $F=3$ , шагом сдвига  $S=1$ , и заполнения нулем  $P=1$ , и 72 фильтрами на каждый сверточный слой получаем 3 массива активаций размерностью  $32 \times 32 \times 72$ . Таким образом общее число параметров равно  $32 \times 32 \times 72 \times 3 = 221184$ . Если каждый параметр имеет тип double размером 8 байт, то объем необходимой памяти для одного изображения равен  $\frac{221184 \times 8}{1024^2} = 1.6875$  или почти два мегабайта. Таких изображений в обучающей выборке CIFAR-10 50000 и для такой сети потребуется  $50000 \times 1.6875 : 1024 \approx 82$  гигабайта. Одним из компромиссов является использование большего фильтра и шага сдвига в первом сверточном слое, или уменьшение размера пакета при пакетном градиентном спуске (MGD) [3].

Некоторые сверточные сети имеют названия. Среди них популярными являются: LeNet [5], AlexNet [6], ZF Net [7], GoogLeNet [8], VGGNet [9].

### ЭКСПЕРИМЕНТЫ

Были проведены эксперименты на базах MNIST и CIFAR-10 с использованием модели сверточная нейронная сеть, а также CPU и GPU. Лучшие результаты для каждой модели приведены в таблице 3 и таблице 4.

**Таблица 3. Лучшие результаты для различных моделей на базе MNIST**

<b>CPU</b>			
	<b>Процент ошибок</b>	<b>Точность, %</b>	<b>Время работы (сек.)</b>
Сверточная сеть	0.94	99.06	68779.2
Линейная регрессия	7.4	92.6	21

**Таблица 4. Лучшие результаты для различных моделей на базе CIFAR-10 для GPU**

<b>Модель</b>	<b>Процент ошибок</b>	<b>Точность, %</b>	<b>Время работы (сек.)</b>
Сверточная сеть	37.91	62.09	131520
Новая сверточная сеть	34.6	65.4	142480

**Выводы**

В данной статье была изучена модель машинного обучения с учителем, а также исследованы вопросы ее практического применения к классификации задач. Рассматривалась сверточная нейронная сеть. Данные модели были применены к базам рукописных изображений MNIST и цветных натуральных изображений CIFAR-10. Модель показала высокую точность классификации на рукописных изображениях MNIST. Высокий результат был зафиксирован для сверточной нейронной сети, 99.06%. Сверточная сеть является многослойной нейронной сетью с специальными слоями свертки и подвыборки. Для базы MNIST экспериментальные исследования были проведены на процессоре и графическом процессоре (GPU). В сравнительном анализе было обнаружено, что нейронная и сверточная сеть используют больше всего времени, около 70000 секунд. На GPU эти цифры уменьшаются до 40 раз. Высокие результаты классификации рукописных цифр объясняются простотой исходных данных. Цифры белого цвета написаны на черном фоне и не имеют множества вариаций как натуральные изображения, поэтому даже простейшие модели достигают высоких показателей.

На базе CIFAR-10 результаты не такие высокие, и простейшие модели не достаточны. Сверточная сеть показала точность 62.09%. На базе сверточной сети была разработана новая модель, которая содержит 6 слоев. Данная модель показала результат, 65.4%. Модель была оптимизирована за 100 эпох для сокращения времени эксперимента. При большем количестве эпох можно значительно повысить результат классификации.

Всего в исследовании было проведено 370 экспериментов на CPU и GPU. На основе экспериментальных данных был проведен сравнительный анализ моделей машинного обучения, и исследовано их поведение в зависимости от значений различных гиперпараметров. Важнейшими гиперпараметрами оказались скорость обучения в алгоритме градиентного спуска, размер пакета при пакетном градиентном спуске и используемая функция активации. Большая скорость обучения ускоряет сходимость алгоритма, и позволяет сократить время обучения. Но при этом вероятность оказаться в локальном минимуме поверхности оптимизации возрастает. При маленьком значении скорости обучения время обучения значительно возрастает, но точность классификации намного выше. Оптимальным значением скорости обучения для нейронной сети и сверточной сети оказалось значение 0.01. Размер пакета также является важным гиперпараметром. При большем размера пакета значение вычисляемого градиента в алгоритме оптимизации более точное, но при этом оптимизация в целом более медленная и требует больше оперативной памяти. Для поиска оптимального значения скорости обучения и размера пакета используют перекрестную проверку, т.е. во время процесса оптимизации на подмножестве обучающей выборки периодически проверяют точность. Если точность для определенного значения гиперпараметра уменьшается, то порог после которого происходит уменьшение является оптимальным значением данного гиперпараметра. Данный процесс требует много времени и ресурсов. Оптимизация гиперпараметров является самостоятельным научным направлением в машинном

обучении. В ходе экспериментов было доказано, что для нейронных и сверточных сетей в качестве функций активации лучше всего использовать гиперболический тангенс или выпрямляющий нелинейный блок (ReLU).

### *Список литературы*

- [1]. Samuel A. L. Some Studies in Machine Learning Using the Game of Checkers - М.: IBM Journal, 1959. - С. 210-229.
- [2]. Mitchell T. M. Machine Learning. – М.: McGraw-Hill, 1997. – С.10.
- [3]. Karpathy A., Li F.-F. Convolutional Networks 2015: [Электронный ресурс]. URL: <http://cs231n.github.io/convolutional-networks/>.
- [4]. A. Courville, Y. Bengio, I. Goodfellow. Deep Learning. – MIT Press. - 2015.- С.100-103
- [5]. LeCun Y., Bottou L., Bengio Y., Haffner P. Gradient-Based Learning Applied to Document Recognition // Intelligent Signal Processing. – 2001. - С. 306-351.
- [6]. Krizhevsky A., Sutskever I., Hinton G. ImageNet Classification with Deep Convolutional Neural Networks // Advances in Neural Information Processing Systems 25. - 2012.
- [7]. Zeiler M. D., Fergus R. Visualizing and Understanding Convolutional Networks // CoRR. - 2013. - abs/1311.2901.
- [8]. Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. Going Deeper with Convolutions // CoRR. – 2014. - abs/1409.4842.
- [9]. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition // CoRR. – 2014. - abs/1409.1556.

## АДАМДАРДЫҢ ТОБЫРЛАНУ ДИНАМИКАСЫН МУЛЬТИАГЕНТТІ МОДЕЛЬДЕУ

*Хасенов А.Ш., Мусабеков А.К.*

*Л.Н. Гумилев атындағы Еуразия ұлттық университетінің Ақпараттық технологиялар факультетінің «Ақпараттық жүйелер» кафедрасы, Астана, Қазақстан*

### **Аннотация**

Бұл мақалада адамдардың тобырлану динамикасының өзара әрекеттесуін модельдеу платформасы ретінде қолданылған мультиагентті модельдеу қарастырылады. Көпшіліктің динамикасына қатысты мультиагентті зерттеулері мен әзірлемелері нақты аналитикалық бағалауларды тудыруы, жаяу жүргіншілердің қауіпсіздік деңгейін және қалалық кеңістіктің жұмыс істеу сапасын арттыруы мүмкін. Сонымен қатар, адамдар көп жиналатын динамика қалалық кеңістік үшін үлкен проблемаға айналы жатыр. Демек, бұл сценарийлер сыйымдылықты бағалау кезінде, әсіресе төтенше эвакуация кезінде, көпшіліктің қауіпсіздігіне көбірек көңіл бөледі. Бұл мақалада соңғы бірнеше жылдағы мультиагентті дамуының қысқаша шолуы берілген. Мультиагентті әдістердің қазіргі жағдайын талқылау мультиагентті әдістің тарихын және мультиагентті модельдеуді дамытуды қамтиды. Бұл мақалада тек микроскопиялық тәсілге назар аудара отырып, көпшіліктің динамикасын модельдеудің мультиагентті әдістерінің қазіргі жағдайы талқыланды. Модельдеуге мультиагенттік тәсілдің артықшылықтары қазіргі зерттеу саласындағы физикалық эксперименттерді алмастыру ретінде нақты әлемдегі көпшіліктің іс-әрекетін талдау мен болжаудың ең жақсы әдіснамалық тәсілдерінің біріне айналды.

**Кілт сөздер:** мультиагентті модельдеу, макроскопиялық модельдер, мезоскопиялық модельдер, интеллектуалды агенттер.

### **Кіріспе**

Қоғамдағы өмір қоғамдық орындарда адамды сөзсіз көп кісілік қозғалысқа әкеледі. Қоғамдық аймақ немесе қоршаған орта, әсіресе терминалда, сауда орталығында, стадионда және көшеде динамикалық модельдердің басқа классификациясы бар, әдетте модельдеу тәсілін таңдамас бұрын анықталады. Жаяу жүргіншілер ағынының модельдерінде тобыр динамикасының үш классификациясы бар, атап айтқанда: макроскопиялық модельдер, мезоскопиялық модельдер және макроскопиялық модельдер. Көпшіліктің динамикалық моделін таңдау модельді зерттеу немесе егжей-тегжейлі жүргізу тәсіліне байланысты. Қазіргі уақытта микроскопиялық модельдер көпшіліктің динамикасын модельдеу кезінде назар аударды. Алдыңғы зерттеулер макроскопиялық деңгейде тобыр динамикасының ұжымдық іс-әрекеті микроскопиялық деңгейде жеке адамдар арасындағы өзара әрекеттесудің пайда болуына байланысты екенін дәлелдеді [1]. Микроскопиялық модельдер жеке тұлғаларды егжей-тегжейлі қарастырады және көпшіліктің ішіндегі жеке тұлғалардың қозғалыс сипаттамаларына қатысты. Әр адам басқа адамдармен және қоршаған ортамен қарым-қатынас жасағанда бақыланады. Көптеген микроскопиялық модельдер тобыр қозғалысын дәл болжауға және тобыр қозғалысындағы әртүрлі өзін-өзі ұйымдастыратын құбылыстарды түсіндіруге арналған және типтік микроскопиялық модельдерге жатады; мультиагентті модельдер [2], агенттерге негізделген модельдер [3], клеткалық автомат модельдері [4], әлеуметтік күш модельдері [5] және басқа бөлшектерге негізделген модельдер.

Бұл мақалада мультиагентті модельдеу көптеген интеллектуалды агенттерден тұратын автономды өзара әрекеттесуді модельдеуге арналған компьютерлік жүйедегі мүмкіндіктері мен артықшылықтарының арқасында көпшіліктің іс-әрекет модельдеу тәсілі ретінде таңдалады және қарастырылады. Компьютерлік технологияларда мультиагент әдісін әзірлеу және қолдану әртүрлі жағдайларда имитациялық талдауда жаяу

жүргіншілерді зерттеуге артықшылық береді, өйткені мультиагентті тәсілі гетерогенді адамдарды талдау мен болжаудың ең жақсы әдістемелік тәсілдерінің бірі болып табылады. Сонымен қатар, мультиагентті жүйесінің тұжырымдамасы агент пен қоршаған ортаға байланысты. Агентті үш санатта табуға болады, атап айтқанда: белсенді агенттер, пассивті агенттер және когнитивті агенттер. Сонымен қатар, орталарды үздіксіз, виртуалды және дискретті сияқты жағдайлардың үш түрінде анықтауға болады. Агент пен қоршаған орта жаяу жүргіншілер траекториясының өзара әрекеттесуіне байланысты тығыз байланысты, жанама түрде жаяу жүргіншілер (агенттер) мен олардың ортасы арасында градиент күшін, итеру күшін, кездейсоқ күш пен қарсылық күшін тудырады [7]. Мультиагент, Махудиннің айтуы бойынша [8], бұл толып кетудің өлшемділігін және стресс пен шаршаудың байланысын білдіретін құралдарды әзірлеу. Мультиагентті қосымшасының егжей-тегжейлі мысалын осы мақаланың «уақыт шкаласының инфографикасы» бөлімінен табуға болады. Сонымен қатар, мультиагентті модельдеу және агентке бағытталған модельдеу негіздері бірдей, дегенмен мультиагентті, агентке бағытталған модельдеуге қарағанда интеллектуалды және егжей-тегжейлі зерттеу жүргізе алады.

Топ динамикасын модельдеу үшін мультиагент сияқты сенімді болжау моделін жасау топ динамикасының негізінде жатқан жергілікті өзара әрекеттесулерді түсінуді қамтиды. Демек, көпшіліктің қозғалыс динамикасы негізінен жаяу жүргіншілер мен олардың қоршаған ортасы арасындағы жергілікті өзара әрекеттесулермен анықталады. Жеке қозғалыстар жеке кеңістікті сақтау үшін барлық жерде қозғалу кезінде басқа жеке траекториялардың қозғалысына әсер етуі мүмкін, өйткені олар көпшіліктің қозғалысын талап ететін әлеуметтік күштерді сезінеді [9]. Мусаидтің тағы бір зерттеуі көпшіліктің динамикасы күнделікті жағдайларда тиімді қозғалысты қолдайтын көптеген өзін-өзі ұйымдастыратын іс-әрекеттерді көрсететіні айтылды [10]. Ең көрнекті мысалдардың бірі - екі бағытты жаяу жүргіншілер ағынында бір бағытты жолақтардың өздігінен пайда болуы. Алайда, қозғалыс тығыздығы жоғары болған кезде жаяу жүргіншілердің тегіс ағыны нашарлауы мүмкін, бұл «тоқта және жүр» және көпшіліктің турбуленттілік толқындары сияқты қозғалыс үлгілерін жасайды. Нәтижесінде жаяу жүргіншілердің бір-бірімен және қоршаған ортамен қалай әрекеттесетінін түсіну үшін осы уақытқа дейін топ динамикасын зерттеудің әртүрлі әдістері ұсынылды. Осылайша, мультиагент тарихы жаяу жүргіншілер ағынын басқару үшін мультиагент дамуын бағалай алады.

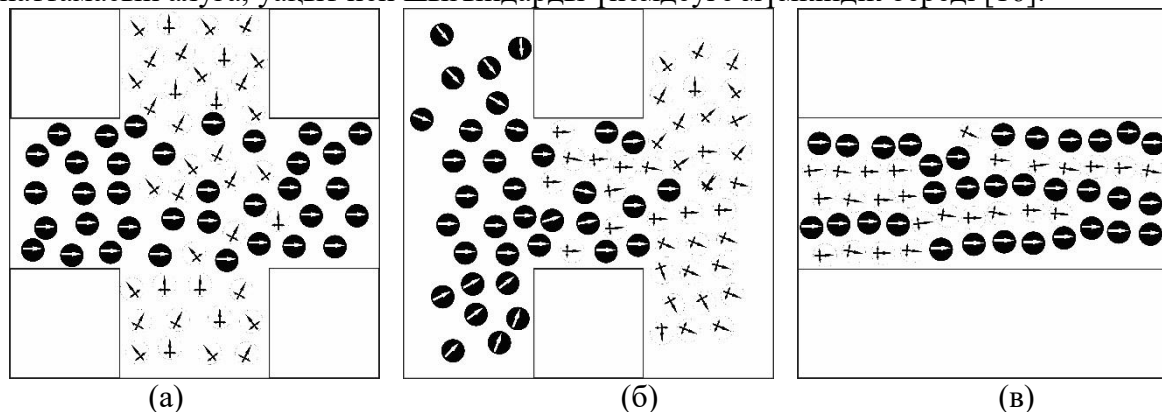
### **1.1. Топ динамикасының мәселелері**

Соңғы бірнеше жылда топ динамикасына қатысты мәселелер (мысалы, адамдар көп жиналатын әрекеттер, табиғи апаттар және шұғыл эвакуация) көптеген модельдеу және имитациялық зерттеулердің басталуына әкелді және проблемаларды шешу үшін әртүрлі жақсартулар жасалды, өйткені жаяу жүргінші тобыр динамикасында болжау мүмкін емес. Осылайша, жаяу жүргіншілердің ағыны мен олардың барлық күрделі ортадағы іс-әрекетіне байланысты жаяу жүргіншінің топ динамикасындағы негізгі рөлін түсіну маңызды. Адамдардың тобырлану динамикасы - бұл адамдардың ниетінен, сенімдерінен, білімінен және қабылдауынан туындайтын әрекеттер жиынтығы, олар зерттеуге әсер ететін әртүрлі топтарды бақылауды қамтиды, бірнеше санаттарға бөлінуі мүмкін, олардың арасында поляризация, бөлісу және экспозицияны жақсарту, доминант топтардың әсері [11]. Сонымен қатар, Самсон және т.б. анықталған топ динамикасы - бұл көпшіліктің жеке қозғалысын, кеңістігін және әрекетін модельдеуді, модельдеуді және түсінуді қажет ететін топтың тығыздығының артуына байланысты пішін мен қозғалыс тұрғысынан топ қозғалысын зерттеу [12]. Сонымен қатар, жаяу жүргіншілерді көпшіліктің қозғалысында қадағалау соңғы кездері топтық динамиканы профильдеу сияқты маңызды және кең қолданбаларының арқасында айтарлықтай назар аударды [13].

Бұрын адамдардың тобырлану динамикасы көпшіліктің белсенділігін модельдеу салаларында зерттелген [14], бірақ нақты жағдайдағы көпшіліктің алдағы жолдарын визуалды тұрғыдан болжауға бірнеше әрекет жасалды [15].

## 1. Модельдеу

Мультиагентті модельдерді немесе агентке негізделген модельдерді, магниттік күш модельдерін, сұйықтық динамикасының модельдерін, әлеуметтік күш моделін, ұялы автомат модельдерін, торлы газ модельдерін, ойын теориясы модельдерін және негізделген тәсілдерді қамтитын топ динамикасын модельдеуі жануарлармен тәжірибелермен негізделген әртүрлі әдістері бар. Бұл мақалада мультиагент таңдалады, өйткені оның артықшылықтары жүйенің нақты пайда болатын құбылысы мен сипаттамасын алуға, уақыт пен шығындарды үнемдеуге мүмкіндік береді [16].



Сурет 1. Жаяу жүргіншілер ағындарының схемасы - қиылысатын ағындар (а), кедергілер (б) және жолақты қалыптастыру (в)

Адамдардың тобырлану динамикасын модельдеу және имитациялау талдауы көрсеткендей, жаяу жүргіншілер ағындарының өзін-өзі ұйымдастыратын құбылыстарына әртүрлі факторлар әсер етеді, мысалы, қиылысатын ағындар (1-сурет(а)), кедергілер (1-сурет(б)), жолақтың қалыптасуы (1-сурет (в)), соққы толқындары, қарсы ағындар, 90 градусқа еңкейту және жаяу жүргіншілердің кемшіліктері. Жаяу жүргіншілердің жүру жылдамдығы, психологиялық қашықтық, көру бұрышы, бұрыштық жылдамдық және жалтару бұрышы сияқты кейбір аналитикалық іс-әрекет сипаттамалары модельдеу және имитациялық әрекеттер процесінде зерттеледі немесе қайта қаралады.

### Қорытынды

Бұл мақалада модельдеу зерттеулерін жүргізетін әртүрлі пәндер мен қауымдастыққа қатысты, мультиагент тарихын және дамуын қамтамасыз ету арқылы топ динамикасының өзара әрекеттесуін зерттеу платформасы ретінде қолданылатын мультиагент тәсілі қарастырылады. Сонымен қатар, бұл мақалада соңғы бірнеше жылдағы мультиагент дамуы көрсетілген. Сонымен қатар, көпшіліктің динамикасын зерттеуде мультиагент дамуы модельдеу зерттеулерін модельдеуде аналитикалық есептеу әдістерін қажет етуі, сонымен қатар әртүрлі салаларды қамтуы мүмкін. Динамикалық модельдеу мен көпшілікті модельдеудің нәтижесі жаяу жүргіншілер инфрақұрылымын (мысалы, сауда орталығы, тротуарлар, стадиондар, жаяу жүргіншілер өткелдері және т.б.) жобалау кезінде тиісті инженерлер үшін анықтамалық ақпарат болады. Топ динамикасын көп агентті модельдеуді дамыта отырып, қоғамдық қалалық кеңістіктердің сапасын аналитикалық бағалауды қамтамасыз етуге, ғимараттардың тиімділігін арттыруға, жаяу жүргіншілердің қауіпсіздігін арттыруға және жарақат алу қаупін азайтуға болады.

### Қолданылған әдебиеттер

1. Abd. Rahman N Abustan M S Che Mat N Gotoh H and Harada E 2015 Applied Mechanics and Materials 802 77-82
2. Schatten M Okreša Đurić B Tomičić, I Ivković N Konecki M Ševa J Terna P 2017 Central European Conference on Information and Intelligent Systems pp 193–200.

3. Helbing D and Balmelli S 2011 How to Do Agent-Based Simulations in the Future: From Modeling Social Mechanisms to Emergent Phenomena and Interactive Systems Design. Santa Fe Institute Working Papers.
4. Alizadeh R 2011 *Safety Science* 49(2) 315-323
5. Zanlungo F Ikeda T and Kanda T 2011 *Europhysics Letters* 93 6
6. Gotoh H Harada E and Andoh E 2012b *Safety Science* 50(2) 326–332
7. Dai J Li X and Liu L 2013 *Physica A: Statistical Mechanics and Its Applications* 392(9) 2202- 2211
8. Mohd Mahudin N D 2012 *Quality of Rail Passenger Experience: The Direct and Spillover Effects of Crowding on Individual Well-Being and Organisational Behaviour* (University of Nottingham: PhD Thesis)
9. Gawroński P and Kułakowski K 2011 *Computer Physics Communications* 182 9
10. Moussaïd M Perozo N Garnier S Helbing D and Theraulaz G 2010 *PLoS ONE* 5(4) e10047
11. Bouchard M Haegele J and Hexmoor H 2015 *Connection Science* 27(2) 164-187
12. Samson B P V Aldanese C R Chan D M C San Pascual J J S and Sido M V A P 2017 *Procedia Computer Science* 108C 195–204
13. Zhou J Chen H Yang J and Yan J 2014 *Mathematical Problems in Engineering* 843096
14. Godoy J Karamouzas I Guy S J and Gini M 2016 *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence* (AAAI Publications)
15. Su H Zhu J Dong Y and Zhang B 2017 *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI)* 1 p 2772–8
16. Bazghandi A 2012 *International Journal of Computer Science Issues* 9(1) 115–119



## РАЗГОВОРНЫЙ ИНТЕРФЕЙС В СРЕДЕ BOT FRAMEWORK

*Есимова Наргиз Бекболатовна, Есендаулетова Жана-Гуль Тлеукуловна  
Карагандинский университет им. Е.А. Букетова, г.Караганда, 100028*

В последние годы мы наблюдаем бурный расцвет искусственного интеллекта, который демонстрирует все новые успехи: лучшее качество распознавания речи и изображений. Такой прогресс обусловлен все большим количеством доступных данных и вычислительных мощностей для их обработки и обучения интеллектуальных моделей. Интеллектуальные и когнитивные технологии считают основным двигателем цифровой трансформации в ближайшие годы, когда многие рутинные операции, выполняемые традиционно человеком, будут заменены ИИ. Если вы планируете заниматься разработкой и созданием продуктов на базе программного обеспечения, способность грамотно включать решения на базе ИИ в свои продукты будет являться важнейшим критерием их успешности.

Последние несколько лет мир информационных технологий движет система интеллектуализации. Мы просыпаемся, используя smart-будильники наряду со smart-часами, которые анализируют фазы сна. Собираясь в институт или на работу, мы смотрим детальные прогнозы погоды. Затем по дороге в очередное место встречи мы пользуемся навигатором, который не просто строит нам качественный маршрут, но и может с нами поговорить. Ежедневно миллионы людей говорят «Привет, Cortana», «Привет, Siri», «О'кей, Google» или разговаривают с Алисой.

Можно просто и быстро сделать свою smart-систему, используя технологии компании Microsoft. Стоит сделать всего один запрос к сервисам Azure, и в считанные минуты, вы напишите код.

Действительно, эти технологии от Microsoft – настоящий прорыв в мире искусственного интеллекта. Всего один веб-запрос – и в ваших руках уже есть что-то smart-подобное! С помощью этих технологий и своими идеями вы быстро построите что-то умное.

Прежде чем начать строить умные приложения с использованием элементов искусственного интеллекта, будет полезно узнать его основы:

- Машинное обучение ,типы машинного обучения;
- Язык - понимание естественного языка, обработка естественного языка (машинный перевод, генерация естественного языка, анализ изменений, распознавание именованных сущностей, извлечение связей);
- Речь - распознавание речи, распознавание голоса, преобразование текста в речь и речи в текст;
- Компьютерное зрение - обнаружение объектов, распознавание образов, распознавание лиц, оптическое распознавание символов;
- Когнитивные службы Microsoft.

### **Создание приложения с разговорным интерфейсом в среде Bot framework.**

Microsoft Bot Framework (<https://dev.botframework.com>) – это набор пакетов SDK с открытым исходным кодом для легкого и быстрого создания программ-роботов корпоративного класса. Посредством нескольких встроенных конструкторов эта платформа ускоряет разработку приложения с разговорным интерфейсом. Это могут быть простые приложения, предназначенные для обработки пользовательских сообщений независимо, или контекстные приложения. Контекстные приложения способны обучаться, анализируя ранее полученные пользовательские сообщения. В результате пользователи будут получать более точные ответы на свои запросы.

Используя инфраструктуру Bot Framework, можно легко отработать логику программы-робота, после чего установить эту программу на одной или нескольких платформах обмена сообщениями. Полнофункциональная среда Bot Framework не

является частью пакета Cognitive Services, но она может эффективно интегрироваться с несколькими API Cognitive Services.

Данная среда благодаря технологиям понимания естественного языка, компьютерного зрения и обработки речи позволяет сделать общение с ботом похожим на общение с человеком.

Бот в контексте инфраструктуры Bot Framework, по существу, представляет собой комбинацию из трех компонентов. Взгляните на рисунок 1, чтобы увидеть связь между этими компонентами.



Рисунок 1 - Схема архитектуры приложения с разговорным интерфейсом на основе инфраструктуры Bot Framework

### Для создания нового проекта приложения Bot Framework:

Запустите программу Visual Studio 2017 и создайте новый проект, выполнив команду меню Файл ► Создать ► Проект. Выбрать в появившемся диалоговом окне шаблон Bot Application Visual C#, как показано на рисунке 2. В результате будет создано приложение ASP.NET MVC с настройками для среды Bot Framework.

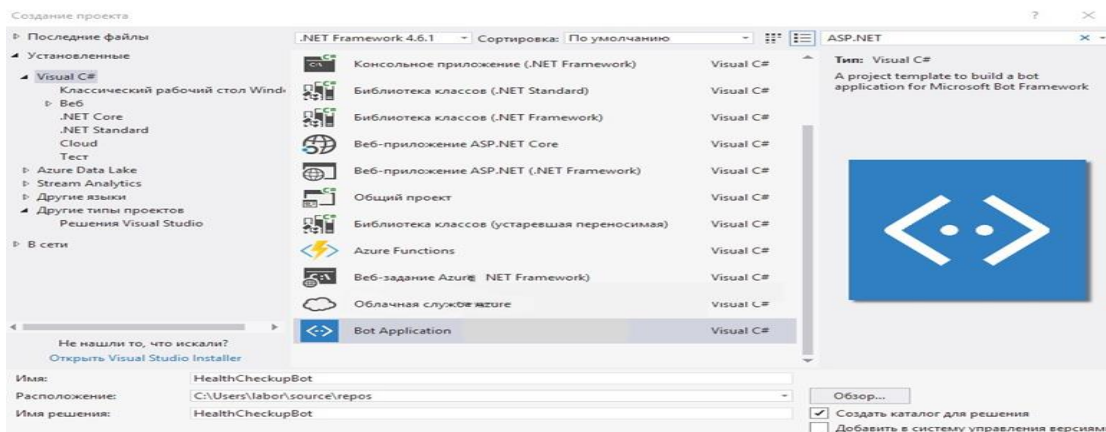


Рисунок 1 - Создание нового приложения Bot Framework в Visual Studio 2017

В приложении Bot Application, которое только что создали, должен быть только один контроллер. Откроем в Обозревателе решений папку Контроллеры и щелкнем мышью на строке MessagesController.cs. В окне Bot Application1 появится исходный код. Начнем его редактирование.

Главным методом здесь является метод Post(). Именно тут содержится вся логика бота. Именно в это место программного кода приходят пользовательские сообщения для обработки, и именно отсюда отправляются соответствующие ответы. В настоящее время бот сконфигурирован на передачу назад того же самого сообщения вместе с длиной этого сообщения в символах, которое бот получает от пользователя.

Модифицируем исходный код, чтобы возвращать приветственные сообщения и сообщения, заданные по умолчанию. При редактировании исходного кода необходимо быть внимательным (см. листинг-1).

### Листинг 1 - Новый исходный код

```
[BotAuthentication]
public class MessagesController : ApiController
{
```

```

        string[] greetingPhrases = new string[] { "привет", "приветствую", "здравствуйте", "как
дела", "доброе утро", "добрый день" };
/// <summary>
/// POST: api/Messages
/// Receive a message from a user and reply to it
/// </summary>
public async Task<HttpResponseMessage> Post([FromBody]Activity activity)
{
if (activity.Type == ActivityTypes.Message)
{
ConnectorClient connector = new ConnectorClient(new Uri(activity.ServiceUrl));
string userText = activity.Text.ToLower();
string replyText = "";
if(this.greetingPhrases.Contains(userText))
{
    replyText = @"Приветствую! Я – бот записи студенческую клинику.
    Буду рад Вам помочь спланировать визит в студенческую клинику.
    Скажите ‘записаться к врачу.’”;
}
else
{
replyText = "К сожалению, я не понял, что Вы сказали.";
}
Activity reply = activity.CreateReply(replyText);
await connector.Conversations.ReplyToActivityAsync(reply);
}
else
{
HandleSystemMessage(activity);
}
var response = Request.CreateResponse(HttpStatusCode.OK);
return response;
}
private Activity HandleSystemMessage(Activity message) {...}
}

```

#### **Построчная разбивка исходного кода следующая:**

public async Task<HttpResponseMessage> Post([FromBody]Activity activity) Этот фрагмент информационного обмена между ботом и пользователем называется действием. Вы получаете объект действия от POST-вызова по протоколу HTTP, направленного на конечную точку обмена сообщениями веб-службы.

Среда Bot Framework поддерживает несколько типов действий, некоторые из которых вы увидите чуть ниже. Прямо сейчас вы заинтересованы в действии с типом Message(Сообщение).

Действие Message представляет текстовое сообщение и некоторую о нем информацию, такую как текстовая строка, отправитель, получатель, метка времени.

```

if (activity.Type == ActivityTypes.Message)
{
ConnectorClient connector = new ConnectorClient(new Uri(activity.ServiceUrl));
string userText = activity.Text.ToLower();

```

Если полученное конечной точкой действие имеет тип Message, то (для простоты сравнения) текст сообщения преобразуется в нижний регистр, и создается экземпляр

соединителя. Этот экземпляр соединителя будет использоваться, чтобы передавать ответ обратно пользователю. Действие `activity.ServiceUrl` вызывает на URL-адрес, куда будет отправляться ответ бота. Каждый канал связи имеет свой собственный URL-адрес обслуживания, где прослушиваются ответы.

```
if(this.greetingPhrases.Contains(userText))
{
    replyText = @"Приветствую! Я – бот записи в студенческую клинику.
    Буду рад Вам помочь спланировать визит в нашу клинику.
    Скажите ‘записаться к врачу.’”;
}
else
{
    replyText = "К сожалению, я не понял, что Вы сказали.";
}
```

Далее проверяем, соответствует ли сообщение пользователя одной из фраз, определенных в строковом массиве `greetPhrases`. Если это так, то назначаем ответу вежливое вводное сообщение. В противном случае отправляем сообщение по умолчанию, чтобы дать пользователю понять, что бот не знает, как ответить на его сообщение.

```
Activity reply = activity.CreateReply(replyText);
await connector.Conversations.ReplyToActivityAsync(reply);
Вы используете метод CreateReply() полученного действия, чтобы создать новое действие для ответного сообщения. Затем отправляем ответ через объект соединителя.
var response = Request.CreateResponse(HttpStatusCode.OK);
return response;
```

После отправки ответа возвращаем из POST-метода веб-службы простейший код состояния HTTP 200, чтобы дать сигнал об успешном завершении этой операции.

### Выполнение, тестирование и отладка бота.

Настало время выполнить приложение, чтобы увидеть, как бот отвечает на пользовательские сообщения. Visual Studio поставляется в комплекте с облегченным программным обеспечением веб-сервера, IIS Express, для выполнения веб-приложений и служб. Веб-сервер IIS Express – это упрощенная версия своего полноценного собрата, IIS.

Чтобы запустить конечную точку обмена сообщениями, откройте меню «Отладка» и выберите элемент «Начать отладку». Или нажать клавишу F5. В результате приложение будет запущено в режиме отладки. По завершении процесса отладки откройте в браузере новую вкладку и введите URL-адрес конечной точки. По умолчанию приложения Bot Framework на порту 3979. Поэтому его URL-адрес будет следующий: `http://localhost:3979`. Здесь должна появиться страница приложения (рис.2).

## HealthCheckupBot

Describe your bot here and your terms of use etc.

Visit [Bot Framework](#) to register your bot. When you register it, remember to set your bot's endpoint to `https://your_bots_hostname/api/messages`

**Рисунок 2** - Страница по умолчанию конечной точки обмена сообщениями.

Если вы это видите, все работает, как надо...

Открытая в браузере страница по умолчанию не дает возможности протестировать бота. Это статическая страница, которая используется только для описания бота. Для тестирования приложения можно применить - эмулятор Bot Framework.

Эмулятор Bot Framework – это бесплатное программное обеспечение, которое специально предназначено для локального или удаленного тестирования приложения Bot Framework. Скачать эмулятор можно со страницы `https://aka.ms/bf-bcemulator`.

Установить полученный эмулятор, запустить его и указать в строке поиска приведенный ниже URL-адрес: <http://localhost:3979/api/messages>  
 Не заполняя поля ввода идентификатора приложения Microsoft App и пароля Password, нажать кнопку Connect(Соединить).  
 Если поместить в исходный код точку останова отладчика, то Visual Studio может выглядеть, как показано на рисунке 3. Можете воспользоваться богатыми возможностями отладки Visual Studio, чтобы просмотреть исходный код. Для этого воспользуемся клавишами F10(Перешагнуть), F11(Шагнуть внутрь) и F5(Продолжить).

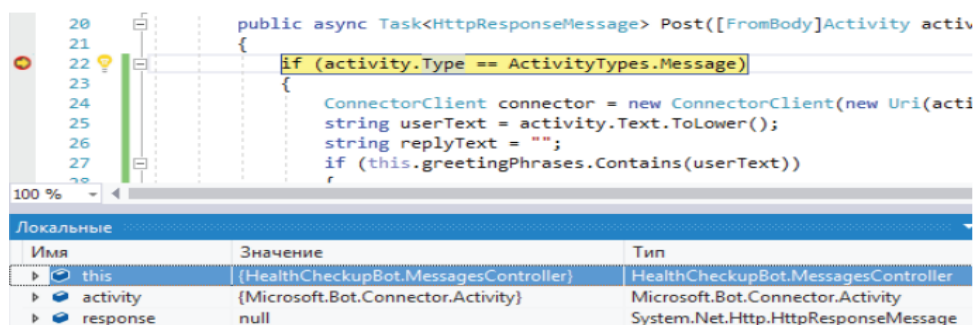


Рисунок 3 - Visual Studio показывает активную точку останова, которая сработала из-за входящего сообщения из эмулятора

В Visual Studio нажать кнопку F5, чтобы закончить отладку, в случае если есть активные точки останова. Теперь снова откроем эмулятор и начнем разговор с ботом. Чтобы начать разговор, необходимо напечатать и отправьте слово «привет». Как видно из рисунке 4, бот ответит на сообщения «привет» и «приветствую», как и ожидалось. И тем не менее он не понял сообщение «привет, бот» и вернул сообщение по умолчанию. Это можно быстро исправить, добавив в класс *MessagesController* следующий ниже метод:

```
private bool MessageHasPhrase(string message, string[] phraseList)
{
    foreach (string phrase in phraseList)
    {
        if (message.Contains(phrase))
        {
            return true;
        }
    }
    return false;
}
```

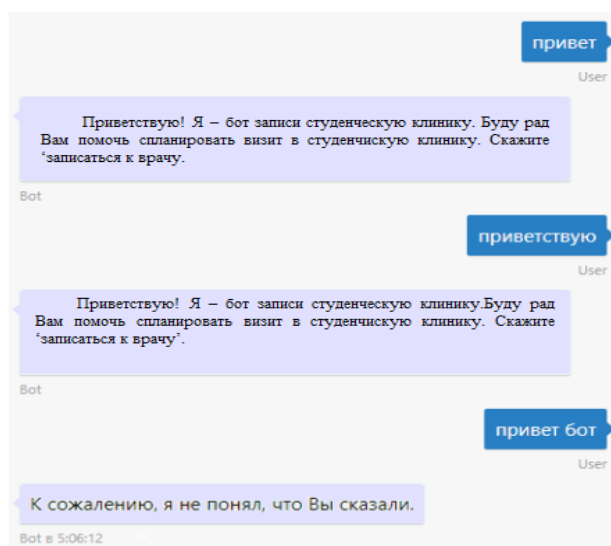


Рисунок 4 - Начало диалога с ботом в эмуляторе

Здесь был создан элементарный бот. Он может правильно выполнять элементарные вещи, но имеет большие резервы для совершенствования. В качестве следующего шага необходимо добавить в бот записи к врачу контекст посредством применения диалогов и подсказок. Кроме того, бот может транслировать свои сообщения на одном или нескольких каналах связи. В существующем сейчас виде бот работает локально на компьютере разработчика, и связываться с ним можно только через эмулятор. Конечная цель, конечно, состоит в том, чтобы сделать бот доступным через такие каналы связи, как Telegram и Facebook Messenger.

#### Список использованной литературы

1. Azure Cognitive Services - <https://azure.microsoft.com/ru-ru/products/cognitive-services/>Развертывание высококачественных моделей ИИ в виде API-интерфейсов (дата обращения: 20.09.2022).
2. TensorFlow - <https://www.tensorflow.org/><https://www.tensorflow.org/>Создание приложений машинного обучения (дата обращения: 20.09.2022).
3. Первые шаги к виртуальной реальности. – URL: [http://www.hwp.ru/Multimedia/Iglasses/indx.html?SHOWALL\\_1=1](http://www.hwp.ru/Multimedia/Iglasses/indx.html?SHOWALL_1=1)<https://www.tensorflow.org/>Первые шаги к виртуальной реальности.
4. Postman - <https://www.postman.com/>(дата обращения: 20.09.2022).
5. Зинченко, Ю.П. Виртуальная реальность в экспериментальной психологии: к вопросу о методологии / Ю.П. Зинченко. – URL: <http://vprosvet.ru/biblioteka/virtualnaya-realnost/> (дата обращения: 20.09.2022).
6. Course Alice. – URL: <https://www2.cs.duke.edu/csed/alice/aliceInSchools/> (дата обращения: 20.09.2022).

## ҒЫЛЫМИ-БІЛІМ БЕРУ ҚЫЗМЕТІН ҚОЛДАЙТЫН ӘКІМШІЛІК ТОПТАМАЛАРҒА ҚОЛЖЕТІМДІЛІГІНІҢ БАСҚАРУ ЖҮЙЕСІ

Серикбаева С.К., Болсынбек М.Қ., Мусабеков А.К.  
Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Қазақстан, Астана  
inf\_8585@mail.ru, [mbolsynbek@bk.ru](mailto:mbolsynbek@bk.ru), [mussabekov.arman@mail.ru](mailto:mussabekov.arman@mail.ru)

Қолжетімділігін басқару жүйесі (ҚБЖ) ұйымға пайдаланушылардың корпоративтік құжаттарға қол жеткізу құқықтарын басқаруды іске асыру үшін енгізіледі.

ҚБЖ келесі міндеттерді шешуге мүмкіндік береді:

- қызметкерлердің есептік жазбаларымен жұмысты ретке келтіру;
- автоматты түрде кадрлық өзгерістерді синхрондауды орындау;
- пайдаланушыға ақпараттық жүйеге қолжетімділік құқығын автоматты түрде беру;
- құқық аудитін жеңілдету;
- пайдаланылмаған есептік жазбаларды автоматты түрде анықтауға мүмкіндік беру;
- деректерді рұқсатсыз пайдалану жағдайларының алдын алу.

Ресурстарға қолжетімділікті басқарудың әзірленіп жатқан моделінде қолжетімділікті басқарудың мандаттық сатылық үлгісі қолданылады. Онда әрбір субъект жүйеде құқықтар (ережелер) жиынтығына ие. Әрбір құқық (ереже) өз кезегінде белгілі бір ресурстарға қолжетімділікті қамтамасыз ететін артықшылықтар жиынтығын қамтиды. Қолжетімділікті басқарудың сатылық үлгісінің бірліктерін қолдана отырып, бұл қатысушыларды басқаруды орталықсыздандыруды қамтамасыз етеді, яғни жүйеде әртүрлі деңгейдегі көптеген әкімшілердің (пайдаланушыларға құқықтар (ережелер) жиынтығын ұсынатын қатысушылар) болады. Жүйе әкімшісі өз өкілеттіктерінің бір бөлігін анағұрлым төменгі деңгейдегі әкімшілерге бере алады. Әзірленген жүйе субъектілердің үш негізгі түрін қолдайды: қатысушылар, қатысушылар тобы және басқалары [1].

Қатысушы тұлға – бұл жүйеге қатысы бар субъектілер. Осы немесе басқа түрдегі қатысушылардың көпшілігі цифрлық кітапхананың ақпараттық ресурстарын пайдаланушылар болып табылады.

Қатысушылар тобы – бұл ақпараттық ресурстармен өзара әрекеттесу кезінде кейбір ұқсас қызметтері бар көптеген субъектілер. Қалғандарының барлығы тек қоғамдық топтамаларға қол жеткізе алатын (ақпаратты іздеу және қарау) тіркелген пайдаланушылар. Субъектілерді басқаруда келесі ұстанымдар қолданылады:

1. Әрбір субъектіге (қатысушы тұлғаға) олардың нысандарға (топтамаларға, құжаттарға) тиісті қолжетімділік орнын көрсететін сыныптау деңгейлері (деңгейлер сатысы) салыстырылады.

2. Әрбір субъект жүйеде ережелер жиынтығына ие, бұл оның әзірленуші жүйедегі мәртебесіне байланысты. Құқықтардың әрқайсысы өз кезегінде белгілі бір ресурстарға (нысандарға) қолжетімділікті қамтамасыз ететін артықшылықтар жиынтығын қамтиды.

3. Кейбір топқа жататын субъектілер осы топтың көптеген құқықтарын, яғни субъектілер тобына берілген құқықтарды иеленеді.

4. Қолжетімділікті басқарудың сатылық үлгісінің бірліктерін пайдалану. Шындығында, бұл ұстанымды қолдану пайдаланушыларды басқаруды орталықсыздандыруды білдіреді, яғни жүйеде әртүрлі деңгейдегі көптеген әкімшілердің (қатысушылар, пайдаланушыларға құқық беретін) болуы. Жүйе әкімшісі өз өкілеттіктерінің бір бөлігін төменгі деңгейдегі әкімшілерге бере алады. Бұл амал, мысалы, төменгі деңгейдегі субъектілердің құқықтарын басқару жауапкершілігін жоғары деңгейге ауыстыруға мүмкіндік береді, осылайша жүйенің қолдау қызметінен жүктемені алып тастайды.

5. Субъектілер құқықтарының едәуір бөлігі оның мәртебесімен немесе субъектінің белгілі бір топқа жатуымен нақты белгіленеді.

6. Қолжетімділік құқығын орнату ресурстардың әрқайсысының баптауларынан бөлек. Басқаша сөзбен айтқанда, әкімшіге қолжетімділік құқығын өңдеу интерфейсін беретін жүйенің компоненті ресурстардың жұмыс істеу логикасына, оның ішінде ресурстар жиынтығының өзіне де байланысты емес. Сол сияқты, ресурстардың өзі қол жеткізуді басқару жүйесінің логикасына тікелей байланысты емес, оған жүгіну ағымдағы пайдаланушының белгілі бір артықшылыққа ие болуын растау үшін мерзімді сұраныстарға ғана азаяды. Сол сияқты, ресурстардың өзі қолжетімділікті басқару жүйесінің логикасына тікелей байланысты емес, оған жүгіну ағымдағы пайдаланушының белгілі бір артықшылыққа ие болуын растау үшін мерзімді сұраныстарға ғана саяды [2].

Цифрлық кітапхананы басқару жүйесінің жұмысын сипаттау үшін негізгі болмыстарды жазу керек: *қызметтік (жүйелік) топтамалар*, көруге, өңдеуге және жабуға арналған *топтамалар*, метадеректерде бар элементтерді келесі сыныптарға бөлінеді: Р – қоғамдық (DC метадеректері), А – автоматтық (иесі мен жазу күні, модификациясы және т.б.), S – қызметтік (қолжетімділік құқығы, метадеректер элементінің сыныбы, элементке қолжетімділік және т.б.), М – мета (беттермен (сөздікпен) байланыс), U - бірегей (ресурс сәйкестендіргіші) және әр сынып өз кезегінде элементке қолжетімділігі бар (Р – барлығы, S - супер әкімші (0 - деңгей), А -әкімшілер (1, 2-деңгейлі) және *құжаттар*.

Топтаманың құрылымы мен мазмұнын өңдеуге арналған әкімші деңгейінің артықшылықтары топтаманың атрибуттарына байланысты. Егер топтамаға қолжетімділік рұқсат етілсе, онда барлығы дерлік топтаманы іздей және көре алады, олай болмаған жағдайда тек осы топтаманың әкімшісі, супер-әкімші және 1-ші деңгейдегі әкімші ғана қарай алады [3].

Әкімшінің әр деңгейі топтаманы қарау және өңдеу үшін қандай құқықтар берілгенін және оның қандай артықшылықтары бар екенін анықтайды, яғни:

- топтаманы (құжаттарды) көруге рұқсат;
- топтаманы өңдеуге, нысандарды қосуға, өңдеуге немесе жоюға рұқсат;
- топтама құжаттарын өңдеуге рұқсат (құжатты жасау, өзгерту және жою);
- топтама нысандарының орындау тәртібін өзгертуге рұқсат;
- топтама нысандарын беру форматын өзгертуге рұқсат;
- жаңа нысанды қосуға рұқсат (метадеректер элементі) және төменгі деңгейдегі

әкімшілерді тіркеуге рұқсат.

Топтаманың құрылымы мен мазмұнын өңдеуге арналған әкімші деңгейінің артықшылықтары топтаманың атрибуттарына байланысты.

Жүйеде әкімшіліктендірудің сатылық алты деңгейлі жүйесі (басқару) іске асырылды (1-кесте).

Кесте 1 – Әкімшіліктендіруді басқару

Әкімшіліктендірудің деңгейі	Қолжетімділігі бар	Жұмыс атқарады (өңдейді)
0	Барлық топтамалар, құжаттар, схемалар, сөздіктер және т.б.	Жүйе және құжаттардың барлық топтамаларымен
1	Қызметтіктен басқа барлық топтамалар мен құжаттар	Барлық топтамалар мен құжаттар (қызметтік пен жүйелік топтамалардан басқа)
2	Тек өзінің және ашық топтамалар	Құрылым (топтаманың метасипаттамасы), құжаттар (шаблондар, қасиеттер) топтама жазбалары



3	Ашық топтамалар	Құжаттардың құрылымы мен жазбалары бар ашық топтамалар (топтаманың метасипаттамасы)
4	Ашық топтамалар	Тек ашық топтамалар жазбалары бар (сондай-ақ жеке жағдайда жеке топтамалар мен құжаттардың құрылымымен жұмыс істей алады)
5	Ашық топтамалар	Ашық топтамалар құжаттарымен (тек өзінің жазбаларымен)

Әкімші әртүрлі артықшылықтары бар түрлі топтаманы қарау және өңдеу үшін және топтама жазбаларын (құжаттарды) өңдеуге қол жеткізе алады. Әр топтамада (түріне байланысты) метадеректердің ең төменгі міндетті жиынтығы бар. Басқару деңгейіне байланысты топтама әкімшісі негізгі каталогтағы қолжетімді метадеректер негізінде топтама метадеректерінің схемасын анықтай алатын мүмкіндікке ие [4].

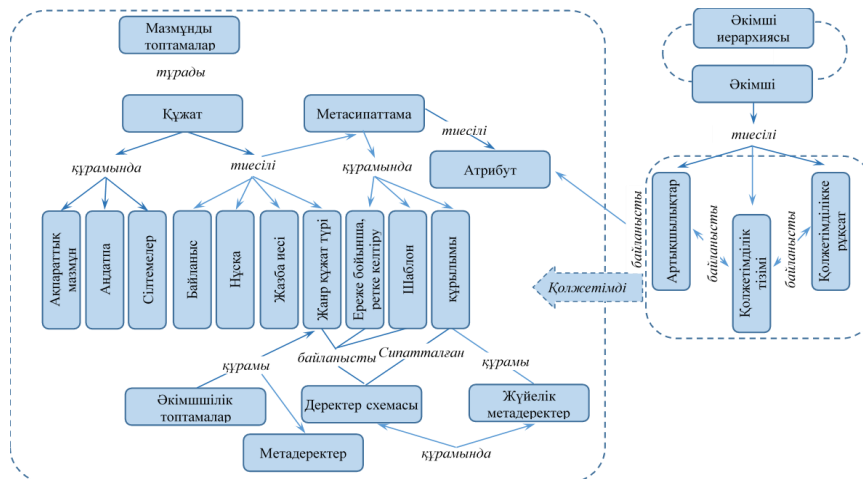
Қолжетімділік құқықтары (артықшылықтары) моделі келесі қызметтерді орындауға мүмкіндік береді:

- жүйе әкімшілерінің есептік жазбаларын құру;
- бірдей әдепкі артықшылықтары бар жүйе әкімшілерінің топтарын құру;
- әкімшілер топтары үшін де, жүйенің жеке әкімшілері үшін де топтамалар мен құжаттарға қолжетімділік құқығын тағайындау;
- топтар үшін де, жеке пайдаланушылар үшін де әртүрлі типтегі құжаттарға қолжетімділік құқығының шаблондарын қалыптастыру;
- жүйенің әкімшісі үшін топтама мен құжат құрылымдарына қолжетімділік құқығын тағайындау.

Әкімшілерді келесі құрылымдар ретінде танылатын әкімшілер тобына біріктіруге болады:

- топтың атауы;
- әкімшілер тізімі;
- нақты топтамалар мен құжаттарға құқықтар жиынтығы (артықшылықтар);
- құжат түрлеріне құқықтар жиынтығы.

Әкімшінің қолжетімділігі басқару деңгейіне (әкімшілер сатысына) және топтама атрибуттарына байланысты. Жоғарыда айтылғандай, топтама жүйесінде шартты түрде *Әкімшілік және Мазмұндық* топтамаларға бөлуге болады. Әрбір *Мазмұндық* топтама құжаттардан тұрады. Құжат, өз кезегінде, Ақпараттық мазмұнды, Аннотацияны және сілтемеден (яғни, толық мәтін сақталатын деректер қоймасына (репозиторий), мұрағаттық файлға, жеке фрагментке (сурет, кесте және т.б.) тұрады. Сондай-ақ, құжатта Метасипаттама, Құжаттың жанрлық түрі, Нұсқалылығы және Жазба иесі бар. Коллекция құжаттарының Метасипаттамасында өз кезегінде Шаблон, Құрылым, Құжатты ретке келтіру және көрсету ережелері бар (1-сурет).



Сурет 1 – Әкімшілердің топтамаларға қол жеткізуді басқару жүйесі

Жүйеде әкімшілендірудің алты сатылық деңгейлі жүйесі (басқару) иерархиялық топтамалары іске асырылды (2-сурет). Сондай-ақ, әкімшілердің әрқайсысы ашық топтамаларды қарауға, басқа төменгі деңгейдегі әкімшілерді тіркеуге, төменгі деңгейдегі әкімшілердің барлық құқықтары мен артықшылықтарына ие.



Сурет 2 – Әкімшілендірудің иерархиялық топтамалары

Қолжетімділікті басқарудың бұл әдісі жүйедегі мәртебе негізінде пайдаланушылардың ақпаратқа қолжетімділігін бақылауға мүмкіндік береді. Бұл әдісті қолдану жүйеде құқықтарды (ережелерді) анықтауды білдіреді. Заң (ереже) ұғымын белгілі бір қызмет түрімен байланысты әрекеттер мен өкілеттіктер жиынтығы ретінде анықтауға болады. Осылайша, әр пайдаланушы үшін әр нысанға қолжетімділігінің барлық түрлерін көрсетудің орнына, нысандарға қолжетімділік ережелерінің жиынтығын көрсету жеткілікті. Ал пайдаланушыларға өз кезегінде олардың артықшылықтарын көрсету.

## Әдебиеттер тізімі

- 1 Барахнин В.Б., Федотов А.М. Исследование информационных потребностей научного сообщества для построения информационной модели описания его деятельности // Вестник Новосибирского государственного университета. – 2008. – Т. 6, №3. – С. 48-59.
- 2 Федотов А.М., Барахнин В.Б., Жижимов О.Л. и др. Модель информационной системы для поддержки научно-педагогической деятельности // Вестник Новосиб. гос. ун-та. – 2014. – Т. 12, вып. 1. – С. 89-101.

3 Шокин Ю.И., Федотов А.М., Жижимов О.Л. Технологии создания распределенных информационных систем для поддержки научных исследований // Вычислительные технологии. – 2015. – Т. 20, №5. – С. 251-274

4 Kurmanbekovna S.S., Gabitovich B.A., Aralbaevna S.M. et al. Development of technology to support large information storage and organization of reduced user access to this information // International Journal of Advanced Computer Science and Applications. – 2021. – Vol. 12, Issue 7. – P. 493-503.

## SIEM – ЖҮЙЕСІ НЕГІЗІНДЕ КОМПЬЮТЕРЛІК ЖҮЙЕНІҢ АҚПАРАТТЫҚ РЕСУРСТАРЫН ТИІМДІ ҚОРҒАУ ӘДІСТЕРІ

Амирбай А.А., Муханова А.А.

*Л.Н. Гумилев аындағы Еуразия ұлттық университеті, Астана, Қазақстан*

*E-mail: [aizat.amirbay@mail.ru](mailto:aizat.amirbay@mail.ru)*

**Аннотация.** Кәсіпорынның компьютерлік жүйесінде ақпараттық ресурстарға қарсы шабуылдаған қатерлердің жүзеге асу мүмкіндігі орын алғанда, шығындану тәуекел деңгейі өсіп, кәсіпорынның ақпараттық қауіпсіздігінің бұзылуына, шығындануға алып келеді. Мақала - қауіпсіздік мониторинг жүйесін құрумен ақпараттық ресурстардың қауіпсіздігін қамтамасыз ету туралы өзекті тақырыпқа арналған. Сондай-ақ кәсіпорынның компьютерлік жүйедегі ақпараттық ресурстардың қатерлі тәуекел деңгейін эксперттік әдіс бойынша бағалау SIEM-жүйесінің жүйелік қосымшаларына сай келетіндей пограммалық орталарда есептеу нәтижелері көрсетілген.

**Түйін сөздер:** қатер, тәуекелділік, онтология, осалдық

**Кіріспе.** Компьютерлік жүйеде ақпараттық ресурстарды қорғауды тиімді жүзеге асыру үшін, қауіпсіздік қауіп-қатер мониторинг жүйесін құрумен байланысқа негізделген бүтін сан қатар есебін шығару қажет. Мониторинг жүйесі ақпаратты қорғауда ағымдық амалын жүзеге асырады және құрудағы негізгі мақсаты компьютерлік жүйеде ақпараттық ресурстарға әсер еткен қатерлерден туындайтын қатерлі тәуекелдер деңгейін төмендету және пайда болған шығындарды азайту.

Қауіп-қатер қауіпсіздік мониторинг жүйесін құрудағы ең маңызды және перспективті бағыттың бірі қауіпсіздік оқиға және ақпаратты басқаруды қамтамасыз ететін жүйе SIEM (Security Information and Event Management) болып табылады. SIEM – жүйесінің орталық компоненті – деректер қоймасы [1]. Деректер қоймасында SIEM – жүйесінің аналитикалық модулінен келіп түсетін сұраныстарды өңдеу мен деректерді сақтау іске асады. SIEM – жүйесі ақпараттық қауіпсіздік деңгейінің көтерілуіне әсер етеді, сонымен қоса қауіпсіздік оқиға мен программалық қақтығыстар қауіп-қатерге ұшырамай тұрып басқаруды жүзеге асырады. Жұмыстың мақсаты компьютерлік жүйеде ақпараттық ресурстарды нақты уақыт режимінде қорғауды болжау және оңтайландыру мәселелерін шешу үшін жеке SIEM – жүйесінің тораптарын толтыру туралы кейбір пікірлерді ұсыну.

Ақпараттық қауіпсіздік ақпараттың иелеріне немесе пайдаланушыларға зиян келтіруі мүмкін бір немесе бірнеше ақпараттық жүйе критерийлерінің (құпиялылық, қол жетімділік, тұтастық) бұзылуына байланысты кездейсоқ немесе қасақана ақпараттық ресурстарды және қолдаушы инфрақұрылымды қорғау жағдайы ретінде түсініледі. Компьютерлік жүйеде ақпараттық ресурстардың бұзылуы сөзсіз шығындарға әкеп соғады: қаржылық, операциялық, тұтынушылық зиян, қызметкерлерге келтірілген зиян. Ең маңызды қауіпсіздік көрсеткіштерінің бірі - қатер.

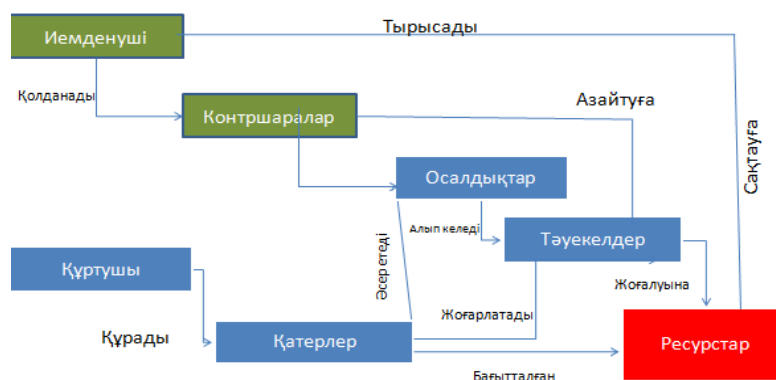
Кез-келген саладағы кәсіпорындардың қызметі белгілі бір дәрежеде тәуекелді болып табылады, бұл адам факторлары мен компания жұмыс істейтін саланың сипатына байланысты [2]. Ақпараттық шабуылдар кәсіпорындардың ішкі және сыртқы бұзушылары және осалдықтарды пайдаланып бағдарламалық қамтамасыз ету бойынша жүзеге асырылады. *Ақпараттық жүйенің осалдығы - кемшіліктер болып табылады, көбіне жүйені жобалау немесе кодтау кезінде жасалған қате болып табылады.* Шабуылдар арқылы жүзеге асырылатын осалдықтар мен қатерлер арасындағы қарым-қатынас тәуекелдің пайда болуы алып келеді, төменде процес 1 - шы суретте көрсетілген.



Сурет 1- Қауіп-қатерлердің пайда болуы мен іске асыру процесі

Ақпараттық технологиялар және ақпараттық қызметтер үшін дамып келе жатқан нарықтар маңызды назар аударуды және зерттеуді талап етеді, себебі құнды және маңызды ақпаратты иелену, пайдалану және беру есебінен компанияға елеулі зиян келтіруі мүмкін бірқатар тәуекелдер болуы мүмкін. Мұндағы, *тәуекел дегеніміз - кәсіпорынның кейбір ресурстарын жоғалту ықтималдығы, кіріс тапшылығы немесе бизнесті жүргізу барысында қосымша шығындардың пайда болуы.*

Құқық бұзушылардың шабуылынан ықтимал шығындарды бағалауға және оны азайту жөніндегі шараларды таңдауға арналған тәуекелге негізделген тәсіл ақпараттық қауіпсіздікте кеңінен қолданылып, тәуекелдерді басқару деп аталды[3]. Тәуекелдерді басқару тұжырымдамасын әртүрлі көздерде әр түрлі болады. Тәуекелдерді басқарудың негізгі кезеңі тәуекелдерді басқару әдістері мен құралдарын таңдау кезеңі болып табылады. Төменде 2-суретке сәйкес, модель мынадай түрде құрастырылады: бөлінген ресурстарға олардың тиісті қаржылық шығындары бойынша да, сондай-ақ ұйымның беделіне залал келтіру, оның қызметін бұзу, құпия ақпаратты жариялаудан материалдық емес зиян. Содан кейін ресурстардың өзара байланысы сипатталған, қауіпсіздік қатерлері анықталады және олардың орындалу ықтималдығы бағаланады.



Сурет 2 - Ақпараттық тәуекелдерді бағалау алгоритмі

Ақпараттық тәуекелдерді талдау - бұл сандық (ақша ресурстарының түрінде) және сапалы (тәуекел деңгейі: жоғары, орташа, төмен) тәуекелдік индикаторларын анықтау арқылы ақпараттық жүйені қорғау деңгейін бағалау процесі. Талдау ақпаратты қорғау процестерін қалыптастырудың әр түрлі құралдары мен әдістерін қолдану арқылы жүзеге асырылады. Тәуекелдерді талдаудың негізі – тәуекелдерді сәйкестендіру процесінде жиналған SIEM-жүйесінің репозиторийінен алынған статистикалық деректер және жүйе талдаушысының жұмысының нәтижелері SIEM жүйесіндегі тиісті түйіндерден немесе процедуралармен пайдаланылады, олар шешім қабылдайды[4]. Қазіргі кезде тәуекелдерді талдау және бағалау үлгілері тәуекелді басқаруды белгісіздік кезінде шешім қабылдау ретінде және даму барысында альтернативті немесе қосымша шешімдер қабылдау критерийлері ретінде сандық тәуекел көрсеткіштерін қарастыратын даму сатысында.

Тәуекелдерді бағалау – тәуекелдерді басқарудың жалпы жүйесінің маңызды құрамдас бөлігі. Бұл тәуекел дәрежесін сандық немесе сапалы түрде анықтау үдерісі. Сапалы талдаудың негізгі міндеті, тәуекелдердің мүмкін түрлерін анықтаудан басқа,

тәуекелдің осы түріне әсер ететін себептер мен факторларды анықтау және сипаттау болып табылады. Сараптамалық бағалау әдісі - белгілі бір мәселе бойынша сараптамалық қорытынды алуға бағытталған логикалық және математикалық процедуралар кешені. Бұл әдістің артықшылығы, басқару шешімдерін оңтайлы шешуге арналған құзыретті маманның тәжірибесін қолдана алады, тәуекелдер сипаттамалары сарапшы құралдармен белгіленуі мүмкін.

SIEM – жүйесін құру және жұмыс істеуін бақылаудағы негізгі мақсат ақпараттық–телекоммуникациялық инфрақұрылымда ақпараттық қауіпсіздіктің деңгейін айтарлықтай көтеру, қауіпсіздік туралы ақпаратты манипуляциялау және қауіпсіз оқиға мен программалық қақтығыстарды “проактивті” басқаруын жүзеге асыру. Мұндағы “проактивті” дегеніміз “қауіп-қатерге ұшырамай тұрып әрекет жасау” деген мағылнаны береді[3-4].

SIEM – жүйесі ақпараттық инфрақұрылысты сақтай алатындай «қосымша сервері» мен «деректер қоймасы» деп аталатын «агенттер» құрылысы бар. *Агенттер* деректердің алғашқы өңдеуі мен сүзбесін, қауіпсіз оқиғалар жинағын орындайды. Жиналған және сүзілген деректерді сақтау үшін деректер қоймасына өтеді, сосын келесі қадамды орындау мақсатында ішкі форматында сақталады. Қосымша сервері болса ақпарат сақтаудың негізгі қызметін атқарады. Ол қоймада сақталған деректерге саралау жөмыстарын жүргізіп, ақпаратты қорғау ескертулерін шығару үшін өндіріске түрлендіреді.

Бірінші деңгейде *деректерді жинау* әрқилы дереккөздердің түрлері арқасында жүзеге асады. Әрқилы дереккөздердің форматы түрліше типте болады. Олардың қатарына: файлдық серверлер, деректер қорының серверлері, Windows – серверлері, желіаралық экрандар, жұмыс станциялары, шабуылдарға қарсы жүйелер (IPS, intrusion prevention systems), вирусқа қарсы программалар және т.б.

Екінші деңгейде қоймада сақталатын, қауіпсіз оқиғалар туралы басқару жүзеге асады. Қоймада сақталынған деректер *деректерді саралау* деңгейінде сұралатын сұраныс нәтижесінде алынады [5]. Бұл деңгейде тәуекелдерді сандық және сапалық бағалау сияқты көптеген операциялар жүзеге асады.

Үшінші деңгейде алынатын, SIEM – жүйесіндегі деректерді саралау деңгейі блып табылады. Өңдеу нәтижесі болып алдын-ала сипатталатын ерікті түрдегі отчеттар, оперативті оқиға туралы (on-line) деректер корреляциясы, сонымен қоса on-line режимде өндірілетін ескертулер жатады.

#### **Тәуекелдерді есептеудің экспертті бағалау**

Қолайсыз оқиғаларды іске асырудан ақпараттық жүйелерде әдіснамалық зиянды бағалау тәуекелдерді бағалауға байланысты. Тәуекелдер бойынша есептеулер ақпараттық жүйелерде пайда болатын жағымсыз оқиғалардың ықтималдығы негізінде есептеледі. Қолайсыз оқиғалардың пайда болуының объективті және субъективті ықтималдығы бар. Қолайсыз оқиғалардың пайда болуының объективті ықтималдығын бағалау тәуекелдерді бағалауды есептеудің алгоритміндегі маңызды міндеттердің бірі болып табылады. Қатерлер орын алғанда, қатерлердің жүзеге асу деңгейі артып, тәуекел деңгейі өседі де ықтимал шығындарды есептеуде эксперттік бағалау негізінде тәуекелдерді сандық және сапалық әдіспен есептейміз. Тәуекелдерді бағалаудың негізі - тәуекелдерді өңдеу жүйелігінде жиналған, SIEM-жүйесінің қоймасынан алынған статистикалық деректер пайдаланылады.

SIEM-жүйесінің қоймасында жиналған қатерлер тізімі, ақпараттық ресурстарды елеулі түрде бұзатын келеңсіз оқиғалардың кейбір маңызды жиындары бар. Бұл ішкі жиын келесідей белгіленеді[6-7]:

$$K = \{K_{i_1}, K_{i_2}, \dots, K_{i_m}\}$$

К жиынтығын құрғаннан кейін, компьютерлік жүйенің жүйелік тиімділігінің төмендеуіне әкелетін барлық қатерлердің сандық көрсеткіштеріне негізделген ішкі жиындар элементтерінің қасиеттерін талдауды жалғастырамыз.  $\Delta T$  кезінде келеңсіз

оқиғалардың  $i$ -ші салдарынан туындаған қатердің математикалық күтуі (мысалы, 1 ай) келесі формула арқылы ұсынылуы мүмкін:

$$v(K_i, \Delta T) = M[v(K_i) * f_i], \quad i = 1 \dots m$$

Мұнда,  $v(K_i)$  – қатерлердің жүзеге асу мүмкіндігі;

$f_i$  –  $\Delta T$  кезінде  $i$ -ші қатерлердің кездейсоқ мәні;

$m$  – барлық қатерлердің жалпы саны.

Егер қолайсыз оқиғалар әрбір қатерлердің зақымдануынан тәуелсіз болса, онда

$$v(K_i, \Delta T) = M[v(K_i)] * M[f_i], \quad i = 1 \dots m$$

және көптеген қатерлерден келтірілген шығын осы арақатынасымен анықталады:

$$V(K_i, \Delta T) = \sum_{i=1}^m M[v(K_i) * f_i],$$

Тәуекелдерді бағалау - кәсіпорынның немесе SIEM жүйесінің қоймасындағы, ақпараттық ресурстарға қарсы қатерлі шабуылдардың шығындану мүмкіндігін бағалау үшін жинақталған, талданған, сақталған, өңделген статистикалық деректерге негізделген.

### **Қорытынды**

Жұмысты орындау нәтижесі бойынша кәсіпорынның компьютерлік жүйедегі ақпараттық ресурстарға шабуылдаған қатерлердің жүзеге асу мүмкіндігі орын алғанда пайда болған тәуекел деңгейін эксперттік әдіс бойынша бағалау және шығының есептеу базасы жүйеде құрылып, SIEM-жүйесінің жүйелік қосымшаларына сай келетіндей пограммалық орталарда есептеу нәтижелері көрсетілді.

SIEM-жүйесінің жүйелік қосымшаларының қызметі негізінде жүргізілген есептеулердегі ақпараттық ресурстардың қатерлері мен осалдықтары анықталып, өңделді. Ақпараттық ресурстардың қатерлері мен осалдықтарын саралау, бір түрге жинақтауды визуалды түрде бейнелеп көрсеттік[8]. Жиналып, өңделген және саралау жұмыстары жүргізілген қатерлер жүйесінде сақталып, жоғарыда аталған есептердің толық кешенді алгоритмі орындалып, нәтижесі алынды. Тәуекелдерді бағалау негізінен статистикалық деректерге негізделеді: кәсіпорынның ақпараттық жүйесінде немесе SIEM - жүйесінің деректер қоймасындағы ақпараттық ресурстарға қатысты бұзушылықтардың шабуылы мен пайдаланушылардың қателіктерінен туындаған шығынды бағалау үшін жазылады, жинақталады, талданып, сақталады және шығынды төмендету шаралары таңдалынады.

Атап айтқанда, қолайсыз оқиғалардың туындау мүмкіндігінде объективті ықтималдылықты бағалау, кәсіпорын ақпараттық жүйелеріндегі ақпараттық ресурстардың қауіпсіздігінің бұзылуынан келтірілген зиянның объективті құнын бағалау және зиянды болжау туралы бағалауды алуға мүмкіндік беретін әдіснамалар ұсынылды.

### **Әдебиеттер**

1 Котенко И.В., Степашкин М.В. Метрики безопасности для оценки уровня защищенности компьютерных сетей на основе построения графов атак // Защита информации. Инсайд, - 2009.- № 3. - С.36–45.

2 A. Fedotov. A. Muhanova. Vulnerability Classification of Information Security in Corporate Systems. International Journal of INFORMATION (Indexed by Scopus, JDream, Mathematical Reviews, Zentralblatt MATH, ProQuest, Swets, EBSCO). – 2014. - Vol.17, No.1.- pp.219-228.

3 Дойникова Е.В., Котенко И.В. Расширение методики оценки информационных рисков за счет использования графов зависимостей сервисов // Санкт-Петербург. Издательство Политехнического университета. - 2011. - С.71-72

4 Федотов. А.М., Ревнивых А.В.А.А. Муханова. Классификация угроз и уязвимостей информационной безопасности в корпоративных системах. Вестник. Серия информационных технологий.– Новосибирск: НГУ, - 2013. - Т.11. В. 2. – С.55-72.

5 Муханова А.А., Амирбай А.А. Расчет рисков на основе объективных оценок для уязвимости информации в компьютерной системе предприятия // XVII Международная научно-практическая конференция «Российская наука в современном мире».-Москва: МГУ, 2018. – С.140-142

6 Miller D.R., Harris Sh., Harper A.A., VanDyke S., Black Ch. Security Information and Event Management (SIEM) Implementation. McGraw–Hill Companies. 2011. - 430 p.

7 Дервянко А.В. Построение эмпирических моделей для управления сложными Чечулин А.А., Котенко И.В. Анализ происходящих в реальной сети событий на основе использования системы моделирования сетевых атак // VII Санкт-Петербургская межрегиональная конференция «Информационная безопасность регионов России (ИБРР–2011). 26–28 октября 2011 г. - С.97–98.

8 Абденов А.Ж., Абденова Г.А., Амирбай А.А., Кулбаев Д.Р. Маркетинговые информационные услуги в SIEM-системах // Вестник ЕНУ. Серия естественно-технических наук. - Астана: ЕНУ, 2017. -№4(119). – С.24-35.



# АВТОМАТИЗАЦИЯ ИНТЕРПРЕТАЦИИ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ МЕДИЦИНСКИХ ЛАБОРАТОРИИ

<sup>1</sup>К. Кадиркулов

<sup>2</sup>А. Исмаилова, доктор PhD, ассоциированный профессор  
<sup>1</sup>ТОО «Smart Soft Kazakhstan», г. Астана, kkuanysh@gmail.com  
<sup>2</sup>КАТУ им. С. Сейфуллина, г. Астана

## Введение

Государственная программа «Цифровой Казахстан» реализуемая в период с 2018 по 2022 гг. нацелена на повышение уровня жизни населения страны за счет использования цифровых технологий [1]. В рамках программы производится реализация электронного паспорта здоровья населения РК, где автоматизация медицинских лабораторий занимает приоритетную роль. И как показывает практика, не все лаборатории имеют автоматизацию и централизованное хранение данных, то есть каждый результат лабораторного исследования остается на локальных хранилищах анализатора. Соответственно, результаты анализов выдаются на стандартных бланках анализатора (зачастую на английском языке) или же результаты переписываются на бумажный носитель, что может привести к ошибкам человеческого фактора, разночтения показателей и потери данных. В рамках программы «Цифровой Казахстан», а именно в блоке лабораторных исследований электронного паспорта здоровья населения необходимо произвести комплексную автоматизацию операционных процессов лаборатории с внедрением лабораторной информационной системы (далее - ЛИС), с помощью которой производится автоматизация процессов от регистрации до выдачи результатов.

Президент Республики Казахстан Касым-Жомарт Токаев 4 марта 2020 года на совещании по реализации Государственной программы «Цифровой Казахстан», отметил, что цифровизация направлена не для развития одного сектора, а всей экономики государства и в преобразовании общества в целом. Президент подчеркнул, что в Республике Казахстан имеется зависимость от зарубежных разработок и технологий, в связи с чем, было поручено на законодательном уровне поддержать IT-компании казахстанского производства и обеспечить им приоритетность в конкуренции в государственном сектор [2].

## Основная часть

Целью представленной работы является процесс автоматизации интерпретации результатов лабораторных исследований. Система современного здравоохранения ставит перед клинической лабораторной диагностикой задачу быстрого и качественного определения результатов. Решение такой актуальной проблемы лежит в автоматизации процессов на всех этапах диагностических исследований, что позволит сократить время получения результатов, минимизировать участие человека и получить достоверные результаты.

Основной целью интерпретации результатов является первичное диагностирование и трактовка показателей на основе ререференсных значений каждого исследования, где отклонение от референсных значений будет считаться наличием проблем в организме пациента (рис. 1).

В работе описывается полный цикл автоматизации процессов от регистрации биоматериала до выдачи результатов лабораторных исследований с протоколом интерпретации результатов. Автоматизация производится путем внедрения лабораторной информационной системы [3]. Лабораторные информационные системы (ЛИС) обеспечивают критически важные возможности для удовлетворения широкого спектра потребностей современных лабораторий в обработке информации. Архитектуры ЛИС включают конфигурации мэйнфрейма, клиент-сервера и тонкого клиента. Программное

обеспечение базы данных управляет данными лаборатории. Словари ЛИС – это таблицы базы данных, которые лаборатория использует для адаптации ЛИС к уникальным потребностям лаборатории [4].

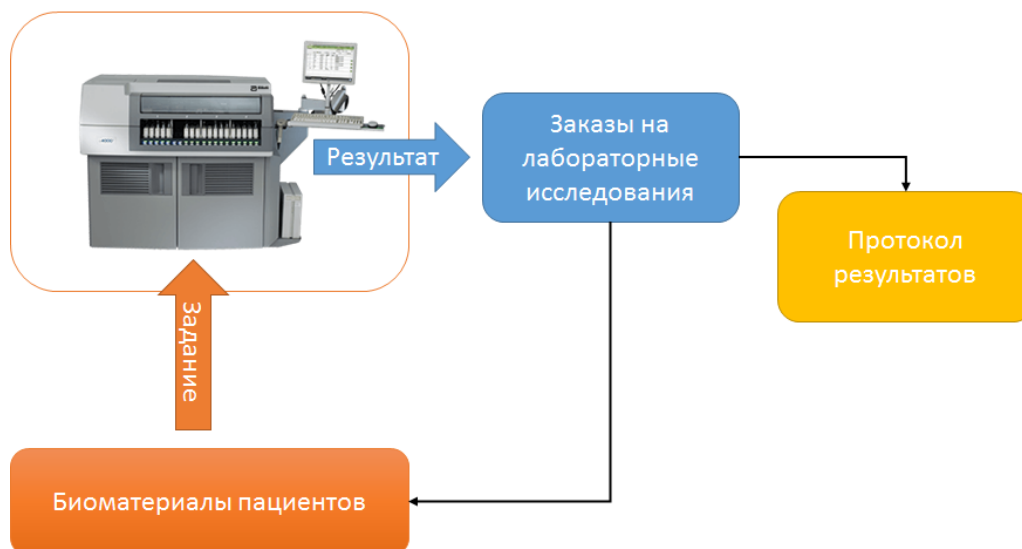


Рисунок 1 - Краткая схема автоматизации

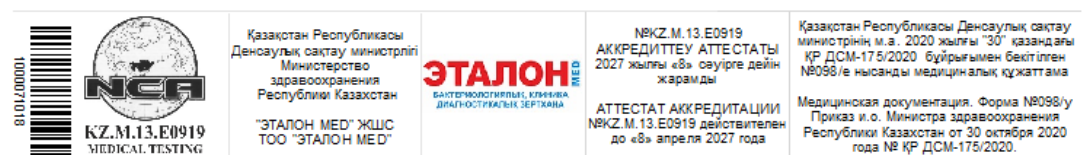
Преимущества алгоритмов и интерпретаций включают предотвращение ошибочных диагнозов, сокращение количества необходимых лабораторных тестов, сокращение количества процедур, переливаний крови и госпитализаций, сокращение времени, необходимого для постановки диагноза, сокращение ошибок при заказе тестов и предоставление дополнительной информации о том, как результаты лабораторных исследований могут повлиять на другие аспекты ухода за пациентом. Сочетание алгоритмов и интерпретаций для лабораторных исследований имеет множество преимуществ для медицинского обслуживания пациента [5] (рис. 2).



Рисунок 2 - Схема интерпретации результатов

Интеграция взаимодействия с лабораторным оборудованием с использованием обмена данными по протоколу RS232/TCPIP в формате ASTM (англ. American Society for Testing and Materials – «Американское общество по испытанию материалов») [6] и HL7 (англ. Health Level 7) [7]. Представленное решение является программным модулем разработанным для лабораторной информационной системы SmartLAB [8] (далее - ЛИС),

который производит комплексную автоматизацию диагностических лабораторий в соответствии с ISO 15189:2012 (Международный стандарт по аккредитации медицинских лабораторий) [9], соблюдая все рабочие процессы для получения достоверных результатов путем непосредственного взаимодействия с лабораторным оборудованием и автоматическому выявлению отклонений от нормативных величин.



**Қанның биохимиялық талдауы**  
**Биохимический анализ крови**

Тегі А.Ж.   Ф.И.О.	[REDACTED]			№	1000071018
Туған күні   Дата рождения	26.08.1991	Жынысы   Пол	М.	ЖСН   ИИН	[REDACTED]
Мекен-жайы   Адрес	ОКРУГ: КУРЫК, АУЛ(СЕЛО): КУРЫК, УЛИЦА: АВТОБАЗА, ДОМ: 55А				
Ұйым атауы   Организация	ГКП на ПХВ "Каракиянская центральная районная больница" УЗ МО				
Бөлім	Отделение специализированной помощи			Дәрігер   Врач	ТҮРІКПЕНБАЕВ А.Қ.
Подразделение					
Биоматериалды алу мерзімі	24.09.2022 08:18	Нәтижені қалыптастыру мерзімі	24.09.2022 12:08		
Дата и время взятия биоматериала		Дата и время готовности результата			
Жолдаманы тіркеу мерзімі	24.09.2022 08:18	Зерттеудегі материал	Сыворотка		
Дата и время регистрации заявки		Исследуемый материал			

Компоненттер, элементтер Компоненты, элементы	Нәтиже Результат	Комментарии	Қалыпты мөлшер Нормативная величина	Бірліктер Ед.изм.
Аланинаминотрансфераза (АЛТ)*	18.3		0 - 45	Ед/л
Аспаратаминотрансфераза (АСТ)*	17.2		0 - 40	Ед/л
Билирубин общий*	6.8		<22,2	ммоль/л
Глюкоза (сахар крови)*	5.26		3.3 - 5.6	ммоль/л
Креатинин	66		62 - 106	ммоль/л
Мочевина	9	▲	2.5 - 6.5	ммоль/л
Общий белок*	79.3		65 - 85	г/л

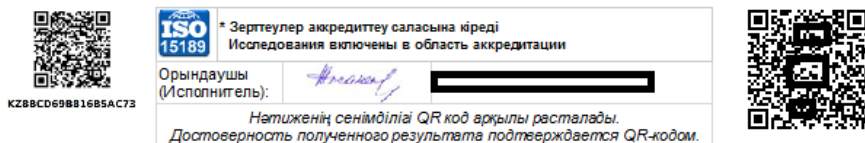


Рисунок 3 - Бланк результата лабораторного исследования

### Закключение

Представленная работа функционирует в виде программного модуля в лабораторной информационной системе SmartLAB. На данный момент времени модуль апробирован и внедрен в двух диагностически лабораториях гг. Актау, Караганда для автоматизации процессов интерпретации результатов исследований.

Материалы, представленные в данной работе, являются результатом практического применения решений в сфере автоматизации лабораторной диагностики, а также демонстрационные данные являются обезличенными во избежание отражения персональных данных пациентов, которые защищаются законом о персональных данных [10].

### Список использованной литературы

1. Государственная программа “Цифровой Казахстан”. О программе: [Электронный ресурс] // Официальный интернет-ресурс Государственной программы «Цифровой Казахстан». Н., 2018-2019. URL:<https://digitalkz.kz/o-programme/>.
2. Глава государства провел совещание по реализации Государственной программы «Цифровой Казахстан» События [Электронный ресурс]. URL:

[https://www.akorda.kz/ru/events/akorda\\_news/meetings\\_and\\_sittings/glava-gosudarstva-provel-soveshchanie-po-realizacii-gosudarstvennoi-programmy-cifrovoi-kazahstan](https://www.akorda.kz/ru/events/akorda_news/meetings_and_sittings/glava-gosudarstva-provel-soveshchanie-po-realizacii-gosudarstvennoi-programmy-cifrovoi-kazahstan).

3. К.К. Кадиркулов, А.А. Исмаилова, Г.Ж. Солтан, А.А. Муханова, М.Маханов, Автоматизация идентификации отклонения результатов лабораторных исследований, 2020. Вестник КазННТУ, ISSN 1680-9211, №4(140), стр. 127-133.
4. Henricks, W. H. (2016). Laboratory Information Systems. Clinics in Laboratory Medicine, 36(1), 1–11.
5. Van Cott, E. M. (2014). Laboratory test interpretations and algorithms in utilization management. Clinica Chimica Acta, 427, 188–192.
6. International Association for Testing Materials: [Электронный ресурс]. 2001-2022. URL: [https://en.wikipedia.org/wiki/ASTM\\_International](https://en.wikipedia.org/wiki/ASTM_International).
7. Health Level Seven International: [Электронный ресурс]. 2013-2022. URL: [https://wiki.hl7.org/Main\\_Page](https://wiki.hl7.org/Main_Page).
8. ТОО “SmartLab Kazakhstan”: [Электронный ресурс]. А., 2015-2022. URL: <http://lis.kz>. (Дата обращения: 17.09.2022).
9. ISO 15189 Medical laboratory accreditation: [Электронный ресурс]. 2020. URL: <https://anab.ansi.org/iso-15189-medical-labs>
10. Закон Республики Казахстан от 21 мая 2013 года № 94-V «О персональных данных и их защите» (с изменениями и дополнениями по состоянию на 03.07.2020 г.): [Электронный ресурс] // Официальный интернет-ресурс zakon.kz, А., 2020, URL: [https://online.zakon.kz/document/?doc\\_id=31396226](https://online.zakon.kz/document/?doc_id=31396226)

# AN INTELLIGENT AGENT OF FINITE STATE MACHINE IN EXAMPLE OF GAMES

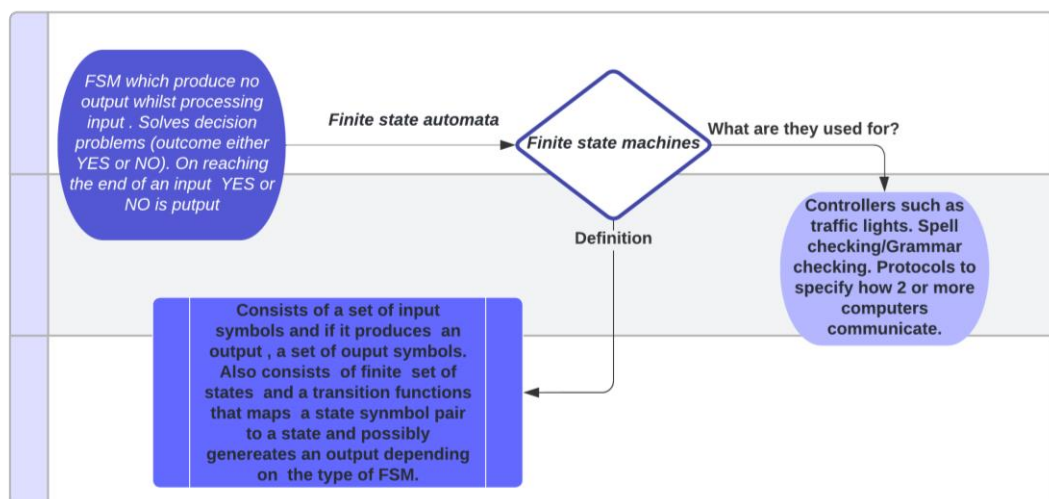
A Shukurova<sup>1, a)</sup>, B Kumalakov<sup>2, b)</sup>, T Zhukabayeva<sup>1, c)</sup>

<sup>1</sup>L. N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan

<sup>2</sup>Astana IT University, Nur-Sultan, Kazakhstan

**Abstract.** An increasing number of artificial intelligence applications are now being used in both computer games, analysis and real-time systems. This paper focuses on Artificial Intelligence as computer systems that can successfully achieve objectives in a 'complex environment'. Game-based simulation learning is designed to simulate existing problems in order to gain knowledge that can be used to solve a problem. What we know from algorithm theory is that any algorithm has a morphism into a finite automaton. On this basis, at a certain level of abstraction, we can assume that a mind is equivalent to a finite automaton. On this basis, at a certain level of abstraction, it is proposed to explore the hypothesis that reason is equivalent to a finite automaton.

Paper is at first motivated by an endeavor to demonstrate social behavior using a finite state machine that could be at that point interpreted to an agent based framework [1]. Game education is exceptionally attractive for improvement. With the assistance of finite state machines - uncommon calculations mindful for the state of in-game objects and characters - it is conceivable not as it were to control the conduct of the AI, but moreover to set models for the working of interfacing. There are a number of points of interest of instructive recreations over conventional learning methods. One of the most points of interest of instructive diversions is the representation of real-world issues [2-4]. A simulation-based instructive amusement is outlined to recreate existing issues in arrange to pick up information that can be utilized to illuminate the issue [4,5]. Social intelligence [6] is regularly defined as Nash balance in a multi-agent framework; at the same time, dialect capacity is defined as a combination of objectives such as parsing, part-of-speech labeling, lexical examination, and estimation investigation; but perceptual capacity is defined as division and protest acknowledgment. Considering an elective speculation: that the common objective of maximizing remunerate is adequate to drive the behavior that most, in the event that not all, of the capacities considered in common and manufactured insights show. For a common understanding of the term limited machines, what they are utilized for, the definition can be seen within the figure underneath (fig.1).



**FIGURE 1.** Finite State Machines mindmap.

## **MATERIALS AND METHODS**

### **Related Work**

Before studying other works and writing the current paper, terminology such as game agent (intelligent agent) and Finite State Machine were studied. In artificial intelligence, intelligent agent is an autonomous entity which observes and act on the environment, which requires agents and directs its activities to achieve the goal of rational [7,8].

The real world is set up in such a way that all processes run in parallel with each other. For example, a person can breathe, listen to a report and think about a thought at the same time. To this are added processes that occur outside the current context.

Games use a mathematical model of the world that depicts continuity and parallelism. In order to successfully emulate these properties of reality, parallel processes must be combined into a single continuity.

Intelligent agents can also learn or use knowledge to achieve their goals. Intelligent agents must demonstrate the following characteristics [9,10]:

1. Being able to analyze their own in terms of behavior, error and success.
2. Learn and improve interaction with the environment (embodiment).
3. Learning quickly from large amounts of data.

From these characteristics can be concluded that the agent is a function that can provide an action in such condition that looks like itself in determining a course of action or decision in carrying members of his functions in an autonomous entity which observes and act on the environment [11,12].

When considering the work of Ilya Prigozhin in [13,14], a definition of complexity is given, from which a whole course within the ponder of strategies for planning and executing items with components of manufactured insights - multi-agent frameworks - begins. Agreeing to [15], when an specialist or gather of specialists employments a technique, and are aware of what that methodology is, at that point that procedure contains a know-how technique. Based on the writing examined, numerous modular coalitional control rationales for know-how techniques have as of now started to be examined by other researchers recently. The center of a multi-agent framework is an cleverly operator [16], which has the taking after common highlights:

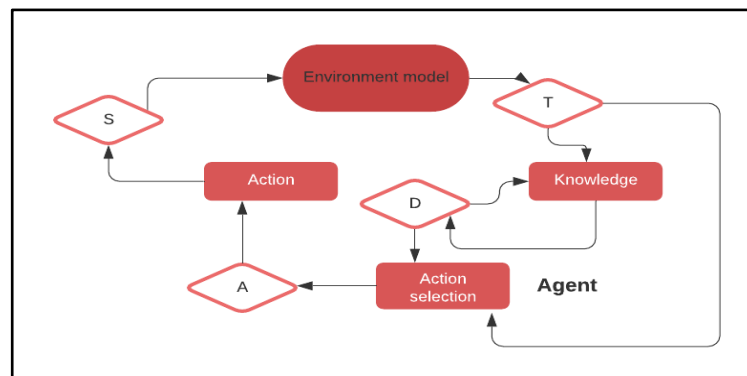
- Independence, which is showed not as it were at the level of framework working, but too of its person specialists (from each other);
- Autonomy. Agents don't have a centralized control instrument. Instep, their behavior is directed by a set of imperatives and rules;
- Emergence. That is, the behavior as a result of uncontrolled interaction of distinctive agents. In practice, rise is additionally called accomplishing of a result by a framework within the case when none of its components has the calculation for its arrangement. In other words, collective work of specialists understands the issues, which separately they are not able to solve;
- Self-organization. On the off chance that a non-standard circumstance emerges, the multi-agent framework must improve itself, in this manner evolving.

From a formal point of view, in any case, agent definition isn't a unimportant assignment. Due to the disparate research bunches within the field, there's no single definition of the term agent. However, numerous researchers concur that the taking after definition covers the fundamental pith: "an agent is an typified computational entity that's in some environment and is able to perform adaptable and autonomous activities to attain a goal" [17]. It is evident from the definition that each agent must have a user-defined objective and perform a "adaptable" grouping of activities to realize it. That's , be free and alter its activities depending on the circumstance in which it needs to work. It takes after that the agent must be able to associated with the environment by getting information through interfacing and reacting to the environment by accessible implies. For illustration, a robot controlled by an agent may see the environment through the utilize of a camera and respond utilizing mechanical body parts such as arms and legs.

The agent's activities are commonly classified as receptive and proactive [16]. An agent's receptive behavior is the behavior in which its activities are a reaction to changes within the environment. Proactive behavior is the behavior in which he performs steady activities to advantage from the expected changes within the environment within the future. Based on this classification, it is troublesome to distinguish an cleverly agent from other computational substances, since real-time frameworks moreover have properties of proactivity and nearly all known programs are receptive in one way or another. Be that as it may, an intelligent agent moreover has social behavior, which is reflected within the arrange and semantic substance of actions.

### Main Part

Reactive agents have a primitive inside show of the outside world. They are characterized by the utilize of the concept of state and the only rules of behavior of the "stimulus-response" sort. Reactive agents are broadly utilized in automata models. They can utilize rules, calculations and transient rationales. Their advantage is straightforwardness and unquestionable status [18]. The most "primitive", but effortlessly executed on an mechanical scale, is the "reactive agent" (Fig. 2) [19].



S - the set of states of the external environment (the agent exists in the environment); D - the database where the agent's knowledge is stored; T - the set of states of the external environment that the agent is able to perceive; A - the set of actions that the agent is able to perform

**FIGURE 2.** Reactive Agent.

- The behavior of this sort of agents is characterized by coordinate execution of modified activities when certain signals are gotten. The most thought communicated in Brooks' works is that intelligent behavior can be accomplished without depicting the rules of equivocal translation of occasions and without programming calculations for elective behavior. It is based on a stack of assignments that are executed in a specific arrangement one after another. Among the points of interest of this course of agents are the relative straightforwardness of depicting frameworks utilizing finite automata, the accessibility of improvement instruments, counting mechanical level, and the speed of decision-making by agents.
- BDI (belief, desire, intention) agents are based on modular rationale portraying the properties of conviction, want and deliberate from a philosophical point of see [20]. Conviction in this setting alludes to information almost the world around, which is accessible to the agent, but may be deficient or off base. The agent's wants are the objectives set by the engineer, whereas the eagerly are the steps essential to realize the wants. For the consistent arrangement and reformulation of eagerly, the agent makes an inner instrument. It moreover shapes a arrangement of eagerly, moreover called a plan.
- The conspire of behavior of BDI agent looks as follows (fig. 3);
- The agent has a specific desire and forms an intention that can realize that desire;
- He commits to act according to the intention;
- Constructs trees of subgoals leading to the realization of the intention, and selects the optimal subtree;

- Selects a subgoal from the goal tree and performs the action leading to its attainment.
- When the agent accomplishes the objective, or gets to be persuaded of its unattainability, he reshapes his behavior agreeing to the same standard cycle. From our point of see this sort of agent is the foremost promising for realization of the marvels of self-organization, development and independence.

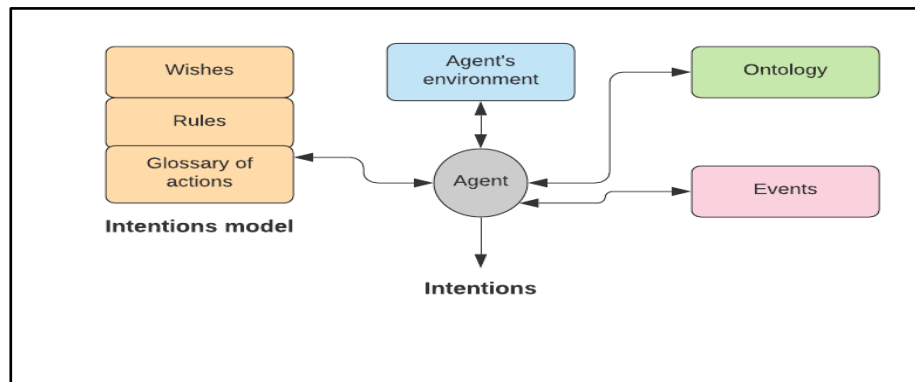


FIGURE 3. BDI Agents.

The issue of modeling the wonders: "belief", "desire" and "intention" could be a generally open address. In this paper, an endeavor is made to demonstrate the wants of adversary agents by modeling the amusement "Mafia". To this conclusion, the taking after formal device is utilized: let us take the definition of an agent [21] as a set:

$$ag = (S, A, env, see, IB, bel, brf, ID, des, drf, plan, prf) \quad (1)$$

where  $S$  - finite set of states of the external environment,  $A$  - non-empty finite set of agent's actions,  $env$  - function of behavior of the external environment,  $see$  - perception of states of the external environment by the agent,  $IB$  - set of representations of the agent,  $bel$  - current representation of the agent,  $brf$  - update function of the representations,  $ID$  - set of wishes,  $des$  - current wish,  $drf$  - wish update function,  $plan$  - current action plan of the agent,  $prf$  - plan update function.

Multi-agent system:

$$MAS = (S, AG, env) \quad (2)$$

where  $S$  is the finite set of states of the environment,  $AG$  is the finite set of agents (corresponding to the definition (1)),  $env$  is the function describing the reaction of the environment to the agents.

In our case, the diversion comprises of discrete stages - days. Each night one of the agents "murders" one resting agent. Amid the taking after day, the objective of the whole community of agents is to distinguish the killer, and the objective of the killer is to mislead the others. In this way there's a articulated hostility between the eagerly of the agents beneath the common discreteness of the diversion model.

Let us denote the set of agents whose goal is to identify the killer as a coalition  $C \subseteq AG$ , and the antagonist agents as  $C \triangleq AG \setminus C$ . Thus the modeling of the opponent's intentions is reduced to collective cooperation. To to represent the collective nature of the phenomena, let us give the following definitions:

- The representation of an agent will be called the Cartesian product of the representations of all agents in the coalition  $bel_{ag} \subseteq S \times ACS \times S$ , and the possible actions of the group of antagonist agents the coalition is sure of as  $sbel_{ag} \subseteq S \times ACS$ . Thus, each agent makes its own plan, which, however, takes into account the perceptions of the other agents. Thus, the search for the "killer" boils down to comparing the actions of the individual coalition members  $sbel_{ag}$  and thus identifying member  $C$ . To this end, let us extend the agent model as follows:

$$ag = (S, A, env, see, IB, bel, brf, ID, des, drf, plan, prf, sbel_{ag}, sbelf), \quad (3)$$

where  $sbel_{ag}$  is the function of matching the behaviors of the agents with respect to  $sbel_{ag}$ .



In order to further discretize the model and model the phenomena at the level of finite automata, we define an extended-interacting finite automaton with time as:

$$M = (I, O, S, D, \sigma, \langle s_0, x_0 \rangle, X, F), \quad (4)$$

where:

$I$  is a finite non-empty alphabet of input symbols;

$O$  - there is a finite non-empty alphabet of output characters;

$S$  - there is a finite non-empty set of states of the automaton;

$D = D_1 \times \dots \times D_n$  - an  $n$ -dimensional set describing possible values of the variables;

$\sigma : S \times I \times X \times D \rightarrow 2S \times O \times m$  is a transition and output function that determines for the current internal state  $s \in S$ , the current values of the variables  $d \in D$ , and the observed input symbol  $i \in I$  at time of move  $x \in X$  the set of possible internal state changes and writes them to output queues. If the return set always consists of a single element, such an automaton is deterministic, otherwise the automaton is nondeterministic;

$\langle s_0, x_0 \rangle \in S \times D$  - the initial state of the automaton and initial values for all variables;

$X = X_1 \times \dots \times X_m$  -  $m$ -dimensional set describing the values of time variables;

$F \subseteq S$  is the set of finite states of the automaton.

In this case, the alphabet  $I$  is defined by the set of values returned by the agent function *see*. The alphabet  $O$  is partially defined by the set of values returned by the agent functions *brf*, *drf*, *prf* and *sbelf*. The set  $S$  is formed from the Cartesian product of values  $IB \times ID \times sbela \times plan$ .  $D$  is formed from  $IB \times ID \times sbelag$ , while  $X$  is the set of all game stages.

## CONCLUSION

Finite automata can be used for a wide range of tasks. They are used not only to describe the behaviour of an AI, but also to operate interfaces and so on. In addition, finite automata can be extended as needed to add a large number of additional states. Agreeing to [22], expanding the quality of the arrangement and the strength to startling occasions requires a better computational taken a toll. Hence, a limited machine competent of describing intelligent BDI agent. Encourage it'll be utilized for the discrete portrayal of agents' intentions, including the calculation for distinguishing adversary agents within the amusement, in which they are not characterized from the starting. Inside the system of the over hypothesis, the concept of intellect was formalized within the system of formalization of the objective accomplishment handle and the method of cognition based on the concept of a finite automaton and an unique hypothetical development realizing this concept was constructed.

## References

1. Kumalakova, Bulletin of KazNPU. Series of "Physical-mathematical sciences" **1(49)**, pp. 52–57 (2015).
2. Kim, D. K., Dinu, L. F., & Kim, C. G. (2018). Adoptability of E-Textbooks Featuring Educational Online Games. In *Gamification in Education: Breakthroughs in Research and Practice* (pp. 338-355). IGI Global.
3. Thornton III, G. C., Mueller-Hanson, R. A., & Rupp, D. E. (2017). *Developing organizational simulations: A guide for practitioners, students, and researchers*. Routledge.
4. Richter, G., Raban, D. R., & Rafaeli, S. (2015). Studying gamification: the effect of rewards and incentives on motivation. In *Gamification in education and business* (pp. 21-46). Springer, Cham.
5. Andrea, R., & Kopel, M. (2018). Design and development of "battle drone" computer-based trading card game (CTCG). In *International Conference on Multimedia and Network Information System* (pp. 574-585). Springer, Cham
6. D. Silver, S. Singh, D. Precup and R. S. Sutton, "Reward is enough," in *Artificial Intelligence* 299 (2021).

7. Yannakakis, G. N., & Togelius, J. (2018). Game AI Panorama. In *Artificial Intelligence and Games* (pp. 259-277). Springer, Cham
8. Helgadóttir, H. E., Jónsdóttir, S., Sigurdsson, A. M., Schiffel, S., & Vilhjálmsson, H. H. (2016). Virtual general game playing agent. In *International Conference on Intelligent Virtual Agents* (pp. 464-469). Springer, Cham.
9. Barrera, R. (2018). *Unity 2017 Game AI Programming-: Leverage the power of Artificial Intelligence to program smart entities for your games*. Packt Publishing Ltd.
10. Rabin, S. (2017). *Game AI Pro 3: Collected Wisdom of Game AI Professionals*. AK Peters/CRC Press
11. Helgadóttir, H. E., Jónsdóttir, S., Sigurdsson, A. M., Schiffel, S., & Vilhjálmsson, H. H. (2016). Virtual general game playing agent. In *International Conference on Intelligent Virtual Agents* (pp. 464-469). Springer, Cham.
12. Gregory, J. (2017). *Game engine architecture*. AK Peters/CRC Press.
13. Prigogine, *Is Future Given?* (World Scientific Publishing Co., Singapore, 2003), pp. 232.
14. G. Nicolis and I. Prigogine, *Exploring Complexity: An Introduction* (WH Freeman and Co, 1989), pp. 247.
15. P. Naumov and Y. Yuan, *Journal of Artificial Intelligence Research* **71** (2021).
16. M. Wooldridge and N. Jennings, *The Knowledge Engineering Review* **10(2)**, pp. 115–152 (1995).
17. M. Wooldridge, *IEEE Proceedings SoftwareEngineering* **144**, pp. 26–37 (1997).
18. S. V. Bulgakov, “Primenenie mul'tiagentnyh sistem v informacionnyh sistemah,” in *PNiO* 5(17), 2015, <https://cyberleninka.ru>.
19. R. Brookes, “Intelligence without Representation,” in *Artificial Intelligence* 47, pp. 139–159 (1991).
20. A. Rao, “Georgeff M. BDI Agents: from Theory to Practice,” in *Proceedings of the 1st International Conference on Multi-Agent Systems*, pp. 312–319 (1995).
21. D. Ju. Bugajchenko, “Razrabotka i realizacija metodov formal'no-logicheskoy specifikacii samonastraivajushhihsja mul'tiagentnyh sistem s vremennymi ogranichenijami,” Ph.D. thesis, Spb, pp. 261, 2007.
22. De Filippo, M. Lombardi and M. Milano, *Journal of Artificial Intelligence Research* **70** (2021).

# БЕТТЕРДІҢ БЕЙНЕ КЕСКІНДЕРІН ТАЛДАУ НЕГІЗІНДЕ ҚАТЫСУШЫНЫҢ ОНЛАЙН ІС-ШАРАДАҒЫ ЗЕЙІНІН АНЫҚТАУ

К.М. Сабырбаева<sup>1</sup>, С.С. Жүзбаев<sup>2</sup>

<sup>1</sup> магистрант, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Қазақстан, Нұр-Сұлтан қ., e-mail: kamshat\_804@mail.ru

<sup>2</sup> физика-математика ғылымдарының кандидаты, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Қазақстан, Нұр-Сұлтан қ., e-mail: juzbayev@mail.ru

**Аңдатпа.** Адамның зейінін анықтау көптеген салаларда маңызды, олардың ішінде көлік жүргізушінің шаршауын анықтау, қызметкерлердің жұмысқа зейінін бақылау немесе студенттің сабаққа қатысуын бағалау. Бұл мақалада біз бейнені талдау арқылы зейінді анықтаудың нақты уақытта (real-time) жұмыс істейтін әдісін ұсынамыз. Бұл әдісте алгоритмнің орындалу жылдамдығын арттыру мақсатында классификатор ретінде MobileNetV2 архитектурасын, ал детектор ретінде RetinaFace архитектурасын қолдандық. Ұсынылған әдіс Student Engagement in the Wild [1] деректер жинағында кадр және бейне деңгейінде 0.07 MSE (Mean Squared Error) шамасын көрсетті.

**Кілтті сөздер:** Конволюциялық нейронды жүйелер, рекуррентті нейрондық жүйелер, детектор, регрессор, трэкер.

## **Кіріспе.**

Интернеттің ыңғайлылығының, тиімділігінің және өткізу қабілетінің жылдам өсуінің арқасында онлайн білім беру танымал бола бастады. Интеллектуалды оқыту жүйелері (ITS) және жаппай ашық онлайн оқыту курстарының (MOOCs) дамуы студенттерге барлық материалдарды оңай меңгеруге және онлайн сертификаттар алуға мүмкіндік берді. Сонымен қатар, COVID-19 пандемиясына байланысты, бүкіл әлем бойынша мектептер жабылды, ал білім беру түбегейлі өзгерді, соның нәтижесінде оқыту қашықтан және цифрлық платформаларда жүзеге асырыла бастады. Көпшілік университеттер, орта мектептер, бастауыш мектептер және басқа да білім беру мекемелері толығымен онлайн оқытуға көшті, ал мұғалімдер онлайн режимінде оқушыларды білімге тарту үшін күресе бастады.

Дәстүрлі сынып жағдайында мұғалім мимика, оқушылардың есінеуі, көз және дене позициясы арқылы оқушылардың сабаққа қатысу деңгейін бағалай алады. Дегенмен, студенттердің қатысуын онлайн режимінде анықтау әлдеқайда қиын. Және де зерттеулер студенттердің 25%-дан 60%-ға дейінгі бөлігі онлайн курстардан шаршайтынын көрсетеді [7]. Нәтижесінде, автоматты зейінді бағалау арқылы онлайн білім берудің сапасын анықтаудың педагогтар мен студенттер үшін маңызы артты.

Зейін оқушы мен ресурс арасындағы эмоционалдық, когнитивтік және мінез-құлықты байланыс ретінде анықталады.

## **Қолданылған деректер жинағы**

Студенттер немесе оқушылардың оқу барысына қатысуын анықтау үшін ашық қолжетімдегі “Student Engagement in the Wild” [1] деректер жиынын қолдану шешілді. Бұл деректер жиынының сипаттамасы келесідей: деректер жиынында барлығы 78 субъект (25 әйел және 53 ер адам) бар. Субъектілердің жас аралығы 19-27 жас. Барлығы 195 бейне жиналады, әрқайсысы шамамен 5 минуттан тұрады. Деректер жинағы шектелмеген ортада, яғни әртүрлі жерлерде жиналады - компьютерлік зертхана, жатақхана бөлмелері, ашық алаң т.б. Бейнелер Microsoft Lifecam кең бұрышты F2.0 камерасына 30 кадр секунд, 640 x 480 пиксель ажыратымдылығында түсірілген. Әр қатысушыны 5 минуттық бейнежазба көріп отырған кезде бейнежазбаға түсірген. Кейін 5 аннотатор әр видеоға 0-ден 3-ке дейінгі аралықта заттаңба берген (1-сурет). Ол заттаңбалар келесідей:

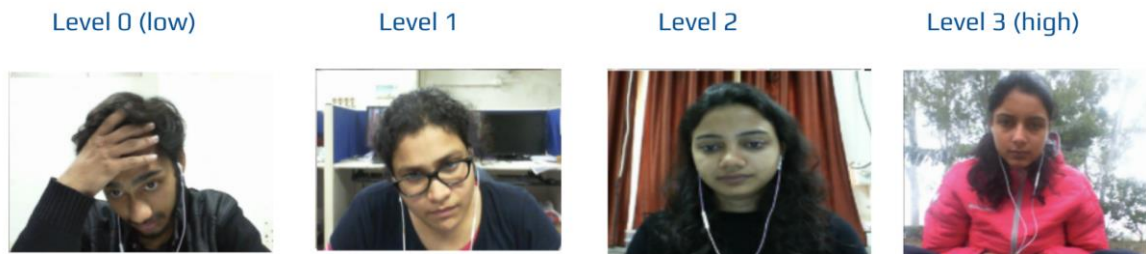
1. 0 - оқу барысына қатысу деңгейі өте төмен
2. 1 - оқу барысына қатысу деңгейі төмен

3. 2 - оқу барысына қатысу деңгейі жақсы

4. 3 - оқу барысына қатысу деңгейі өте жақсы

Деректер жиынындағы бейнелердің таралуы төмендегідей:

0 - деңгейдегі 9 бейнежазба: 1-деңгейде 53, 2-деңгейде 82 және 3-деңгейде 50. Деректер жиыны бір-бірінен тәуелсіз оқыту және тестілеу жиындарына бөлінген. Оқыту жиыны 147 бейнеден, ал тестілеу жиыны 48 бейнеден құралған.



1-сурет. Student Engagement in the Wild деректер жиынынан мысалдар

### Ұсынылған әдіс

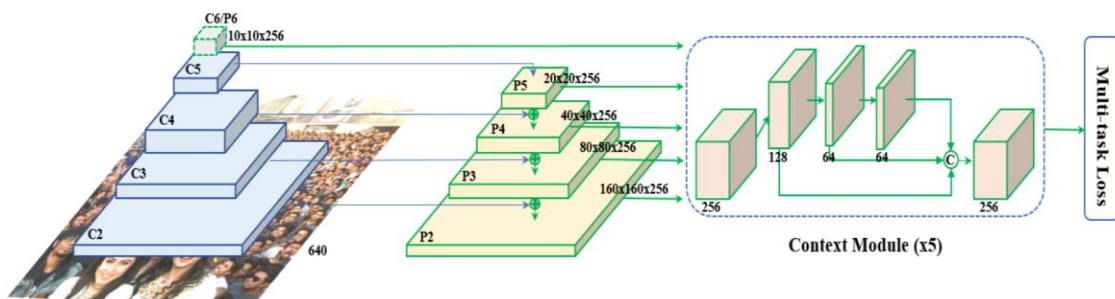
Алгоритмнің жұмыс істеу жылдамдығын арттыру мақсатында рекуррентті нейрондық жүйелерден бас тартып, тек конволюциялық нейрондық жүйелерді қолдандық. Себебі рекурренттік жүйелерде амалдар тізбектей орындалғандықтан жұмыс істеу уақыты ұзағырақ. Алгоритмнің жұмыс істеу барысы келесідей:

1. Әрбір бейнеден тұрақты периодпен кескін алынады
2. Кескіннен детектор алгоритмі арқылы адамның бетінің шекаралары табылады
3. Кескін шекара бойымен қиылып алынады
4. Алынған жаңа кескін зейін регрессиясын жасайтын модельге беріліп, 0-ден 3-ке дейінгі аралықта “зейін деңгейінің болжамы” алынады
5. Кескіндердің болжамдары бейненің жалпы бағасын алу үшін агрегацияланады

Келесі секцияларда қолданылған нақты алгоритмдер тереңірек сипатталады.

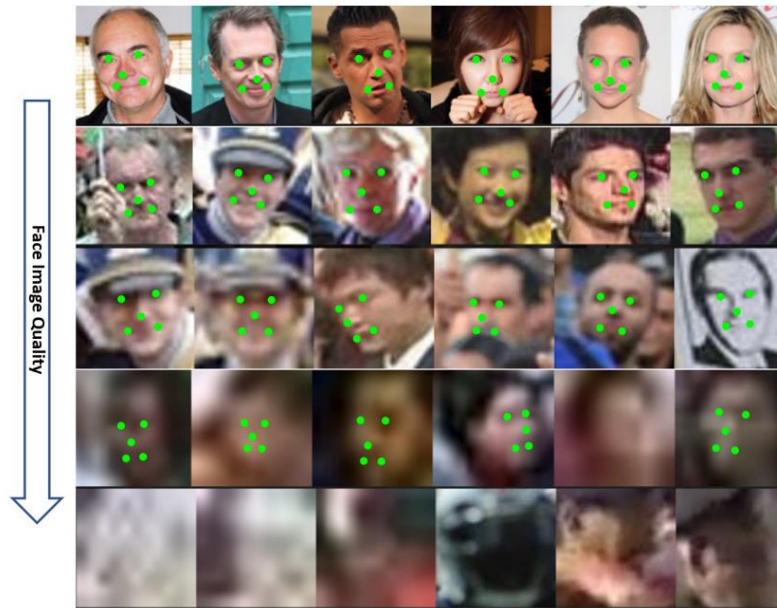
### Детектор

Детектор ретінде RetinaFace [2] моделінің mxnet фреймворкіндегі үлгісі қолданылды. RetinaFace детекторы mtcnn детекторына қарағанда ақырынырақ жұмыс істейді, бірақ эксперименталды түрде RetinaFace дәлірек жұмыс істейтіні анықталды. Сол себепті жылдамдықтан қарағанда дәлдікті таңдадық. Бұл детектордың графикалық процессорда жұмыс істеу уақыты 0.06 секунд. Бұл модельдің архитектурасы 2-суретте көрсетілген.



2-сурет. RetinaFace моделінің архитектурасы [2]

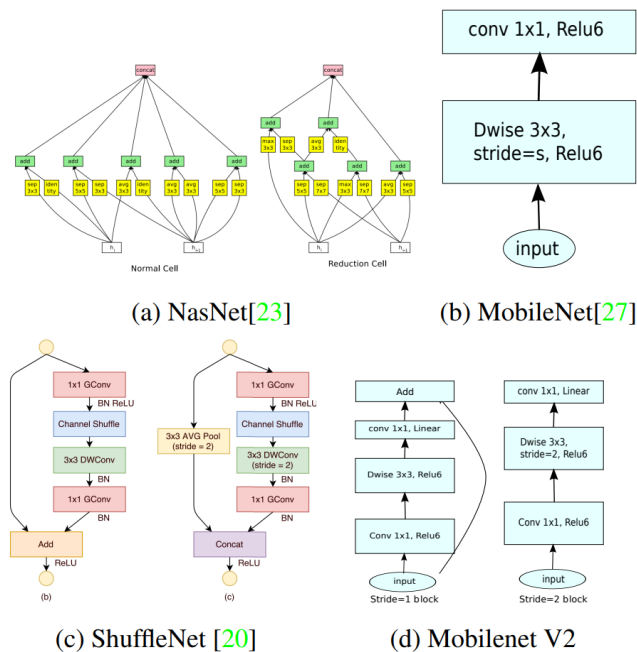
RetinaFace детекторы бет шекарасын табудан басқа беттің негізгі нүктелерін де табады. Негізгі нүктелердің мысалы 3-суретте көрсетілген. Бұл негізгі нүктелер қисық тұрған суретті түзету үшін қолданылады. Түзету модельдің дәлдігін жоғарлататыны эксперименталды түрде дәлелденді.



3-сурет. Негізгі нүктелердің мысалдары [2]

### Регрессор

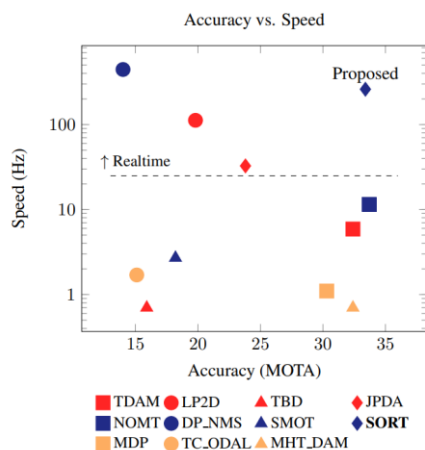
Регрессор моделі ретінде MobileNetV2 [3] моделі қолданылды. Бұл модельдің архитектурасы 3-суретте көрсетілген. MobileNetV2 Depthwise Separable Convolutions [3] негізінде құрылғандықтан resnet және efficientnet архитектураларынан жылдам жұмыс істейді әрі параметрлері әлдеқайда аз. Параметрлердің аздығы графикалық процессордың жадын аз мөлшерде қолдануға мүмкіндік береді. Сол себепті бұл архитектура таңдалды.



4-сурет. MobileNet архитектураларының сипаттамасы [3]

### Бақылау

Бақылау алгоритмі (трэкер) ретінде SORT (Simple online and realtime tracking) [5] алгоритмі қолданылды. Бақылау алгоритмі әр кадр сайын беттерді дұрыс қадағалау үшін қажет.



5-сурет. Бақылау алгоритмін салыстыру диаграммасы

5-суретте SORT алгоритмі жылдамдық және дәлдік параметрлері бойынша басқа бақылау алгоритмдерімен салыстырылған. Бұл сызбадан бұл алгоритмнің секундына 100 кадр өңдей алатынын көре аламыз.

### Регрессия моделін оқыту пайплайны

Әрбір видеодан әр секунд сайын 5 кадр алынады. Әр кадрдан детектор арқылы беттің шекарасы және негізгі нүктелері табылды. Негізгі нүктелердің көмегімен беттер түзетіледі. Әр табылған бетке өз бейнежазбасының заттаңбасы беріледі. Әр сурет 3 каналдан тұратын биіктігі 112 пиксель, ұзындығы 112 пиксель өлшеміне келтірілді. MobileNetV2 моделі осы жинақта 20 кезең оқытылды. “Loss” функциясы ретінде “Mean Squared Error” (MSE) функциясы қолданылды. Оптимайзер ретінде Adam [4] оптимайзері қолданылды. Кадр бойынша алынған  $MSE = 0.07$ . MSE формуласы 6-суретте келтірілген.

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

6-сурет. MSE формуласы

Бұл жерде  $n$  - суреттер саны,  $i$  - суреттер индексі,  $y_i$  -  $i$ -суреттің заттаңбасы.  $\hat{y}_i$  - регрессия моделінің  $i$ -суретке берген заттаңбасы.

Моделді үйрету “mini-batch gradient descent” [6] әдісі бойынша жүзеге асырылды. Бұл әдістің python тіліндегі коды 7-суретте келтірілген. Батч өлшемі 16-ға тең қылып алынды.

```

for i in range(nb_epochs):
    np.random.shuffle(data)
    for batch in get_batches(data, batch_size=50):
        params_grad = evaluate_gradient(loss_function, batch, params)
        params = params - learning_rate * params_grad

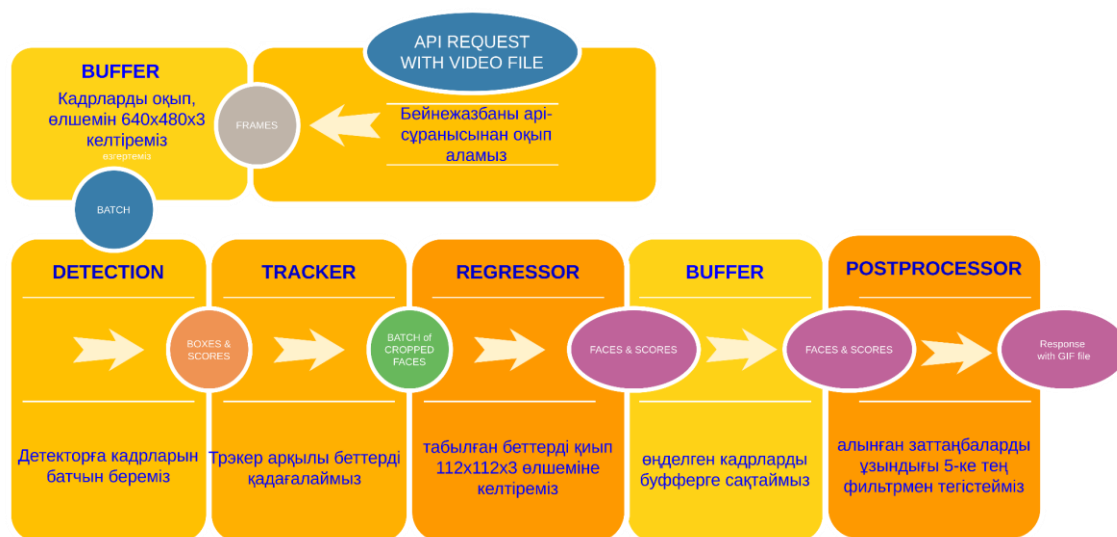
```

7-сурет. Mini-batch gradient descent коды

Үйрету pytorch фреймворкі арқылы жүзеге асырылды.

### Прототип

Прототип арі-сервис ретінде құрылды. Бұл сервис сұраныс ретінде бейнені алып жауап ретінде .gif форматындағы мазмұндама бейнесін алады. Мазмұндамада бейнежазбадағы ең маңызды кадрлар және олардың заттаңбасы келтіріледі. Прототиптің жұмыс істеу сызбасы 8-суретте келтірілген. Прототип FastAPI веб-фреймворк көмегімен жазылды.



8-сурет. Прототиптің жұмыс істеу диаграммасы

Бұл прототиптің Zoom қолданбасымен интеграциясы жасалды.

### Қорытынды

Интернеттің дамуы және COVID-19 тарауының нәтижесінде онлайн курстар саны күрт артты. Бұл үдеріс қатысушылардың зейінін қадағалаудың жаңа тәсілдерін қарастыру қажеттілігін туғызды. Бұл мақалада бұл қажеттілікті конволюциялық нейрондық жүйелер арқылы шешу жолын қарастырдық. Нақты уақытта (real-time) жұмыс істейтін әдісті ұсындық және Student Engagement in the Wild деректер жиынында 0.07 MSE көрсеткішіне қол жеткіздік. Және прототип сервис жүзеге асырылып, Zoom қолданбасымен интеграция жасалды.

### Пайдаланылған әдебиеттер тізімі

1. Kaur A., Mustafa A., Mehta L., Dhall A. Prediction and Localization of Student Engagement in the Wild. // Digital Image Computing: Techniques and Applications (DICTA). - 2018. - pp. 1-8.
2. Deng J., Guo J., Zhou Y. RetinaFace: Single-stage Dense Face Localisation in the Wild. // arXiv preprint arXiv:1905.00641. - 2019.
3. Sandler M., Howard A., Zhu M., Zhmoginov A., Chen L. MobileNetV2: Inverted Residuals and Linear Bottlenecks // arXiv preprint arXiv:1801.04381. - 2019.
4. Kingma D., Lei Ba J. A method for stochastic optimization. // arXiv preprint arXiv: 1412.6980. - 2019.
5. Bewley A., Ge Z., Ott L., Ramos F., Upcroft B. Simple online and realtime tracking. // arXiv preprint arXiv: 1602.00763. - 2017.
6. Ruder S. An overview of gradient descent optimization algorithms. // arXiv preprint arXiv: 1609.04747. - 2017.
7. R. Larson., M. Richards. Boredom in the Middle School Years: Blaming Schools versus Blaming Students. // American Journal of Education 99, 4. - 1991. - pp. 418–443.
8. Zhang K., Zhang Zh., Li Zh. Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks. // arXiv preprint arXiv:1604.02878. - 2016.
9. He K., Zhang X., Ren Sh., Sun J. Deep Residual Learning for Image Recognition. // arXiv preprint arXiv:1604.02878. - 2015.
10. Tan M., Le Q. Efficient Model Sclaning for Convolutional Neural Networks. // arXiv preprint arXiv:1604.02878. - 2020.

## РАЗРАБОТКА МОДЕЛИ ПОРТАЛА НАУЧНЫХ ЗНАНИЙ

Садирмекова Ж.Б.<sup>2,3</sup>, Самбетбаева М.А.<sup>1,2</sup>, Боранкулова Г.С.<sup>3</sup>

<sup>1</sup>Евразийский национальный университет им. Л.Н.Гумилева, г. Астана, Казахстан

<sup>2</sup>Институт информационных и вычислительных технологий, г. Алматы,  
Казахстан

<sup>3</sup>Таразский региональный университет имени М.Х.Дулати, г. Тараз,  
Казахстан

(E-mail:[janna\\_1988@mail.ru](mailto:janna_1988@mail.ru), [madina\\_jgtu@mail.ru](mailto:madina_jgtu@mail.ru), [B.gau@mail.ru](mailto:B.gau@mail.ru))

### Введение

В мировом информационно-образовательном пространстве на базе современных информационных технологий созданы и функционируют различные информационные системы (ИС), связанные с деятельностью вузов. Для реализации современных задач развития информационной среды образования в большей степени пригодны порталные технологии. Порталы предназначены для объединения разнородного контента, сервисов и организации унифицированного доступа к ресурсам и сервисам.

Ресурсы портала должны быть взаимосвязаны с помощью взаимных ссылок. Несмотря на то что подходы к проектированию и реализации порталов достаточно отработаны, на фоне лавинообразного увеличения объема данных в интернете существует необходимость в их качественном улучшении. Именно этим и объясняется **актуальность** этой работы. Так, к основным недостаткам современных порталов мы относим:

- ограниченную поддержку семантики хранимой информации, что затрудняет поддержку средств навигации, поиска и т. д.;

- сложности обеспечения интеграции и обмена данными с другими системами, ввиду невозможности использования стандартных решений для интеграции информации различных веб-ресурсов.

– Одним из способов решения этих проблем является использование технологий Semantic Web, соответствующих стандартов и инструментов при сборе, обработке, хранении, поиске и распространении информации. С помощью этих технологий для описания информационных ресурсов можно использовать модели знаний, представленные в виде веб-онтологии. Порталы, использующие для выполнения своих функций технологии Semantic Web, называются семантическими. Основными преимуществами семантических порталов является реализация следующих услуг:

- классификация веб-ресурсов портала на основе онтологии, поиск и улучшенная навигация по структуре информации;

- поддержание ИС портала путем агрегирования децентрализованных источников;

- возможность интеграции для автоматизированного обмена данными между системами;

- возможность расширения структуры информации.

Порталы часто классифицируются по типам данных и инструментов, которые они предоставляют пользователям. Горизонтальные порталы охватывают различные темы, а вертикальные порталы считаются узкоспециализированными. В зависимости от функциональности различают информационные порталы, порталы совместной работы, порталы экспертизы, порталы знаний.

В данной работе мы обратимся к образовательному portalу – специализированному portalу, обеспечивающему систематическое представление сведений о конкретной области знаний и содержательный доступ к информационным ресурсам, соответствующим выбранной области.

В информационных системах, созданных на основе традиционных технологий создания сайтов, отсутствуют развитые средства описания содержания интернет-ресурсов с целью поиска сторонних систем данных, автоматизированной интеграции данных между ними,



обмена данными. Поэтому перспективными типами порталов являются порталы знаний или семантические порталы.

На основе анализа электронных ресурсов, а также опыта зарубежных информационных систем сформулированы основные требования к содержанию и функциональности портала кафедры [1].

Для успешного выполнения заданий портал должен быть создан как открытая система, модель данных которой должна характеризоваться онтологией предметной области. Онтологии, опубликованные в Интернете, не могут быть получены для формирования модели порталных данных, поскольку они не полностью описывают эту предметную область и не используют словари и классификаторы, часто используемые в международном сообществе, для повышения уровня интероперабельности.

#### **Разработка онтологии портала**

В качестве концептуальной модели портала использовался формализм онтологии. Под онтологией мы понимаем четкое детальное описание (модель) той или иной части мира, относящееся к той или иной области интересов. Основываясь на приведенном определении, мы опишем полученную онтологию с точки зрения многих классов понятий, их свойств, отношений и аксиом, которые накладывают некоторые ограничения на использование классов, свойств и отношений.

Формально онтология портала представляет собой формулу вида:

$$O = \langle C, A, R_C, T, D, R_A, F \rangle,$$

где  $C$  – множество классов, описывающих понятия некоторой предметной или проблемной области;

$A$  – множество атрибутов, описывающих свойства понятий и отношений;

$R_C = \{r_c \mid r_c \subseteq C \times C\}$  – множество отношений, заданных на классах (понятиях);

$T$  – множество стандартных типов значений атрибутов (*string, integer, real, date*);

$D$  – множество доменов (множеств значений стандартного типа *string*);

$R_A = R_{AT} \subseteq R_{AD}$ , где  $R_{AT} \subseteq A \times T$  – отношение, связывающее атрибуты и типы данных, из которых они могут принимать свои значения;

$R_{AD} \subseteq A \times D$  – отношение, определяющее для каждого атрибута его дискретное множество значений (домен);

$F$  – множество ограничений на значения атрибутов понятий и отношений.

С содержательной точки зрения выявленная таким образом онтология может служить для представления понятий, необходимых как для описания определенной области знаний, так и для научной деятельности, осуществляемой в ее рамках. Вводя формальные характеристики понятий предметной области в виде классов объектов и отношений между ними, порталная онтология задает структуры, отражающие фактические данные и их корреляцию. Данные на портале кафедры представлены в виде множества связанных информационных форм.

Методика построения онтологии для портала кафедры, с одной стороны, полностью определяется ее структурой, а с другой стороны, обеспечивается и ограничивается средствами, предлагаемыми редактором онтологии. В процессе развития онтологии выделяются и формально описываются классы понятий и их свойства. Свойства каждого понятия представлены с помощью атрибутов и ограничений, накладываемых на область их значений, а также бинарных отношений, связывающих данное понятие с другими понятиями. Классы понятий строятся с использованием отношения наследования к иерархии ("общие – частные"). Механизм задан таким образом, что из понятия "родитель" в понятие "поколение" передаются не только все его атрибуты, но и отношения.

Структура онтологии портала кафедры. Для удовлетворения целей портала кафедры онтология должна быть хорошо структурирована и адекватно отражать ее

проблемную и предметную область. В качестве основных онтологий были выбраны две онтологии.

Первая онтология характеризует проблемную область системы (рис. 1) и не зависит от предметной области системы. Это онтология верхнего уровня, включающая в себя классы понятий, относящиеся к деятельности кафедры, такие как *Персона*, *Организация*, *Научная деятельность*, *Научные мероприятия*, *Публикация*, *Географическое место*, *Сборник материалов конференции*, *Подразделение*, *Проект*, *Конференция* и т. д. Опишем понятия подробнее:

- ресурс «*Персона*» описывает человека как сотрудника кафедры;
- ресурс «*Организация*» включает вуз, его филиалы, научные базы, научные центры и другие организации, с которыми связана деятельность кафедры. Отражает организационно-структурное деление организации, позволяет получить сведения о ее структурных подразделениях и обеспечивает доступ к информационным ресурсам этих подразделений;
- ресурс «*Публикация*» структурирует данные о разного рода документах и публикациях, как печатных, так и цифровых (в статьях, журналах, книгах, нормативных документах и проч.);
- с ресурсом «*Подразделение*» связана информация о подразделениях организации, различных объединениях сотрудников и учащихся (советы, союзы);
- ресурс «*Проект*» описывает любые проекты или гранты, в рамках которых осуществляется научная работа кафедры;
- ресурс «*Конференция*» описывает любую официальную встречу ученых с целью обмена научной информацией. Например, конференцию или семинар.

Вторая онтология – онтология предметного знания. Она задает метапонятия для описания понятий возможных предметных областей, задающие структуры для описания понятий конкретной области знаний, такие как *Учебный курс*, *Компетенция*, *Задача учебного курса*, *Методы решения задач*, *Результат освоения курса*. Метапонятие «*Учебный курс*», например, обобщает сведения об учебных дисциплинах и курсах. Содержимое учебных курсов (дисциплин) формируется в разрезе специальностей или направлений подготовки. Понятия этих двух онтологий связаны ассоциативными отношениями. Их выбор возможен не только с учетом полноты представления проблемной и предметной областей системы, но и с учетом удобства и функциональности навигации. Такие онтологии не только описывают предметную и проблемную область системы, но и обозначают структуры, отражающие конкретные объекты и связи между ними. Совокупность таких информационных объектов и их связей образует информационное содержание или содержание портала. На рисунке 1 представлены основные онтологии портала кафедры.

Развитие онтологии основано на повторяющемся алгоритме, это особый процесс, поэтому предложенные онтологии кафедры могут служить основой для дальнейшего индуктивного увеличения глубины детализации, реализации систем высших учебных заведений с новыми функциональными возможностями.

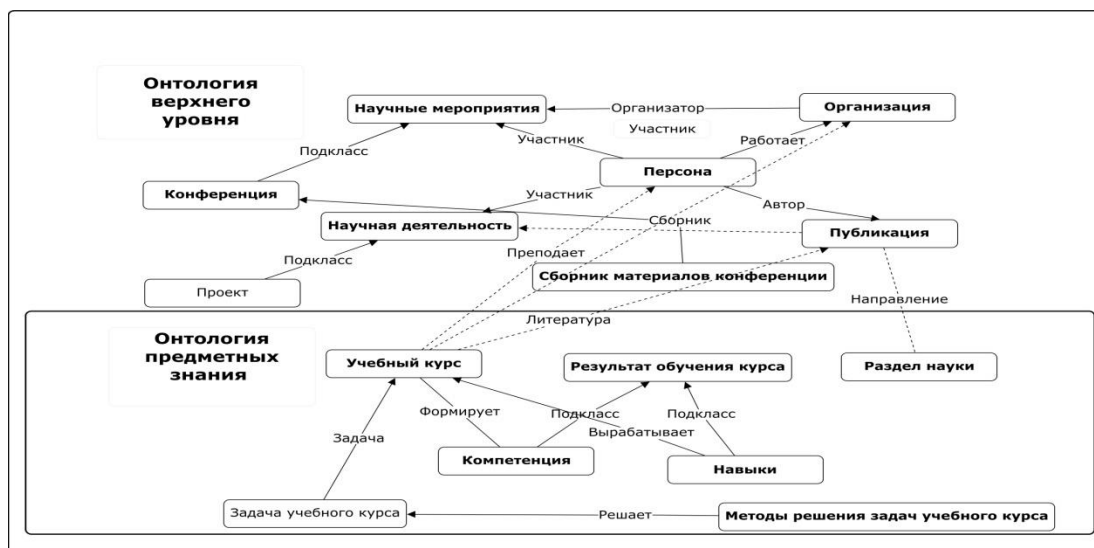


Рисунок 1. Базовые онтологии портала кафедры

Помимо обеспечения пользовательского интерфейса и информационных потоков, при создании онтологии следует также говорить о сохранении данных. В нашем случае в качестве хранилища используется специальная база данных, в которой одновременно хранится содержание портала кафедры и описание онтологии.

Модель RDF была разработана в конце 1990-х годов, а в 2001 г. Т. Бернерс Ли [2] опубликовал статью, положившую начало эре семантической паутины (Semantic Web). Интернет-сообщество наблюдало рост интереса ко всему, что связано со семантической обработкой, в том числе к модели данных RDF, принятой в качестве стандарта комитета W3C в 2004 году. Основание RDF [3] данные, представленные в виде высказываний (тройка, тройка), описывают " субъект-сказуемое-объект", объектно-ориентированную связь от субъекта к объекту. Для определения этой связи используется Uniform Resource Identifier (URI), который является обобщением понятия URL.

Мы сравнили несколько наиболее распространенных инструментов друг с другом на основе критерия, учитывающего возможность выступать в качестве хранилищ RDF [4].

Для портала кафедры необходимо хранилище со следующими требованиями:

1. Поддержка запроса данных через SPARQL Select.
2. Поддержка изменение данных через SPARQL Insert.
3. Поддержка логического вывода.
4. Полное бесплатное использование сервиса.
5. Мультиплатформенность.

В качестве хранилища нами был выбран Jena Fuseki Web Server, так как библиотеки для работы с ним имеют открытый исходный код. С ним удобно работать как через SPARQL HTTP-клиент, так и посредством библиотеки для Jena Framework. Для отлаженного функционирования портала будет достаточно использовать SPARQL HTTP-соединение, а в дальнейшем Jena Fuseki Web Server может быть заменен на любой аналог. Так как в портале кафедры будет наличествовать редактор данных, то есть возможность использовать Virtuoso, так как он поддерживает изменение данных с помощью SPARQL INSERT.

Следует отметить, что подобное хранилище не привязано к данным и не требует дополнительных настроек, зависящих, например, от онтологии, поэтому его можно назвать универсальным и использовать для построения порталов других предметных областей.

### Заключение

При выполнении этой работы была построена онтология портала кафедры. Данные на портале представляются в виде множества связанных информационных объектов. Для

этой области была создана антология в Редакторе protégé с использованием языка Owl, который включает в себя более 52 классов, 1523 объекта, 68 отношений, 1741 экземпляр отношения, 95 атрибутов и 2844 атрибута.

Таким образом, использование онтологии в качестве информационной модели портала позволяет детально описать каталог ресурсов по определенной теме и их взаимосвязь, а также организовать входные и выходные параметры, что значительно повышает производительность системы.

Работа поддержана грантом финансирования научных, научно-технических проектов на 2022-2024 гг. Комитетом науки Министерства образования и науки Республики Казахстан (AP14972834).

#### Список использованной литературы

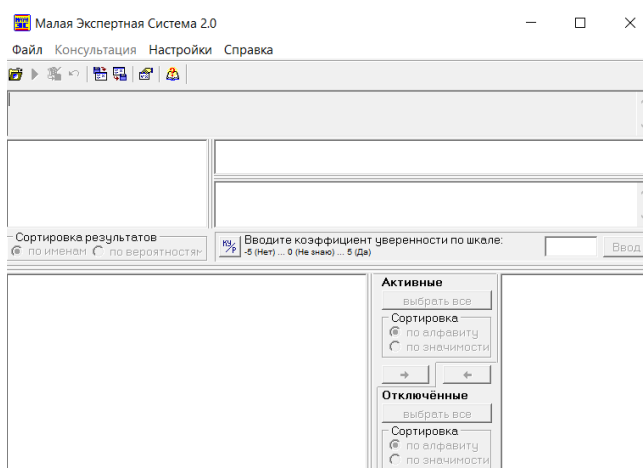
- [1] Sadirmekova Zh.B., Zhizhimov O.L., Tussupov D.A., Sambetbayeva M.A. Requirements for information system to support scientific and educational activities // CEUR Workshop Proceedings . – Novosibirsk, Russia, 2019. – P. 44–47.
- [2] Sadirmekova Zh.B., Tussupov J.A., Sambetbayeva M.A., Altynbekova Zh.T. Development of integrated information systems to support scientific activity // SIST 2021 - 2021 IEEE International Conference on Smart Information Systems and Technologies, art. No. 9465964.
- [3] Serikbayeva S.K., Sadirmekova Zh.B., Batyrkhanov A.G., Sambetbayeva M.A., Yerimbetova A.S. Development of Technology to Support Large Information Storage and Organization of Reduced User Access to this Information // International Journal of Advanced Computer Science and Applications. – 2021. – Vol. 12(7). – P. 493–503.
- [4] Sadirmekova Zh.B., Tussupov D.A., Sambetbayeva M.A., Yerimbetova A.S., Zagorulko Y.A. Features of the development of intelligent scientific and educational internet resources // 6<sup>th</sup> International Conference on Computer Science and Engineering. – Ankara-Turkey, 2021. – P. 392–397.

# БАЙЕС ЛОГИКАЛЫҚ ҚОРЫТУ ЖҮЙЕСІНЕ НЕГІЗДЕЛГЕН САРАПТАМАЛЫҚ ЖҮЙЕНІ ЖАСАУДЫҢ МЫСАЛЫ

*Сарсенбаева А.К., Төлепберген А.А., Сүйіндік Н.М., Алимova Ж.С.  
Павлодар қ, Торайғыров университеті, pgu@tou.edu.kz*

«Сараптамалық жүйе» (expert system) – жасанды интеллект әдістері мен пайдаланушы ұсынатын деректер негізінде жағдайды ұқсастыруға, диагноз қоюға, болжам жасауға, шешімді өзгертуге, әрекетті таңдауға, ұсыныс беруге қабілетті білімдерді қамтитын программалық және аппараттық құралдардың жүйесі. Сараптамалық жүйе, әдетте, кейбір пән саласына (мысалы, жүрек-қан тамырлары ауруларына диагноз қою мен емдеудің сараптау жүйесі) бейімделеді және осы саладан білімді алуға, жинақтауға, түзету енгізуге, белгілі білімдерден жаңа білім тудыруға, осы білімдердің негізінде практикалық мәселерді шешу мен шешу барысын түсіндіруге қабілетті болады [1].

Бұл мақалада, сараптамалық жүйелердің «Кіші сараптамалық жүйе 2.0» – Байес логикалық қорыту жүйесіне негізделген қарапайым бағдарламасын қолдану мысалы қарастырылады (1-сурет). Бұл бағдарламаның басты артықшылықтарының бірі – пайдаланушының өзі жасаған білім базасын құру, өңдеу және өзгерту мүмкіндігі. Сонымен қатар, бағдарламаны орнату үшін нақты жүйелік талаптар жоқ, ең бастысы Windows ОЖ қолдайды. Бағдарламаны орнату MiniES2Install.exe файлын іске қосып, және ұсынылған орнатушының нұсқауларын орындау арқылы жүзеге асырылады. Бағдарламаның авторы – Алексей Бухнин [2].



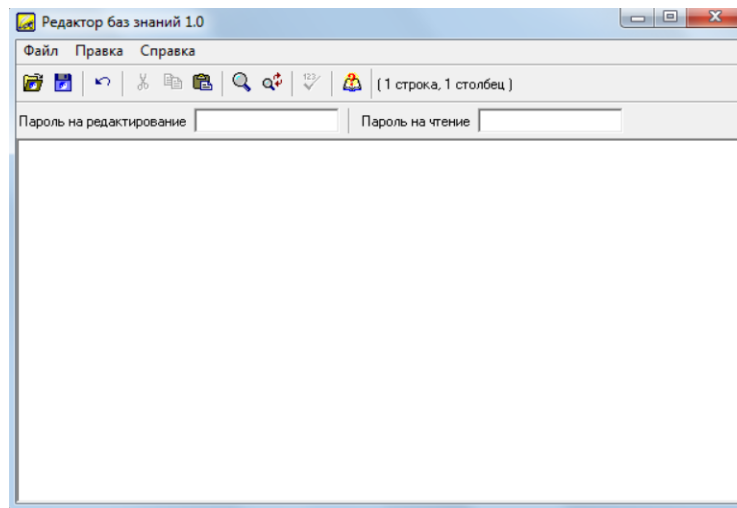
1-сурет – Кіші сараптамалық жүйенің интерфейсі 2.0

Атап өткеніміздей, «Кіші сараптамалық жүйе 2.0» бағдарламасының маңызды артықшылығы – өзіңіздің білім базаңызды құру, өңдеу және пайдалану мүмкіндігі [2]. Бұл тапсырманы жеңілдету үшін жүйемен бірге берілген Білім базасының редакторы 1.0 (2-сурет) қолданылады.

Редактордың мүмкіндіктері:

- 64 килобайттан үлкен файлдарды өзгерту;
- нақты уақытта курсордың орнын көрсету мүмкіндігі;
- мәтін бөліктерін іздеу және өзгерту;
- «Кіші сараптама 2.0 жүйесін» іске қоспай қалыптасқан білім базасын қателерге тексеру;
- білім базасының парольдерімен жұмыс.

Редактор білім базасын өзгерту және оқу үшін құпия сөздерді орнатуға мүмкіндік береді. Білім базасы блокнот бағдарламасында да құрылуы мүмкіндігі бар және бұл жағдайда файлды .mkb форматында сақтау керек.



2-сурет – Білім базасы редакторының интерфейсі 1.0

Білім базасы – бұл үш бөлімнен тұратын құрылымы бар, кейінгі шифрлау мүмкіндігі бар мәтіндік файл [3]:

1) Білім базасының сипаттамасы, автордың аты, түсініктеме және т. б. (жазбаны жалпы ұзындығы 10000 таңбадан аспайтын бірнеше жолда пайдалануға болады; бұл бөлім бірінші бос жолдан кейін аяқталады).

2) № 0 куәлік (жолды ауыстырумен аяқталатын көлемі 1000 символдан аспайтын кез-келген мәтін)

№ 1 куәлік

№ 2 куәлік

...

№ N куәлік (соңғы куәліктен кейін бір бос жол жүреді және екінші секция аяқталады).

3)

Қорытынды № 0, P [ i, P<sub>y</sub>, P<sub>n</sub> ]

Нәтиже № 1, P [ i, P<sub>y</sub>, P<sub>n</sub> ]

Қорытынды № 2, P [ i, P<sub>y</sub>, P<sub>n</sub> ]

Қорытынды № M, P [ i, P<sub>y</sub>, P<sub>n</sub> ]

Соңғы бөлімде шығу ережелерінің сипаттамасы берілген. Сипаттаманың басында нәтиже анықталады, оның ықтималдығы осы ережелерге байланысты өзгереді. Бұл үтірлерден басқа әртүрлі таңбалардан тұратын мәтін. Үтірден кейін осы нәтиженің априорлық ықтималдығы жазылады (P), яғни, қосымша ақпарат болмаған жағдайда, нәтиженің орын алу ықтималдығы. Егер мүмкін нәтиже дұрыс және дұрыс болмаса, осы сұраққа «иә» жауап беру ықтималдығы. Бұл деректер осы нәтижемен байланысты әрбір сұрақ үшін жазылады. Айта кету керек, 0.00001-ден кіші немесе оған тең нөлге тең, ал p 0.99999-дан үлкен немесе оған тең, сондықтан мұндай мәндерді көрсетпеу керек, мұндай априорлық ықтималдығы бар нәтиже өңделмейді [4].

Мысалы ретінде, үй жануарларын белгідері бойынша анықтауды қарастырайық. Білімдер базасын жазуда, инновациялық тұрғыдан жаңартылу жоспарларға сәйкес, латын тілі таңдалды. Төмендегідей білім базасы (3-сурет) қолданылды:

Avtor: Adai Tolepbergen, Sarsenbaeva Aruzhan, Suiindik Nurtas.

Suraqtar:

Tumsyq uzartylgan ba?

Qanattary bar ma?

Qanagattangan diril artqy soqqylarga zhauap bere me?

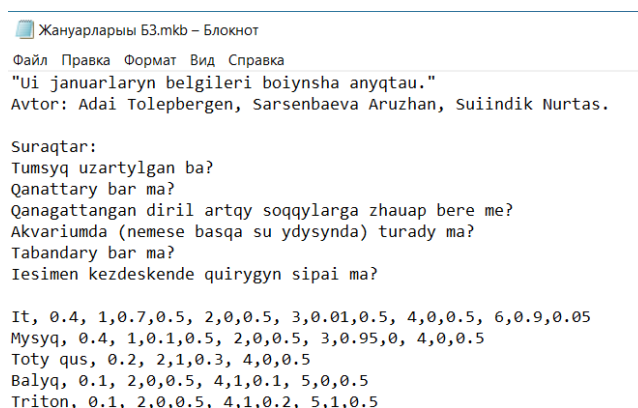
Akvariumda (nemese basqa su ydysynda) turady ma?

Tabandary bar ma?

Iesimen kezdeskende quirygyn sipai ma?

It, 0.4, 1,0.7,0.5, 2,0,0.5, 3,0.01,0.5, 4,0,0.5, 6,0.9,0.05

Mysyq, 0.4, 1,0.1,0.5, 2,0,0.5, 3,0.95,0, 4,0,0.5  
 Toty qus, 0.2, 2,1,0.3, 4,0,0.5  
 Balyq, 0.1, 2,0,0.5, 4,1,0.1, 5,0,0.5  
 Triton, 0.1, 2,0,0.5, 4,1,0.2, 5,1,0.5

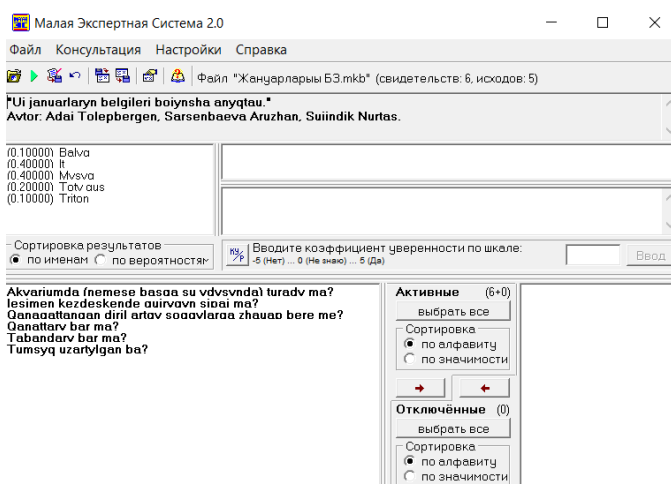


3-сурет – Үй жануарларын белгідері бойынша анықтау үшін құрылған білім базасы

Диагностиканы бастау үшін «Сараптамалық жүйемен кеңес беруді бастау» батырмасын немесе <F3> пернені басу қажет. Бұл нұсқа пайдаланушының екі жауап нұсқасын қолдайды:

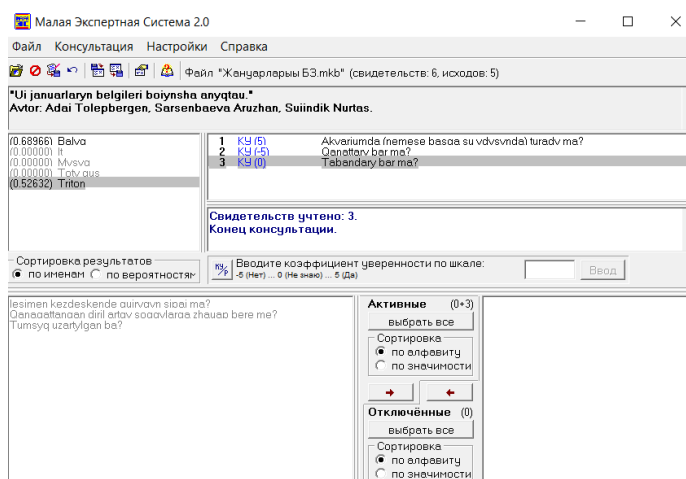
- 1) Белгілі бір шкала бойынша сенімділік коэффициентін орнатамыз (мысалы, «дәл жоқ» дегенді білдіретін -5-тен, «дәл иә» дегенді білдіретін +5-ке дейін);
- 2) Куәліктің шындық ықтималдығын енгізу (сан нөлден бірлікке дейін).

Бірінші және екінші жағдайда да пайдаланушы кез-келген аралық мәндерді таңдай алады. Жауап енгізу шақыруының сол жағында орналасқан ki/P түймесі жауап нұсқалары арасында ауысуға мүмкіндік береді. Мұны <F8> пернесімен де жасауға болады [1]. «Кіші сараптама 2.0 жүйесінде» құрылған білім базасын ашамыз (4-сурет).



4-сурет – «Кіші сараптама 2.0 жүйесінде» ашылған білім базасы

Енді диагностиканы бастауға және жүйенің сұрақтарына жауап беруге болады (5-сурет).



5-сурет – Жүйе диагностикасының мысалы

Диагностика нәтижелерін алфавит немесе ықтималдық бойынша сұрыптауға болады. Бұл жүйе, жоғарыда айтылғандай, Байес логикалық қорыту жүйесін қолданады. Бұл дегеніміз, сараптамалық жүйемен өңделген ақпарат мүлдем дәл бола алмайды, керісінше ықтималды сипатқа ие [5]. Пайдаланушы куәліктің абсолютті ақиқатына немесе жалғандығына толық сенімді бола алмайды, ол тек жүйенің сұрауларына кез-келген сенімділікпен жауап бере алады. Өз кезегінде, жүйе кеңес беру нәтижесін нәтижелердің болжамды ықтималдығы түрінде көрсетеді.

#### Әдебиеттер

1. А.Қ.Құсайынов. Қазақ тілі терминдерінің салалық ғылыми түсіндірме сөздігі: Информатика және компьютерлік техника / Жалпы редакциясын басқарған түсіндірме сөздіктер топтамасын шығару жөніндегі ғылыми-баспа бағдарламасының ғылыми жетекшісі, педагогика ғылымдарының докторы, профессор, Қазақстан Республикасы Мемлекеттік сыйлығының лауреаты А.Қ.Құсайынов. – Алматы: «Мектеп» баспасы» ЖАҚ, 2002 жыл. – 456 б.
2. «Кіші сараптамалық жүйе 2.0 білім базасының редакторы 1.0», Пайдаланушы нұсқаулығы, 2000 ж.
3. Сараптамалық жүйелерді құруға арналған қабықшалар [Электрондық ресурс] // Bourabai Research. URL: <http://bourabai.ru/alg/expert22.htm>.
4. Новикова В. А., Андреева Е. Ю., Туйкина Д. К., жасанды интеллект және сараптамалық жүйелер [Электрондық ресурс] // сараптамалық жүйелер. URL: [http://expro.ksu.ru/materials/ii\\_i\\_es/book.html](http://expro.ksu.ru/materials/ii_i_es/book.html).
5. Столярова А.И., Севостьянова А. В. медициналық ақпараттық жүйелерді енгізу мәселелері туралы / / «ғылымдағы, басқарудағы, әлеуметтік саладағы және медицинадағы Ақпараттық технологиялар» халықаралық конференциясының еңбектер жинағы 2015 ж. – 338-340 б.
6. Петеляк В. Е. Автоматты-алгоритмдік схемаларды қолдана отырып, орташа және аз күрделіліктегі бағдарламалық жасақтаманы модельдеу // ғылым мен білімнің заманауи мәселелері: Хгу оқытушыларының университетшілік ғылыми конференциясының XLV материалдары. – Магнитогорск: МаГУ, 2007. – 257-258 б.



## METHODS AND SYSTEM FOR ANALYSIS OF BIOMEDICAL IMAGES IN DYNAMIC CHANGES

Ainur Orazayeva, Jamalbek Tussupov, Sergii Pavlov, Natalia Babyuk  
L.N. Gumilyov Eurasian National University, Republic of Kazakhstan  
e-mail: oaris.83@gmail.com, tussupov@mail.ru

Vinnytsia National Technical University, Vinnytsia, Ukraine, psv@vntu.edu.ua

**Introduction.** In the theory of digital signal processing, an image is considered to be an informative two-dimensional matrix that contains all the characteristics of the displayed object, which are determined by the methods and processes of the received image [1, 2]. The elements of this matrix are the values of the color intensity function. Here, the value of color is understood as generalized, which is dependent on the type of images. For example, in the case of halftone images, the matrix element is the intensity (hue, gradation) of gray. Additional parameters can be considered clarity, contrast, optical density, sharpness, etc. [3, 4].

For color images, the element of the matrix is determined by the triad of parameters [1, 2] of the color model, which can be combined into one whole number with the help of defined additive rules. Integer representation is commonly used for image storage or transfer, or in entry-level graphics digital processing systems (GDSs). In professional processing, color is mainly represented by three basic parameters: tone, saturation and intensity, or parameters of the color representation model. The color tone is determined by the peak and distribution in the spectrum of visible optical radiation. Saturation (intensity of a certain tone) is the degree of difference between the brightness of a chromatic color and gray. A completely desaturated color will be a shade (gradation) of gray [5,6,7].

In addition to the specified parameters, the following are often used:

- lightness parameter (clarity) – the inverse degree of difference from white color;
- the concept of color depth - a term used to indicate how many bits are needed to store information about the color of a pixel. Color depth is measured in bits per pixel [8,9,10].

### **Methods of preprocessing of biomedical information.**

With the development of image analysis tools, the requirements for pre-processing results are increasing. Traditional methods of processing one-dimensional signals are less and less meeting the criteria of intellectual analysis. As a result, performing scientific and engineering tasks while working with visual data requires more effort [3, 6]. This is most clearly manifested in the case of the creation of new types of distributed information and analytical systems, systems of intelligent management and remote monitoring, since their work is based on the processing of visual information [2, 3].

In some cases, pre-processing methods can be used at the post-processing stage for further application of algorithms in artificial intelligence systems.

It should be noted that during image processing, the main task is to improve the quality of images. This task is complex and includes not only tasks of improving clarity, contrast, noise removal, etc. [11, 12], but also image scaling.

Since the process of obtaining graphic information is, as a rule, multi-level, it is appropriate to distinguish four levels of processing and analysis of biomedical images and their corresponding information processes (Table 1)

Where  $R(x,y,t)$  is the reflecting capacity of the object;  $I'(x,y,t)$  is the image at the output of the image forming system;  $I(x,y,t)$  – image of improved quality;  $f\{.\}$  – processing operator;  $P(x,y,t)$  is a set of image parameters;  $C_v(x,y,t)$  is a set of features describing the  $v$ -th object;  $K_v$  is a decision based on  $C_v$  features.

Table 1. Stages of processing and analysis of biomedical images [13]

	Stages of processing and analysis of biomedical images	Processes conversion	Target purpose

		information	level
1.	Preprocessing	$R(x, y, t) \rightarrow f\{I'(x, y, t)\} \rightarrow I(x, y, t)$	Improve image quality
2.	Segmentation (general and detailed analysis)	$I(x, y, t) \rightarrow P(x, y, t)$	Image analysis objects
3.	Feature selection and analysis	$P(x, y, t) \rightarrow C_V(x, y, t)$	Identification
4.	Classification	$C_V(x, y, t) \rightarrow K_V$	Solutions

**Structural implementation of a system for processing biomedical images.**

The structural diagram of the system for processing biomedical images shown in fig. 1 contains subsystems that provide for the following stages of image processing [12]:

- receiving (entering) an image;
- conversion and editing;
- selection of objects or phases in the image;
- determination of textural areas and features;
- saving and printing images and analysis results.

One of the most important (in terms of the value of the information obtained) and complex (in terms of implementation) stages of biomedical information analysis is segmentation. The purpose of segmentation is to reduce the volumes of processed visual information and to ensure the possibility of applying different processing methods during further analysis of each section.

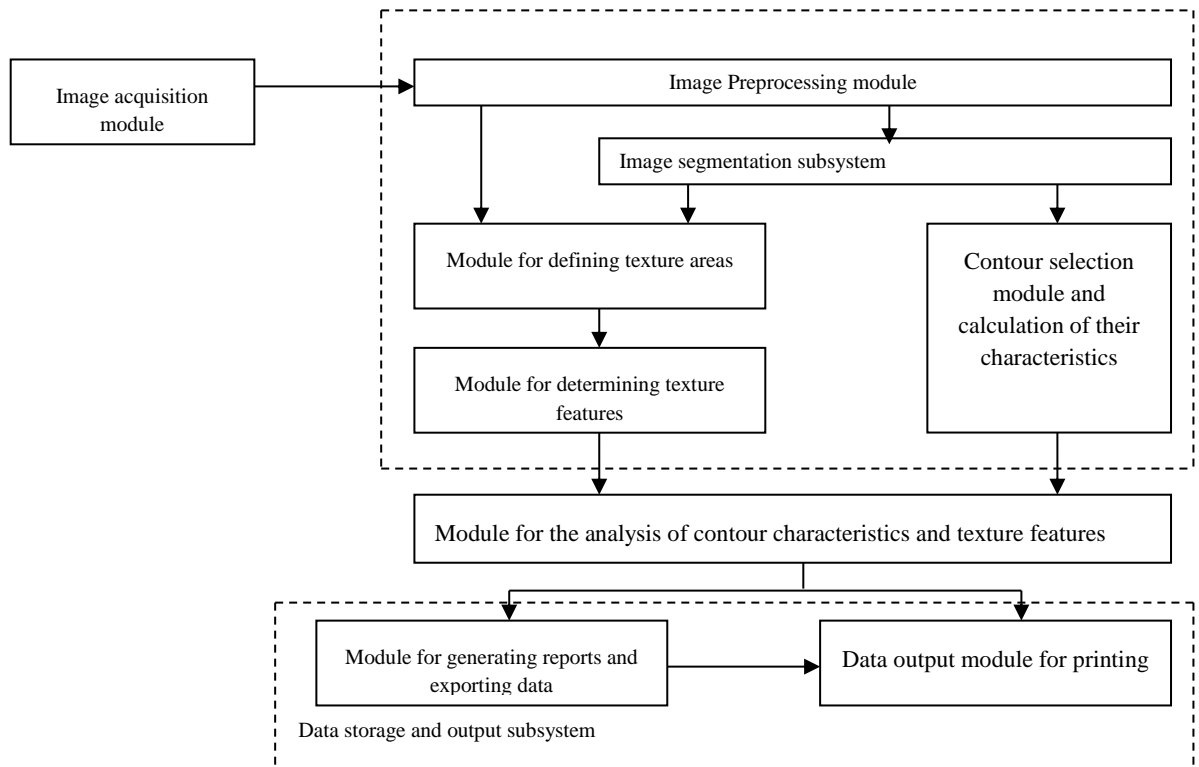


Figure 1 - Structural diagram of a system for processing biomedical images [11,12]

When processing graphic information, as a rule, a statistical texture model is used. According to this model, the texture is presented in the form of a random field, the numerical characteristics change at the boundaries of the areas that are homogeneous for the texture. For the segmentation of such textures, the amplitude-detector method [11,12] is used, which consists of the following: a local statistical evaluation is performed; mathematical expectation, which is subtracted pixel by pixel from the input image; a local statistical evaluation of the root mean square deviation is performed. For this, depending on the properties, either an amplitude or a

quadratic detector is used.

As a result, the matrix of the obtained values can be considered as a matrix of intensities; contour segmentation of the received field is performed.

Studies have shown that this method has high speed, accuracy and interference resistance. In Fig. 2. the developed biomedical image processing scheme is given [7]. Stages and processing tasks are classified for each of the processes.

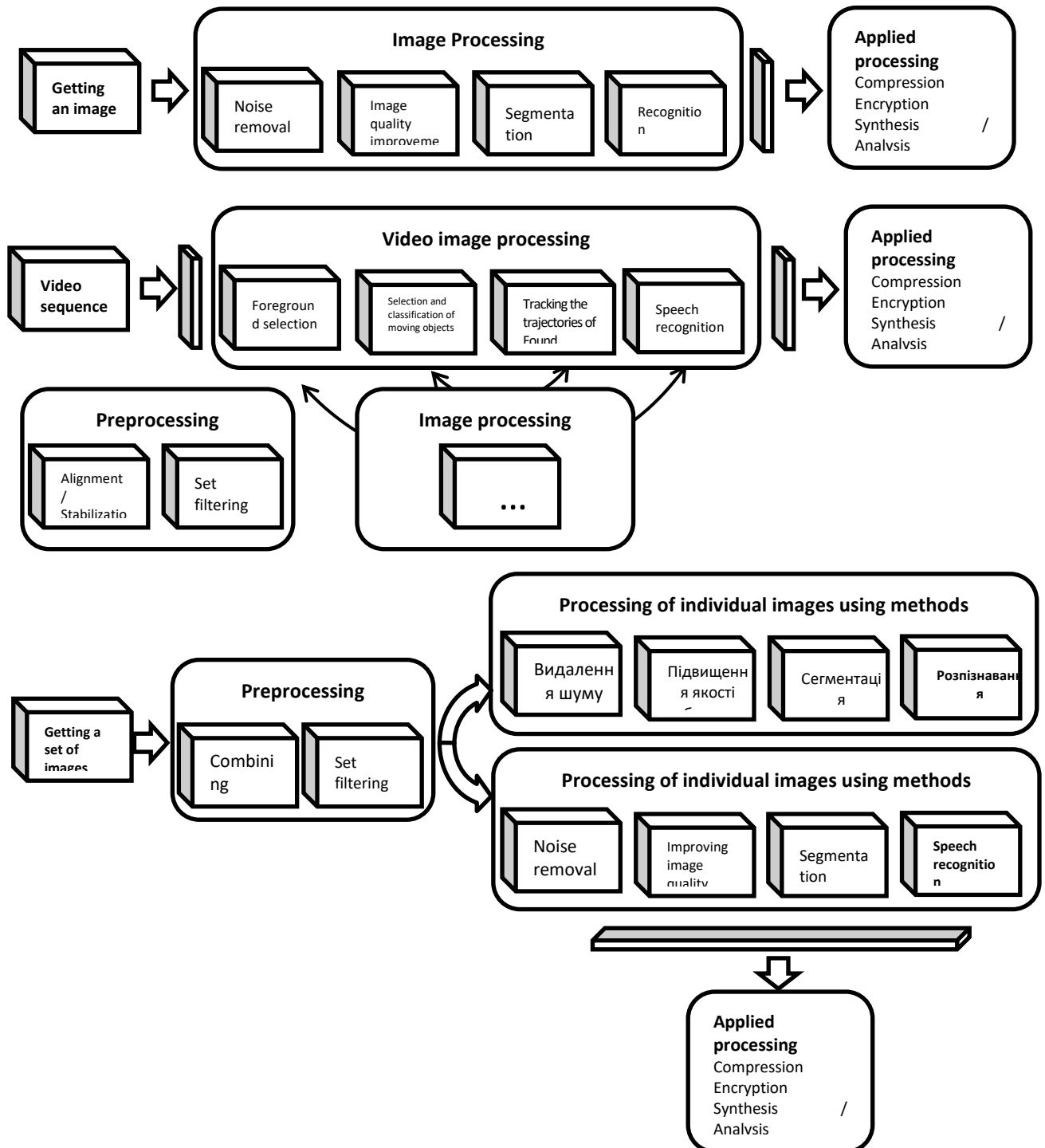


Figure 2 – General scheme of biomedical image processing

**Biomedical image processing algorithm.** To eliminate defects, improve quality and additional processing in information technology, image preprocessing functions are necessary:

- selection of a part of the image;

- image scaling;
- conversion from one color base to another;
- image brightness correction;
- color gamut correction.

In order to calculate contour features, information technology should include additional processing of the received data. Functions that provide additional analysis include (Table 2):

- contour passing function;
- major axis selection function;
- feature selection of characteristic points;
- contour approximation function.

Table 2. Comparison of algorithms for selection of biomedical objects [12]

	Algorithm	Accuracy	Speed selection	Level automation
1.	Pixel-by-pixel selection	the highest	low	missing
2.	Block selection	high	high	high
3.	Selection based on characteristic points	high	the highest	the highest

Based on the analysis of existing methods and systems of biomedical image processing, it can be concluded that they are used to achieve certain tasks: improving the quality of images, compensating defects in the registration system, reducing noise; calculation of clinically important quantitative parameters (distance, area, volume, etc.); facilitation of interpretation (recognition of structures and neoplasms); establishing feedback.

**Experimental results.** The development of systems for evaluating biomedical images, namely for determining important elements of images, such as neoplasms and other diagnostic features, is a promising direction for research [9, 10].

During the experiment, 30 X-ray images of the mammary glands were taken as initial images. Each of the images has dimensions of 1024x1024 (fig. 3).

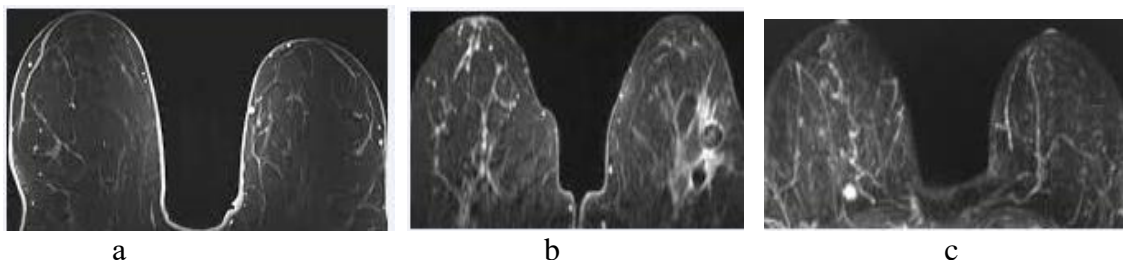


Figure 3 - a) sample №1 of the image of the mammary gland without pathology; b) sample № 2 of the image of the breast on the right side with different stages of pathologies; c) sample № 3 the image of the breast on the left side with different stages of pathologies

**Conclusion.** The article shows that the task of creating a system for processing biomedical images of pathologies in the analysis of mammary gland cancer is urgent and needs to be solved, since the existing methods and evaluation systems, as well as the existing methods of processing biomedical images, do not meet the modern requirements for such systems according to accuracy, reliability of processing of biomedical images, which leads to errors in making a diagnosis in mammographic studies.

#### References

1. S. O. Romanyuk, "Approximation of bidirectional reflectance distribution function for highly efficient shading", in Monography *Information Technology in Medical Diagnostics*, W.

- Wójcik and A. Smolarz, London: England: CRC Press, 2017, chapter 2, pp. 27-49. doi:10.1201/9781315098050.
2. A. Maier, S. Steidl, V. Christlein, and J. Hornegger. *Medical Imaging Systems*, Erlangen-Nürnberg, Germany: Springer, 2016.
  3. Гонсалес Р. Цифровая обработка изображений Москва: Техносфера, 2005. - 616с: 874-892.
  4. Коренной А.В. Математические модели полутоновых изображений / А.В. Коренной // Радиотехника. – 2007. - №8. – С. 79-81.
  5. Воробель Р. Повышение контраста изображений с помощью модифицированного метода кусочного растяжения / Р. Воробель, И. Журавель // Отбор и обработка информации. – 2000. – №14(90). – С. 116-121.
  6. Цмоць І.Г. Алгоритмічні та матричні НВІС–структури пристроїв ділення для комп'ютерних систем реального часу: Научно-технический журнал “Автоматика. Автоматизация. Электротехнические комплексы и системы” Херсон, 2004. – №1(13). – С. 97-105.
  7. Ильясова, Н.Ю. Информационные технологии анализа изображений в задачах медицинской диагностики / Н.Ю. Ильясова, А.В. Куприянов, А.Г. Храмов. - М.: Радио и связь, 2012. - 424 с.
  8. Zorina Nizhynska-Astapenko, Sergiy Pavlov, Oleg Vlasenko, Waldemar Wojcik, Ainur Orazayeva, and etc. "Information medical fuzzy-expert system for the assessment of the diabetic ketoacidosis severity on the base of the blood gases indices", Proc. SPIE 12126, Fifteenth International Conference on Correlation Optics, 2021, 1212626).
  9. Orazayeva A. R., Tussupov D.A., Pavlov S.V., & Abdikerimova G.B. (2022). Efficiency of processing breast cancer biomedical images using filters. Proceedings of the National Academy of Sciences of the Republic of Kazakhstan. Physics and Information Technology Series, (1), 2022, 69–76.
  10. Wójcik, W., Pavlov, S., Kalimoldayev, M. (2019). Information Technology in Medical Diagnostics II. London: Taylor & Francis Group, CRC Press, Balkema book. – 336 Pages, <https://doi.org/10.1201/9780429057618>. eBook ISBN 9780429057618.
  11. Pavlov S. V. Information Technology in Medical Diagnostics //Waldemar Wójcik, Andrzej Smolarz, July 11, 2017 by CRC Press - 210 Pages. <https://doi.org/10.1201/9781315098050>. eBook ISBN 9781315098050.
  12. Pavlov Sergii, Avrunin Oleg, Hrushko Oleksandr, and etc. System of three-dimensional human face images formation for plastic and reconstructive medicine // Teaching and subjects on bio-medical engineering Approaches and experiences from the BIOART-project Peter Arras and David Luengo (Eds.), 2021, Corresponding authors, Peter Arras and David Luengo. Printed by Acco cv, Leuven (Belgium). - 22 P. ISBN: 978-94-641-4245-7.

## ҮЛГІНІ ТАНУДА ҚОЛДАНЫЛАТЫН ӘДІСТЕРДІ ҚАРАСТЫРУ

Қасенхан А.М., Ускенбаева Р.К., Нуралықызы С., Джунусова С.М.  
Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті,  
akassenkhan@gmail.com

### Андатпа

Үлгілерді тану статистика мен инженериядан басталады; үлгіні танудың кейбір қазіргі тәсілдері үлкен деректердің қолжетімділігінің ұлғаюына және өңдеу қуатының жаңа молдығына қатысты болғандықтан, машиналық оқытуды пайдалануды қамтиды. Алайда, бұл әрекеттерді бір ауқымның екі аспектісі ретінде қарастыруға болады және олар бірге соңғы бірнеше онжылдықта айтарлықтай қарқынмен өсуде. Бұл мақалада негізгі үлгіні тану үрдісіндегі басты алгоритмдері мен әдістері жайлы қысқаша тоқталып кеткен. Зерттелу аумағы үлкен болғандықтан, ең көп қолданылатын әдістерге анализ жасап кету керек екеніне бел будық.

**Кілттік сөздер:** үлгіні тану, деректерді талдау, деректерді өңдеу, үлгіні танудағы әдістер

### Кіріспе

Есептеуіш техниканың жетілуі мен өмір сүру үрдісінде шығатын бірқатар мәселелерді шешуге және оның нәтижесін жеңілдетуге, өңдеу жылдамдығын арттыруға және сапасын жақсартуға мүмкіндік туды. Мысалы, тіршілікті қамтамасыз етудің әртүрлі жүйелерінің жұмысы, адам мен компьютердің өзара әрекеттесуі, роботты жүйелердің пайда болуы және т.б. Алайда, қазіргі уақытта кейбір мәселелерді, соның ішінде қозғалыстағы бейнені тану, және оның нәтижесін өңдеу әлі де қиын, тіпті мүмкін емес екенін байқаймыз. Жалпы айтқанда, бұл мақала тапсырманы орындау үшін есептеулерде қолданылатын әдістер мен принциптерді талқылауды ұсынады [1].

Үлгі бойынша кескінді іздеу мәселесі үлгіні танудың неғұрлым жалпы тапсырмасының бөлігі болып табылады. Объектілер жиынтығынан "ұқсас" объектілерді жүйеленбеген және бағытталмаған іздеу кезінде олар шексіз тізімге енгізілуі мүмкін және берілген ықтималдықпен аяқталмайды. Нақты жағдайларда объектілер пішіні, түсі, орналасуы, қозғалғыштығы сияқты идентификациялық өлшемдермен, айқындалатын белгілерімен, олардың әр түрлі сипаттарымен және т.б. көрсетіледі.

Мәселені қарастырған кезде әрқашан ең қарапайымнан бастанған жөн. Ажыратымдылығы жоғары бейнетаспаны алу өте жоғары ажыратымдылық алгоритмін жасаудан әлдеқайда оңай. Оптикалық тану әдістерінде қатаң проблемалық мәлімдеме жүйелік бағдарламалау мәселелеріне қарағанда әлдеқайда маңызды:

### Әдістер

Тану мәселелерінде әмбебап шешімдер жоқ. Бірден "кез келген бейнені танитын" алгоритм жасау оңай емес. Көшедегі белгі мен мәтін парағы түбегейлі әртүрлі объектілер. Әмбебап алгоритм жасау мүмкін, бірақ бұл үлкен топтан көп жұмыс жасауды қалайды және ондаған түрлі ішкі бағдарламалардан тұрады.

Қандай да бір әмбебап кеңес беру және де компьютерлік көрудің ерікті мәселелерінің шешімін құруға келетін құрылымды қалай құру керектігін түсіндіру өте қиын. Бұл мақаланың мақсаты - қолдануға болатын әдістер мен тәсілдерді көрсету.

Қолданыстағы әдістерді үш топқа бөліп қарастып шығайық. Бірінші әдіс – ең алдымен сүзгілеу және кескінді өңдеуге келістіру. Екінші әдіс – сүзгілеу нәтижелерін логикалық өңдеу. Үшінші әдіс – логикалық өңдеуге негізделген шешім қабылдау алгоритмдері. Әдістер арасындағы байланыс өте ерікті. Мәселенің соңына жету үшін барлық топтардың әдістерін қолдануды қажет етпейді, бір немесе екі жақсы келетін әдістерді қолданған әлдеқайда тиімдірек [2].

## 1. Фильтрация.

Бірінші сүзуді қарастырайық. Бұл топқа бейнелердегі қызықты аймақтарды қарастырмай таңдауға мүмкіндік беретін әдістерді орналастырдық. Бұл әдістердің көпшілігі кескіннің барлық нүктелеріне біркелкі түрлендірудің қандай да бір түрін қолданады. Сүзгілеу деңгейінде бейне қарастырылмайды, бірақ сүзгіден өткен нүктелерді арнайы қалпы бар аймақтар ретінде қарастыра аламыз. Ішіндегі жеңілдетілген түрлендіру – кескінді сүйретінің шегі бойынша бинаризациялау. RGB (қосымша түсті модель) және сұр реңді кескіндер үшін түстің мәні болады. Мұндай түрлендіру жеткілікті болатын тамаша мәселелер бар. Мысалы, автокөліктердің нөмірлерін анықтау [3].

Сегменттеу және тану процедуралары екілік кескінмен, яғни тек қара және ақ пикселдермен жұмыс істейді. Сондықтан жұмысты осы процедураларға көшірмес бұрын бастапқы түсті кескінді екілік пішінге келтіру керек. Бұл тапсырма екі кезеңде шешіледі. Бірінші кезеңде түсті кескін ақ-қара түске ауыстырылады және сұр реңкте көрсетіледі. Әрбір пиксель үшін оның жарықтығы нөлден 255-ке дейінгі диапазонда есептеледі. 0 жарықтық деңгейі қара түске, 255-деңгейі ақ түске сәйкес келеді. Осылайша, кескінді сақтау үшін пиксельге бір байт пайдаланылады [2].

Бинаризация гистограммалармен жұмыс істегенде өте қызықты нәтижелер бере алады, оның ішінде суретті RGB-сыз, HSV (түсті модель) түрінде қарастыратын болсақ. Бұл принцип бойынша жазбалар детекторын да, адам келбетінің детекторын да құруға болады. Радар мен белгілерді өңдеуден алынған классикалық сүзгілеу әдістері әртүрлі үлгіні тану мәселелерінде сәтті қолданылуы мүмкін. Радардағы дәстүрлі әдіс, оның таза түрінде кескіндерде ешқашан қолданылмайтыны - Фурье түрлендіруі (нақтырақ айтқанда, FFT). Бір өлшемді Фурье түрлендіруі қолданылатын бірнеше ерекшеліктердің бірі кескінді қысу. Кескінді талқылау үшін әдетте бір өлшемді түрлендіру жеткіліксіз, қорды әлдеқайда көп қажет ететін екі өлшемді түрлендіруді қолдану керек. Оны нақты есептейтін зерттеушілер аз, әдетте жоғары (HPF - High pass filter) немесе төмен (LPF - Low pass filter) жиіліктерге дейін өткірілген қалыптасқан сүзгінің көмегімен қызығушылық аймағының біріктірілмесін қолдану әлдеқайда тез және оңай. Мұндай әдіс, әрине, спектрді талдауға мүмкіндік бермейді, бірақ бейнені өңдеудің нақты мәселесінде әдетте талдады керек етпейді, негізінде нәтижесі болғаны әлдеқайда маңыздырақ. Төмен жиіліктерді (Гаусс фильтрі) және жоғары жиіліктерді (Габор сүзгісі) баса көрсететін сүзгілердің қарапайым көрінісі деп атасақ болады. Әрбір бейне нүктесі үшін терезе таңдалады және бірдей өлшемдегі сүзгі арқылы көбейтіледі. Мұндай конвульсияның нәтижесі нүктенің жаңа мәні болады [3].

Классикалық қолданбасы алмасуы немесе оптикалық ағындарды табу үшін бейне ағынының корреляциясы. Ең қарапайым алмасу детекторы да белгілі бір мағынада айырмашылық корреляторы болады. Бейнелер сәйкес келмейтін аумақта қозғалыс болғандығын байқайды. Бейнелерді сүзу кезінде бұл қажет құрал.

Сүзгілердің арнайы тобы - *сүзу функциялары*. Бұл кескіндегі жай математикалық функцияны анықтауға мүмкіндік беретін таза математикалық сүзгілер (шаршы, куб және тағы басқа формаларды анықтау). Жинақтаушы кескін құрастырылады, ол жерде бастапқы кескіннің әрбір нүктесі үшін оны тудыратын функциялар жиынтығы жазылады. Ең классикалық түрлендіру сызықтар үшін Хафа түрлендіруін (Hough Transform) атап айтып кетуге болады. Кездескен пішіндерді іздеуге мүмкіндік беретін өзгертілген түрлендіру бар. Бұл түрлендіру математика мамандарына қатты ұнайды. Бірақ кескіндерді өңдеу кезінде ол, барлық жағдайда жұмыс жасай бермейді. Өте баяу жылдамдықта және бинаризация сапасына қатты әсер етеді. Тіпті идеалды жағдайларда басқа әдістермен айналысуды жөн деп ойлаймыз.

Сүзгілердің тағы бір жеке тобы шекаралық және контурлық сүзу. Жолдардағы кескінмен жұмыс істеуден сол кескіндегі объектімен жұмыс істеуге ауысқымыз келгенде өте тиімді. Егер объект өте күрделі, бірақ жақсы ажыратылған болса, ол кезде жұмыс істеудің жалғыз әдісі көбінесе оның шекараларын таңдау болады. Шекараны сүзу

мәселесін шешетін бірқатар алгоритмдер бар. Олар: Оператор Canny, С.Фельдман операторы, Laplace операторы, Prewitt операторы, Mathieu Roberts операторы. Ең жиі қолданылатын Canny, ол жақсы жұмыс істейді және жүзеге асырылғандығы көп жағдайда көрсетілген. Мысалы, OpenCV (Open Source Computer Vision Library) кітапханасынан таба аламыз[4].

Жоғарыда көрсетілген сүзгілер модификациялары тапсырмалардың көп бөлігін шешуге көмектеседі. Дегенімен басқа да ауқымды мәселелерді шешкенде кездесетін және аз қарастырылатын сүзгілер жеткілікті. Оларға жататын итеративті сүзгілер (айқын бейне түрлерін анықтайтын), сонымен қоса Rf (Radon) түрлендіру аумағында кездесетін қарапайым толқындық сүзгілеу, одан қалса өңдеу қорытпасы дейтін риджлет және қисық түрлендірулер қызығушылық тудырады. Beamlet трансформациясы шекараларды бөлектеуге мүмкіндік беретін толқындық түрлендіру және логикалық талдау шекарасында тамаша жұмыс істейді.

## **2. Филтрлеу нәтижелерін логикалық өңдеу.**

Екінші бөлім ретінде сүзу нәтижелерін логикалық өңдеуді қарастырайық.

Сүзу өңдеуге керек деректер жинағын береді. Дегенімен бұл деректерді талдамай қолдана беруге болмайды. Бұл бөлімде кескіннен объектілердің қасиеттеріне немесе объектілердің өзіне өтуге мүмкіндік беретін бірнеше классикалық әдістер болады.

Сүзгілеуден логикаға айналдыру математикалық морфология әдістері деп айтсақ болады. Шын мәнінде, бұл екілік кескіндерді үлкейту және эрозиялаудың ең қарапайым операциялары. Бұл әдістер қол жетімді элементтерді ұлғайту әлде кеміту арқылы екілік кескіннен ұажет емес аумақты алып тастауға қолайлы жағдай тудырады. Математикалық тұрғыда шекаралау алгоритмдері көптеп кездеседі, дегенімен олар гибриді алгоритмдердің әлде басқа да бір алгоритмдердің түрлерін қосып пайдаланады [5].

Келесі бір талдау әдісі шекаралық талдау. Филтрлеу бөлімінде шекараларды алу алгоритмдері қарастырылған. Анықталған шекаралар жай ғана контурға айналады. Кенни алгоритмінде бұл автоматты қалыпта орын алады, ал кез келген басқа алгоритмдер үшін қосымша бинарлауды талап етеді. Бинарлау үшін шекараны белгілеп, соны қарастыруға болады, мысалы, қоңыз (алгоритм Жука) алгоритмін пайдалану арқылы. Шекара - бұл нысанның айрықша бейнесі. Көбінесе бұл шекара шетімен нысананы анықтауға болады. Осы анықтауды іске асыруға мүмкіншілік тудыратын жақсы математикалық құрылғы бар. Құрылғы шекаралық талдау деп аталады. Контурлық талдаудың үлкен кемшілігі тым идеалды жағдайлар қажет. Себебі нақты шекара әр кезде бола бермейді, немесе керек емес заттардың көп болуы. Дегенімен, идеалды кезде бір затты анықтау керек болса, онда шекаралық талдау керемет нәтиже береді. Ол тез жұмыс істейді, математикасы мен логикасы түсінікті.

## **Нәтиже**

### **3. Шешім қабылдау.**

Үшінші бөлімде кескінмен тікелей жұмыс жасаймыз, бірақ шешім қабылдауға мүмкіндік тудыратын әдістерді қарастырамыз. Жалпы бұл машиналық оқытудың және шешім қабылдаудың әртүрлі әдістері.

Көп кезде тану мәселесінде оқытудың мәні осылай болады: Объектілердің бірнеше кластары бар және сынақ үлгісі бар. Бұл суреттегі нысанның болуы немесе болмауы болсын деп алсақ. Әрбір кескін үшін Haar, HOG, SURF немесе қандай да бір волвлет (Wavelet transform) болсын, кейбір мүмкіндіктермен ерекшеленген мүмкіндіктер тобы бар. Оқыту алгоритмі осындай модельді құруы керек, сонымен қатар ол жаңа кескінді өңдейді және кескінде қандай нысандардың тұрғанын анықтай алады [5].

Қалай жасалыну логикасын теория жүзінде қарастырсақ. Сынақ кескіндерінің әрқайсысы функция кеңістігіндегі нүкте болады. Оның координаттары кескіндегі әрбір мүмкіндіктің салмағы болады. Бастысы, классификатордың атқаратын қызметі нысандардың классификациясына сай келетін белгілер кеңістігінде аумақтарды салады. Көптеген классификаторлар бар. Олардың әрқайсысы өзінің кейбір тапсырмаларында жақсырақ



жұмыс істейді. Белгілі бір тапсырма үшін классификаторды таңдаудың өзі қосымша білімді талап етеді. Оны келесі мақалаларда қарастыруға болады.

### **Қорытынды**

Қорыта кетсек, бұл мақалада біз ең басты үлгіні танудың негізгі логикасын қарастырдық. Бұл жайында ақпаратта жеткілікті, дегенімен жалпы шешімдерді қорытып шығарған мәлімет тапшы. Жоғарыда көрсетіліп кеткен ұсыныстар теория жүзінде көрсетілген, себебі зерттеу жалғасуда. Келесі мақалаларда әр алгоритмге жеке көңіл бөліп, артықшылығы мен кемшіліктеріне жеке тоқталуға болатындығын қарастыруға бел буамыз.

### Қолданылған әдебиеттер тізімі:

1. Материал из Национальной библиотеки им. Н. Э. Баумана. Распознавание образов.
2. Пару слов о распознавании образов <https://habr.com/ru/post/208090/>
3. Распознавание автомобильных номеров <https://sprecord.ru/files/manual/ipvr/3.4.html>
4. Гонсалес Р., Вудс Р. Цифровая обработка изображений (3-е изд., 2012) Москва: Техносфера, 2012.
5. Яне Б. Цифровая обработка изображений. Москва: Техносфера, 2007 г.

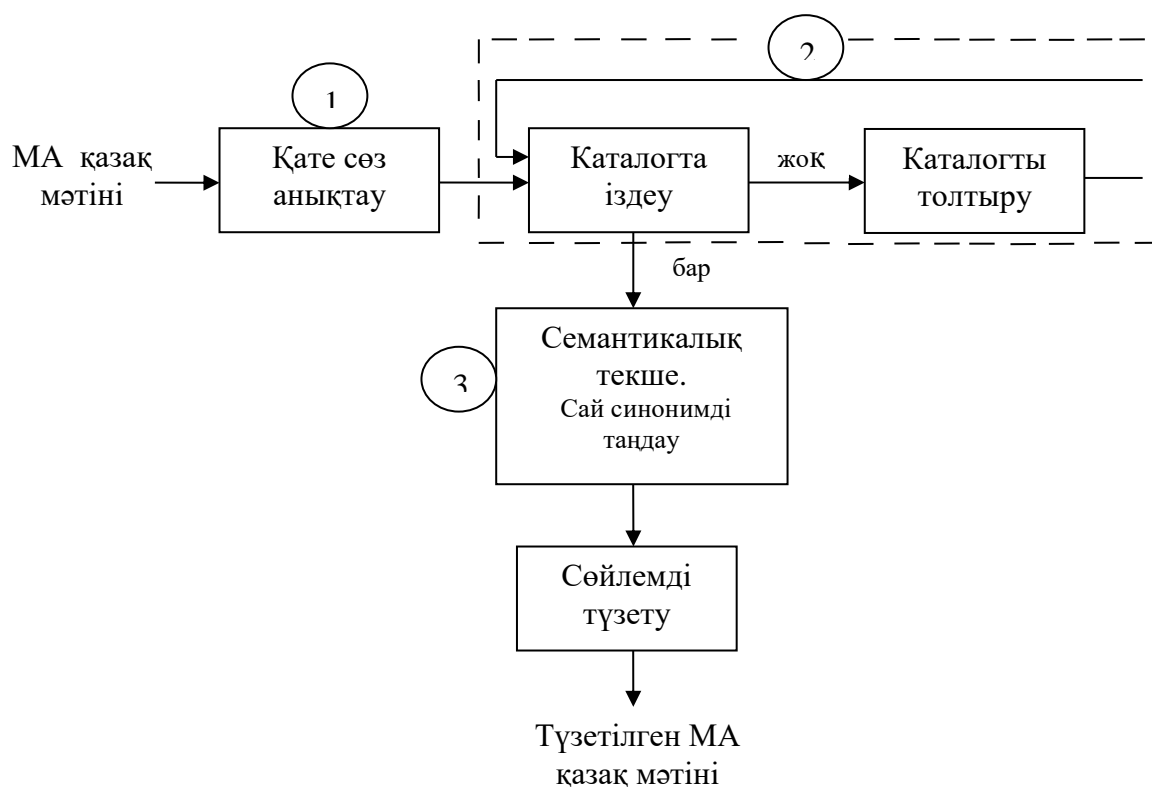
## АҒЫЛШЫН-ҚАЗАҚ МАШИНАЛЫҚ АУДАРМА ПОСТ-РЕДАКЦИЯЛАУ ТЕХНОЛОГИЯСЫ

Шормакова А.Н.,Тукеев У.А.

Әл-Фараби атындағы ҚазҰУ, Алматы қаласы, [shormakovaassem@gmail.com](mailto:shormakovaassem@gmail.com)

Машиналық аударма жасанды интеллекттің алдыңғы қатарлы бағыттарының бірі болып табылады. Машиналық аударманың сапасы жылдан-жылға артып келеді, бірақ кәсіби аударма сапасына әлі жеткен жоқ. Машиналық аударманың сапасын жақсарту үшін маңызды және практикалық әдістердің бірі – пост-редакциялау әдісі. Машиналық аударманы қолмен пост-редакциялау біршама еңбекті қажет ететін үрдіс. Сондықтан да автоматты пост-редакциялау маңызды [1-2]. Машиналық аударманың сапасы, тіпті нейромашиналық аударманың жетістіктерімен де кәсіби аударманың сапасына әлі жеткен жоқ. Осыған орай, Баоның (Бао 2015) [3] еңбегінде лексикалық қателердің қанша түрі анықталғанын көруге болады. Жұмысы барысында корпустан қате іздеу кезінде 633 лексикалық қате, оның ішінде сөздердің орын тәртібінің 17 қатесі, 116 сөз табының қатесі, ауыстырудың 209 қатесі, 96 эллипс қатесі, 100 артық қате, 3 қайталанатын қате, 92 семантикалық көпмағыналық қателер анықталды. Ол жиілігі жоғары (ең көп кездесетін) лексикалық қателердің 4 түрін ғана зерттеген. Ал біздің жұмыста біз ауыстыру қателігін қарастырамыз.

Пост-редакциялау әдісін қолдана отырып, біз келесі блок-схемада (1-сурет) көрсетілген технологияға (PE-LC Post-Edit Lexical Choice) қол жеткіздік.



1-сурет. Пост-редакциялау технологиясының блок-схемасы

Негізі жұмыстың мақсаты ағылшын тілінен аударылған қазақ тіліндегі дұрыс аударылмаған сөзді тауып, оның синонимдер каталогынан тауып, дұрыс түзетілген сөйлем шығару болып табылады. Қазақ тіліндегі қате сөздерді табу үшін Еспла тобы бастаған испан тобы [4-5] жұмыстары қолданылды. Яғни, бірінші тапсырмада машиналық аудармадағы қазақша сөйлемдегі қате аударылған сөздерді анықтау үшін Esplà-Gomis және т.б. (2012),(2015) еңбектері қолданылып, аударма жадының орнына кері аудару

әдісін қолдану арқылы жақсартылды. Содан кейін қазақ тіліндегі дұрыс емес аударылған сөздерді автоматты түрде тапқан соң, осы табылған сөздерден автоматты түрде синонимдер каталогы құрылды. Синонимдер каталогын құру үшін онлайн сөздік қолданылды (thesaurus.com). Онлайн сөздіктен ағылшын сөздері қазақшаға аударылып, қазақша синонимдер сөздігіне жазылды[6-7]. Нәтижесінде автоматты синонимдер каталогы құрылды.

Анықталған дұрыс емес сөздің орнына дұрыс нұсқасын қою үшін синонимдер каталогын қолданып, синонимдер тізімінен мағынасы жағынан ең тиімді сөзді таңдалынып толық түзетілген сөйлем шығарылды. Ол үшін максимум энтропия әдісі негізінде лексикалқ таңдау әдісі қолданылды. Қазақ тілінің қосымшаларын дұрыс шығару мақсатында Апертиум платформасы қолданылды[8-9].

Статистикалық маңыздылық екі жүйе (Google мен PE-LC) үшін де есептелінді. МА бағалау көрсеткішіндегі жақсарту – байқалған жақсартудың  $p$  кездейсоқ нәтиже болу ықтималдығы (нөлдік гипотезаның ықтималдығы)  $p$  мәнінен аз болғанда статистикалық маңызды болып табылады. Мұнда экстремалды – сынақ статистикасының нөлдік гипотезадан ауытқу дәрежесі. BLEU (Papineni K. et al. 2002)[10] сенімділігі жұптастырылған жүктеу жолағын қайта үлгілеу (bootstrap resampling) арқылы есептелінді. Содан кейін Google және PE-LC жүйелері арасында статистикалық маңыздылық тексерілді. 1-кестеде  $n$ -грамм сөздері үшін BLEU, TER және chrF2(статистикалық маңыздылық тест көрсеткіші) нәтижесі алынды. Google үшін PE-LC үлгісінде статистикалық маңыздылық тесті жасалынды. Нәтижелер келесі 1-кестеде көрсетілген.

Кесте 1 – Google , PE-LC үлгісі үшін статистикалық маңыздылық көрсеткіш кестесі

BLEU ( $\mu \pm 95\% CI$ )	TER ( $\mu \pm 95\% CI$ )	chrF2 ( $\mu \pm 95\% CI$ )
Google 41.0103 (40.9405 $\pm$ 6.0153)	Google 40.6349 (40.6292 $\pm$ 4.9062)	Google 70.3059 (70.2557 $\pm$ 3.4295)
PE-LC model 47.4809 (47.4171 $\pm$ 5.8691) <b>(<math>p = 0.0010</math>)*</b>	PE-LC model 35.9788 (35.9826 $\pm$ 4.8104) <b>(<math>p = 0.0010</math>)*</b>	PE-LC model 75.6458 (75.6220 $\pm$ 2.9849) <b>(<math>p = 0.0010</math>)*</b>

МА мен сөйлемнің пост-редакциялау сапасы TER, WER, BLEU жүйесі арқылы анықталды (Papineni K. et al. 2002). Салыстыру үшін Google аударма жүйесі де қолданылды. 2-кестеге сәйкес, қазақ тіліндегі ішінара пост-редакцияланған мәтін мен Google аударған мәтін арасындағы BLEU метрикасының айырмашылығы 6,47 пайызды құрады. Бұл деректер 2022 жылдың 18 ақпанында жиналды. Зерттеу нәтижелері жақсы көрсеткіштерді көрсетті.

Кесте 2 – Бағалау көрсеткіш нәтижелері

Машиналық Аударма құралы	Машиналық аударманы бағалау, BLEU	Машиналық аударманы бағалау, TER	Машиналық аударма сапасын бағалау (weighted average WER)
Google Translate	41.01%	40.63%	29.59%
PE-LC system (Google-ға қолданылған)	47.48%	35.97%	25.31%

Сондай-ақ аударма қатесінің деңгейі (TER) метрикасын қолданылды, бұл әдісті МА мамандары пост-редакциялаудың қажеттілік көлемін анықтау үшін қолданады. Ал ұсынылған PE-LC жүйесінен кейін Google Translate жүйесі PE-LC жүйесінен 4,66 пайызға артық көрсеткішті көрсетті. TER көрсеткішінің азайуы - бұл жақсырақ дегенді білдіреді; PE-LC жүйесінде TER бар, ол 4,66 пайызға төмен.

WER сөз қатесі деңгейін есептеу үшін дұрыс мәтін мен Google шығарған мәтін қарастырылды. 1-2 кестедегі нәтижелерді алу үшін 100 сөйлем WER, TER және BLEU метрика мәнін Sacrebleu онлайн бағдарламасы қолданылып, мәндері есептелінді. Жүйелердің пайыздық көрсеткіштерін есептеу үшін синонимдері бар сынақ корпусындағы сөйлемдерде лексикалық таңдау мәселесінің қаншалықты шешілгенін көру үшін пайдаланылды. Google-дың орташа өлшенген WER көрсеткіші нәтижесі 29,59 пайызды құрады. WER көрсеткіші ағылшын тілінен қазақ тіліне аударылған мәтінге арналған PE-LC технологиясы Google Translate аудармасының мәтінін 4,28 пайызға жақсартқанын көрсетті.

#### Әдебиеттер тізімі

1. Абеустанова А.Н. "Машиналық аударманың нарықтағы және Қазақстандағы күйі". *ҚазҰТУ хабаршысы* № 6(106), 2014. –150-152 б.
2. Shormakova A. "Machine translation and post-editing". *Материалы международной конференции студентов и молодых ученых «Мир науки»*, 17-19 апреля 2013г. – Алматы: Қазақ университеті, 2013. – стр. 222
3. Bao, X. L. (2015). Analysis on Lexical Errors in College English Writing. *Canadian Social Science*, 11(12), 127-130. Available from: <http://www.cscanada.net/index.php/css/article/view/7775> DOI: <http://dx.doi.org/10.3968/7775>
4. Esplà-Gomis, M., Sánchez-Martínez, F., Forcada, M.L. (2012), "A Simple Approach to Use Bilingual Information Sources for Word Alignment", *Procesamiento del Lenguaje Natural*, 49, 93-100
5. Esplà-Gomis, M., Sánchez-Martínez, F., Forcada, M.L. (2015), "Using Machine Translation to Provide Target-Language Edit Hints in Computer Aided Translation Based on Translation Memories", *Journal of Artificial Intelligence Research*, 53, 169-222
6. Абеустанова А.Н. "Қазақ тіліндегі көпмағыналы сөздердің бірін анықтаудың бір болжамы". *ҚазҰТУ хабаршысы* №4(110) 2015. –625-628 б.
7. Абеустанова А.Н. "Ағылшын тілінен қазақ тіліне аударылған қазақша қате сөздерді анықтау және баламалар каталогын құру". *ҚазҰТУ хабаршысы* №6 2017. –313-317 б.
8. Abeustanova A., Tukeyev U. "Automatic Post-editing of Kazakh Sentences Machine Translated from English". *Studies in Computational Intelligence/Advanced Topics in Intelligent Information and Database Systems*, vol. 710 – Springer International Publishing, 2017. – pages. 283-295.
9. Mikel L. Forcada, Mireia Ginestí-Rosell, Jacob Nordfalk, Jim O'Regan, Sergio Ortiz-Rojas, Juan Antonio Pérez-Ortiz, Felipe Sánchez-Martínez, Gema Ramírez-Sánchez, Francis M. Tyers, "Apertium: a free/open-source platform for rule-based machine translation", *Machine Translation*, (Special Issue on Free/Open-Source Machine Translation) 25:2, 127-144
10. Papineni K, Roukos S, Ward T, Zhu WJ (2002). BLEU: a method for automatic evaluation of machine translation. In: *Proceedings of the 40th annual meeting of association for computational linguistics (ACL 2002)*, Association for Computational Linguistics, Philadelphia, Pennsylvania, USA, pp 311–318.

# ПРОГРАММНАЯ РЕАЛИЗАЦИЯ КОМБИНАЦИИ АДАПТИВНЫХ МЕТОДОВ ПОВЫШЕНИЯ КОНТРАСТА ДЛЯ УЛУЧШЕНИЯ РЕНТГЕНОВСКИХ ИЗОБРАЖЕНИЙ

*Г.С.Омарова, Е.Тұрсынбай*

*Евразийский национальный университет имени Л. Н. Гумилева, г.Астана*

*\*e-mail: ogs12@mail.ru*

Идея методов улучшения изображений заключается в том, чтобы проявить детали объектов, которые скрыты, или просто выделить определенные особенности изображения. Один из примеров улучшения — повышение контраста изображения путем растяжения его динамического диапазона значений яркости. Термин "контраст", наблюдаемый на цифровых изображениях описывается отношением яркости темных и светлых областей, присутствующих на изображении [1,2].

Целью данной статьи является исследование методов повышения контраста изображений и определение их возможностей для улучшения малоконтрастных рентгеновских снимков. Для этого оцениваются адаптивные методы повышения контраста, выполняется выбор их параметров и разрабатывается методика их применения. Для выбора метода повышения контраста, а также его параметров, необходима оценка изображения. Проведение эксперимента опирается на алгоритмы объективной неререференсной оценки качества NIQE[3] и BRISQUE[4]. Эксперименты используют предположение того, что данные объективные оценки уменьшаются в значении при визуальном увеличении контраста.

Для достижения поставленной цели были поставлены следующие задачи:

- проанализировать основные этапы гамма-коррекции и CLAHE для обработки рентгеновских медицинских изображений;
- разработать методику использования адаптивных методов контрастного усиления.

Для выполнения экспериментов по применению методов преобразования яркости изображения использовались несколько десятков рентгеновских снимков из базы Kaggle [5]. Целью экспериментов является поиск метода повышения контраста рентгеновских изображений легких для их более информативного представления. Сущность методов повышения качества медицинских изображений состоит в следующем: применить к изображениям малого контраста математические методы и улучшить контраст цифрового медицинского изображения для повышения точности диагностики.

Проведен ряд экспериментов применения функции преобразования яркости полутоновых изображений к нескольким рентгеновским снимкам для выбора наиболее подходящих входных параметров[6]. Эксперименты проводились в пакете MATLAB.

В ходе экспериментов были обработаны светлые, темные и нормальные рентгеновские снимки. Применение метода объективной оценки к обработанным изображениям показало следующие результаты. В результате исследования вариантов преобразования тестовых изображений, для получения рентгеновских снимков с максимальным контрастом рекомендуется:

- 1) построить гистограмму изображения и определить его общий уровень яркости (рис.1);

На Рис.1. показано начальное окно приложения, где выполняется загрузка изображения и определение ее гистограммы. Далее выполняется выравнивание гистограммы изображения с помощью функции `histeq`. Также вычисляются неререференсные оценки NIQE для исходного и преобразованного изображений.

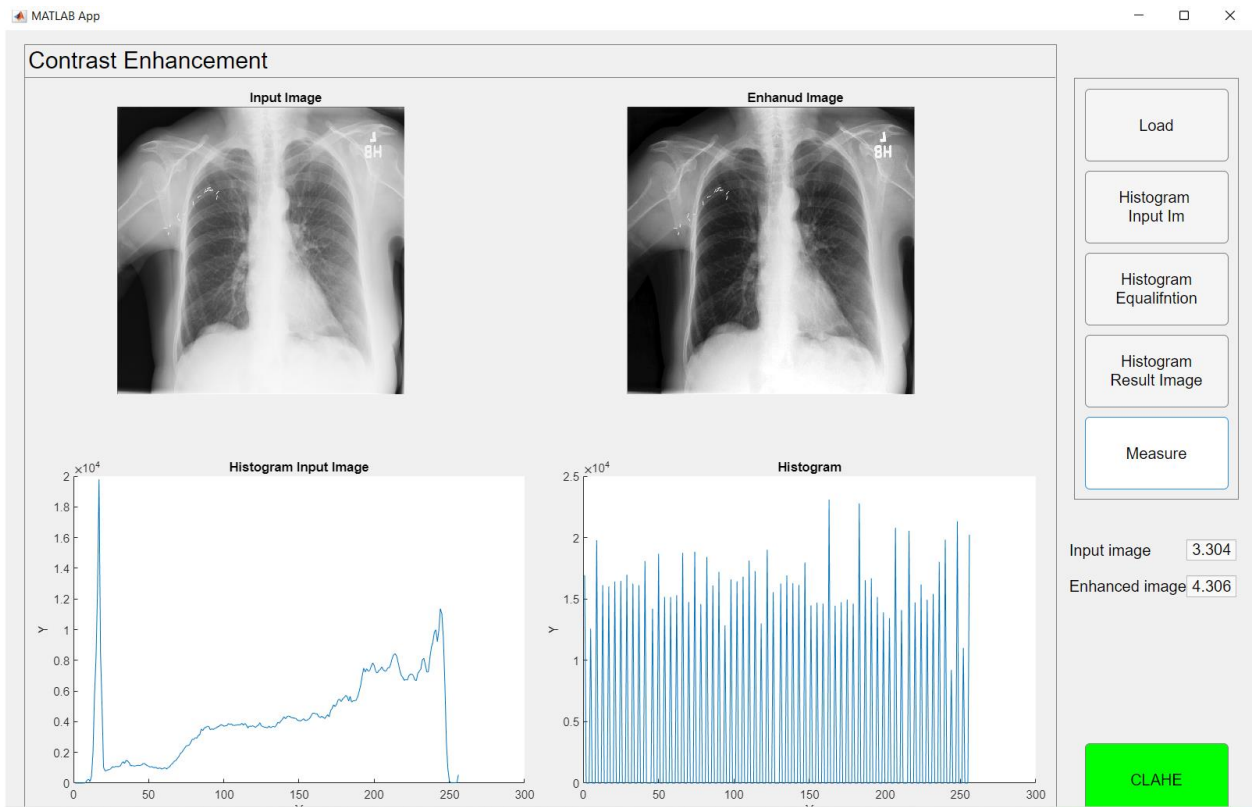


Рисунок 1. Начальное окно приложения. Загрузка изображения и определение гистограммы изображения.

- 2) применить процедуру адаптивного выравнивания гистограммы изображения с ограничением контрастности (рис.2), значением параметра `distribution` выбрать 'exponential' и значения параметра `cliplimit` выбирать с интервала  $[0, 0.02]$  с шагом 0.01;
  - 3) все преобразованные изображения оценить нерепренской оценкой NIQE и определить изображение, соответствующее минимальной оценке NIQE;
- На рис.2 изображено продолжение работы приложения. Здесь можно выполнить подбор параметра `cliplimit`, также можно вычислить значение этого параметра для изображения с минимальной оценкой NIQE.

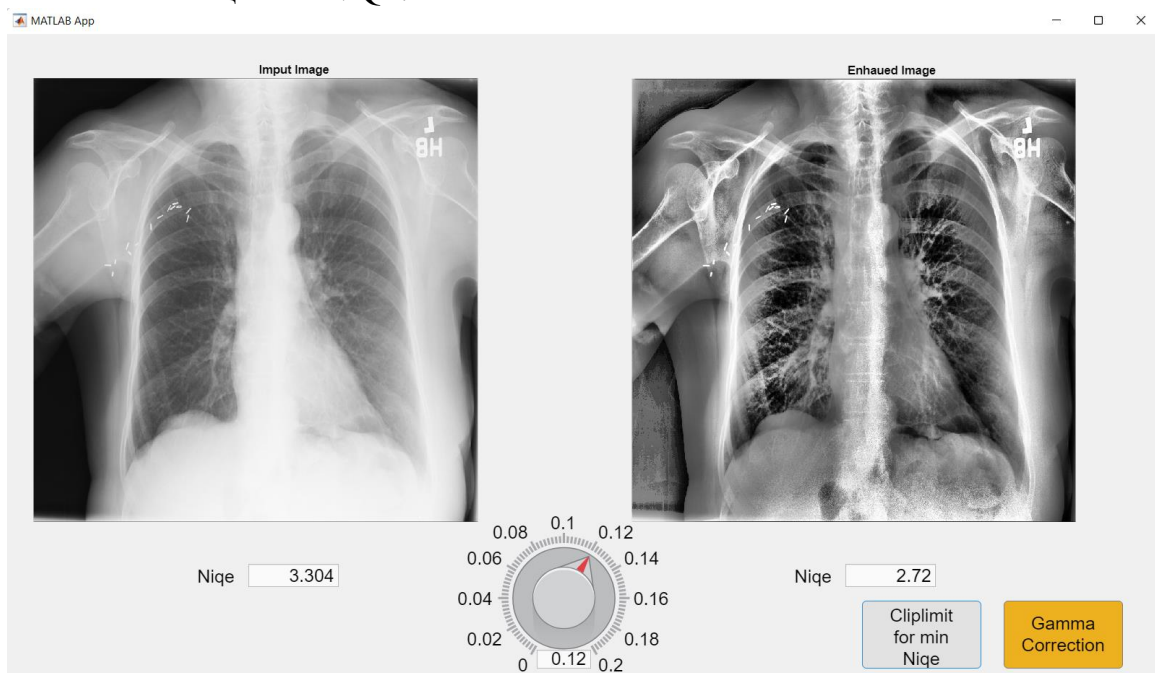


Рисунок 2. Реализация метода CLAHE.

- 4) после применения метода CLAHE[7] применить функцию `imadjust`:
- в случае если исходное изображение  $I$  содержит больше светлых оттенков, то входные параметры для функции `imadjust` выбрать в следующем виде:  
 $J = \text{imadjust}(I, [\text{low\_in}, 1], [0, 1], \gamma)$ , где  $0.4 \leq \text{low\_in} \leq 0.7$ ,  $0 < \gamma \leq 3$   
 дают более лучшие результаты;
  - в случае если исходное изображение  $I$  содержит больше темных оттенков, то входные параметры для функции `imadjust` выбрать в следующем виде:
  - $J = \text{imadjust}(I, [0, \text{high\_in}], [0, 1], \gamma)$ , где  $0.4 \leq \text{high\_in} \leq 0.7$ ,  $0 < \gamma \leq 3$

На рис.3 показано выполнение метода гамма-коррекции для рентгеновского изображения, которое было преобразовано методом адаптивного выравнивания с ограничением контраста. Это окно вызывается со второго окна приложения, где выполняется метод CLAHE. Здесь выполняется подбор значений таких параметров, как входные диапазоны (значения `low_in`, `high_in`) для изображения, преобразуемого функцией `imadjust`, и значение параметра  $\gamma$ . При нажатии по соответствующей кнопке выполняется гамма-коррекция изображение с выбранными параметрами.

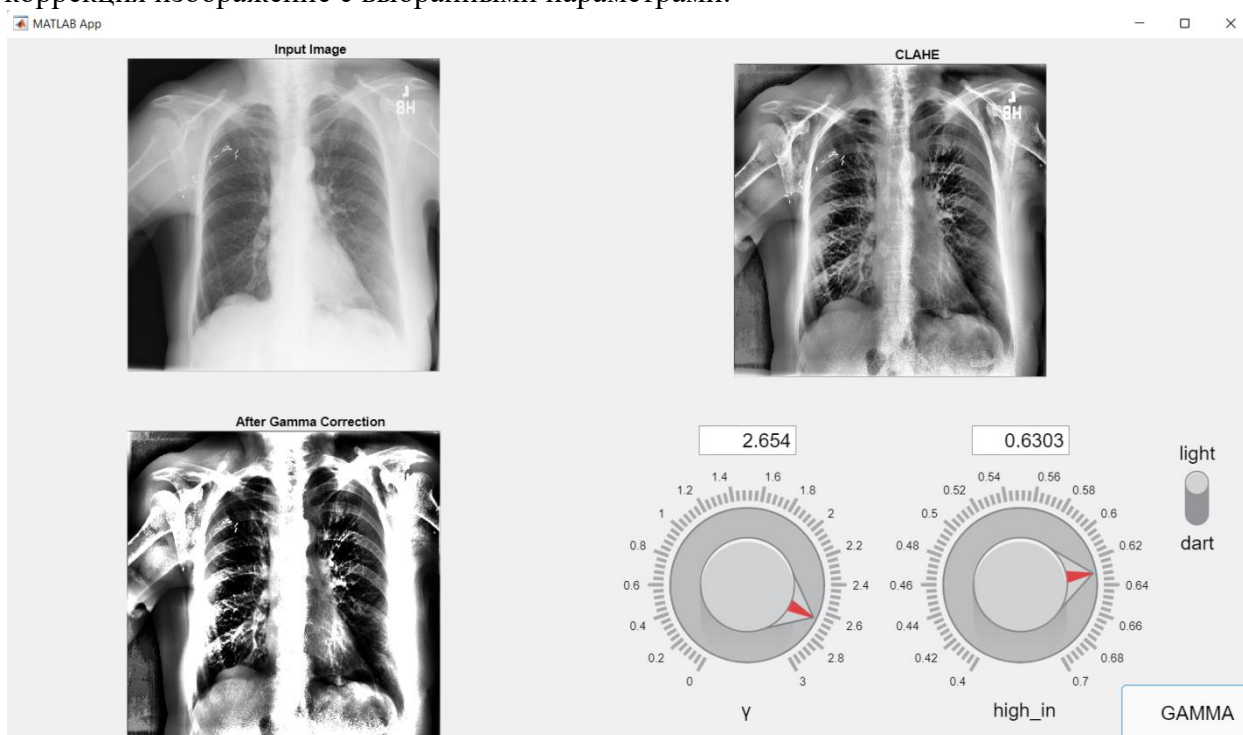


Рисунок 3. Окно реализации метода гамма-коррекции изображения.

В результате выполненных исследований показано, что целесообразно использовать комбинацию метода гамма-коррекции с методом адаптивного выравнивания гистограммы, в котором усиление контраста ограничено во избежание возникновения или усиления шумов на изображении.

В исследовании выполнен анализ возможностей методов гамма-коррекции [8] и CLAHE для повышения контраста рентгеновских изображений. В ходе выполнения экспериментов выполнен подбор значений необходимых параметров, при которых субъективные и объективные оценки одинаково показали положительный результат улучшения качества рентгеновских изображений. Эксперименты доказали целесообразность применения комбинации метода гамма-коррекции с адаптивным выравниванием гистограммы с ограничением контраста.

В результате выполненных экспериментальных исследований сформулирована методика применения комбинации метода гамма-коррекции с адаптивным выравниванием гистограммы с ограничением контраста[9]. Данная методика предусматривает выполнение повышения контраста рентгеновских изображений в два этапа. На первом

этапе исходное изображение преобразовывается методом CLAHE с выбранными параметрами, второй этап улучшает полученное изображение методом гамма-коррекции. Экспериментальные результаты показали, что предлагаемая методика позволяет получить рентгеновские изображения с усиленным контрастом.

#### Список литературы

- 1 Гонсалес Р., Вудс Р. Цифровая обработка изображений.- Издание 3-е, исправленное и дополненное. –М.: Техносфера, 2012. – 1104 с.
- 2 Гонсалес Р., Вудс Р., Эддинс С. Цифровая обработка изображений в среде Matlab. – М.: Техносфера, 2006.-616 с.
- 3 Mittal, A., R. Soundararajan, and A. C. Bovik. "Making a Completely Blind Image Quality Analyzer." IEEE Signal Processing Letters. Vol. 22, Number 3, March 2013, pp. 209–212.
- 4 Mittal, A., A. K. Moorthy, and A. C. Bovik. "No-Reference Image Quality Assessment in the Spatial Domain." IEEE Transactions on Image Processing. Vol. 21, Number 12, December 2012, pp. 4695–4708.
- 5 <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>.
- 6 Омарова Г., Старовойтов В. Увеличение контраста рентгеновских изображений на основе гамма-коррекции. «Физико-математические науки». 77, 1 (апр. 2022). DOI:<https://doi.org/10.51889/2022-1.1728-7901.32>.
- 7 Ma J., Fan X., Yang S.X., Zhang X., Zhu X. Contrast Limited Adaptive Histogram Equalization Based Fusion for Underwater Image Enhancement // Preprints [Электронный ресурс] 2017, URL: <https://www.preprints.org/manuscript/201703.0086/v1>
- 8 Omarova G. S, Starovoitov V. V, Aitkozha Zh. Zh, Bekbolatov S, Ostayeva A. B and Nuridinov O, “Application of the Clahe Method Contrast Enhancement of X-Ray Images”(2022), International Journal of Advanced Computer Science and Applications(IJACSA), 13(5),. <http://dx.doi.org/10.14569/IJACSA.2022.0130549>
- 9 Omarova, G., Aitkozha, Z., Sadirmekova, Z., Zhidekulova, G., Kazimova, D., Muratkhan, R., Takuadina, A., & Abdykeshova, D.. “Devising a methodology for X-ray image contrast enhancement by combining CLAHE and gamma correction” (2022), Eastern-European Journal of Enterprise Technologies, 3(2 (117), 18–29. <https://doi.org/10.15587/1729-4061.2022.258092>



## NAMING GAME МОДЕЛІ НЕГІЗІНДЕ ПІКІР АЛМАСУ

<sup>1</sup>Мурзахметов А.Н., <sup>2</sup>Бапанов А.А., <sup>3</sup>Бейсов Н.Қ., <sup>4</sup>Тасжуреков Ж.Қ.

<sup>1,2,4</sup> М.Х.Дулати атындағы Тараз өңірлік университеті

<sup>3</sup>Әл-Фараби атындағы Қазақ ұлттық университеті

<sup>1</sup>email: [aslanmurzakhmet@gmail.com](mailto:aslanmurzakhmet@gmail.com); <sup>2</sup>email: [arseniy.bapanov@gmail.com](mailto:arseniy.bapanov@gmail.com)

<sup>3</sup>email: [beisov.nurbol@gmail.com](mailto:beisov.nurbol@gmail.com); <sup>4</sup>email: [tashhurecov.88@gmail.com](mailto:tashhurecov.88@gmail.com)

**Аңдатпа.** Жұмыста Naming Game моделі негізінде қоғамда пікір алмасу және онымен келісу немесе келіспеу үрдістері сипатталған. Бұдан басқа, біз келісімге жақындау үшін желілік стратегияларды қарастырамыз.

**Кілттік сөздер:** Naming Game, әлеуметтік қоғам, ақпараттық жүйелер, инновация, ақпараттың таралуы.

**Кіріспе.** Әлеуметтік қауіпсіздікті анықтайтын процестердің динамикасын сипаттау үшін әлеуметтік жүйедегі тиісті процестердің таралу жылдамдығын ескеру қажет. Әлеуметтік жүйеде, соның ішінде әлеуметтік қауіпсіздікті анықтайтын процестердің кейбірін инновация ретінде сипаттауға болады. Жалпы, инновация дегеніміз – жаңашылдықты, идеяны, тәжірибе немесе объекті сезінетін жеке тұлға немесе топ» [1]. Жаңашылдық дегеніміз - белгілі бір әлеуметтік пән үшін жаңа идеялар, өнімдер, шешімдер, технологиялар және т.б. Диффузия - бұл инновацияның уақыт өте келе әлеуметтік жүйенің мүшелері арасында белгілі бір арналар арқылы таралу процесі. Сонымен қатар, диссертациялық жұмыста ақпарат ретінде, инновация, идея және пікір сияқты түніктерді білдіреді [2-3].

Инновация диффузиясының теориясы әлеуметтік жүйеде инновациялардың таралуын қарастырады. Осы саланың зерттеушілері [4-6] әлеуметтік жүйенің мүшелері жаңашылдықты қандай ықтималдықпен қабылдайтындығын және оның таралу жылдамдығын түсіндіруге тырысуда. Инновацияны әлеуметтік жүйеге инноваторлар енгізеді, содан кейін біртіндеп инновация туралы ақпаратты бір-біріне беретін көптеген агенттер қабылдайды, осыған байланысты әлеуметтік жүйенің табиғаты маңызды рөл атқарады [7]. Инновацияның таралуына көптеген факторлар әсер етеді: агент сипаттамалары, инновациялық сипаттамалар және әлеуметтік жүйенің сипаты [8-9]. Диффузияны қарапайым түрде зерттеу - бұл әлеуметтік жүйе агенттерінің инновацияны қабылдауға әсер ететін осы және басқа факторлардың өзара әрекеттесуін зерттеу.

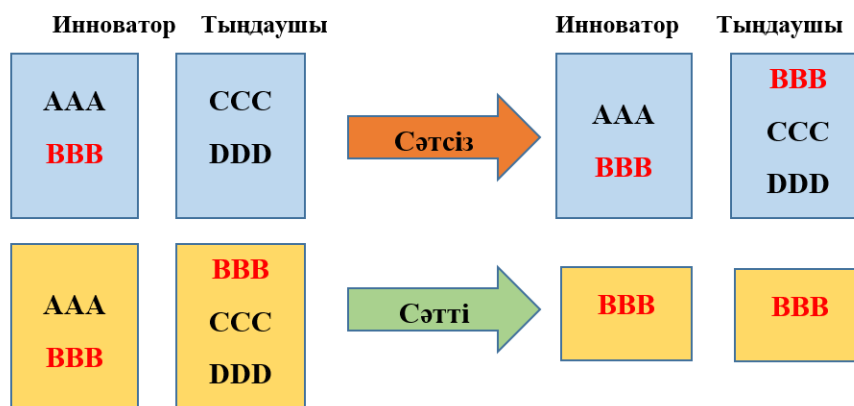
Интернеттің қарқынды дамуымен біз түрлі құрылғылар арқылы ғаламдық желіге қосылып, адамдар бір біріне жақындай түсті. Түрлі қосымшалар арқылы адамдар жылдам ақпарат алмасып, түрлі мәселелерді талқылай алады, олармен бөлісе, қабылдай алады. Кейбіреулер, мәселені талқыға салып, өз ойын ортаға салады, біреулер олармен келісіп, енді біреулер қарсылық білдіріп жатады. Мұндай сан-алуан ақпараттармен бірге инновациялар да таралады. Мәселен, Bitcoin, Selfie, MOOC сияқты жаңа сөздер мен қысқартулар пайда болды. Бұлар өз кезегінде әлеуметтік қоғамдағы инновациялардың таралу үрдістерін зерттеуге әртүрлі ғылыми қауымдастықтардың қызығушылықтарын арттырды. Бұл үрдістер биологияда [10], әлеуметтік ғылымда [11], пікірлер немесе келіссөздердің таралуы [12-16], жасанды интеллектте [17] әртүрлі салаларда [18, 19] қоғамды қалыптастыру сияқты жұмыстарда мұқият зерттелген және пайдаланылған. Кешенді желілер мен әлеуметтік динамиканы біріктіре отырып, Naming Game (NG) математикалық модельдеу және бәтуаға келу процестерін модельдеудің тиімді тәсілін ұсынады. Атап айтқанда, NG – бұл кейбір атаулар туралы келісімге келуге тырысатын агенттердің өзара әрекеттесуін зерттейтін модель. Соңғы уақытта NG кездейсоқ графикалық желілер [20-22], кішігірім желілер [23, 24] және әртүрлі күрделі желілік модельдерде [25, 26] зерттелді. Naming Game моделі мүмкіндігінше аз есептеу қуатын пайдаланады және осылайша танымдық күрделілік пен өнімділік үшін төменгі сызықты белгілейді. Тілдік өзін-өзі ұйымдастырудың басқа модельдерінен айырмашылығы,

агенттер жеке сөздердің сәттілік дәрежесі туралы ақпаратты сақтамайды және ең жақсы сөз немесе крест-ситуациялық оқуды таңдау сияқты зияткерлік эвристиканы пайдаланбайды.

Naming Game моделінің ережесі бойынша N агенттері қатысады және олар өз сөздігіне M сөздерді жүктеуге тырысады, осылайша бір агент басқа біреудің назарын аудара алады. Сөздер – бұл адамдар, физикалық нысандар, қатынастар, веб-сайттар, суреттер, музыка файлдары сияқты кез-келген нысандар бола алады. Барлық агенттердің  $t=0$  уақытында бос сөздігі бар. Әр қадамда ( $t=1,2,\dots$ ) кездейсоқ екі ойыншы таңдалады, олардың біреуі баяндаушы, ал екіншісі тыңдаушы ретінде қарастырылады. Алғашында, егер сөздік бос болса, ол сөздіктен жаңа сөз құрастырады, содан кейін оны тыңдаушыға белгісіз заттың атауы ретінде береді. Бірақ егер баяндаушының жадында бірнеше сөздер болса, ол кездейсоқ тыңдаушыға айту үшін жадынан бір сөз таңдап алады. Егер берілетін сөз тыңдаушының жадында болса, онда екі агент те келісімге келеді де екі агент те осы сөзді сақтайды.

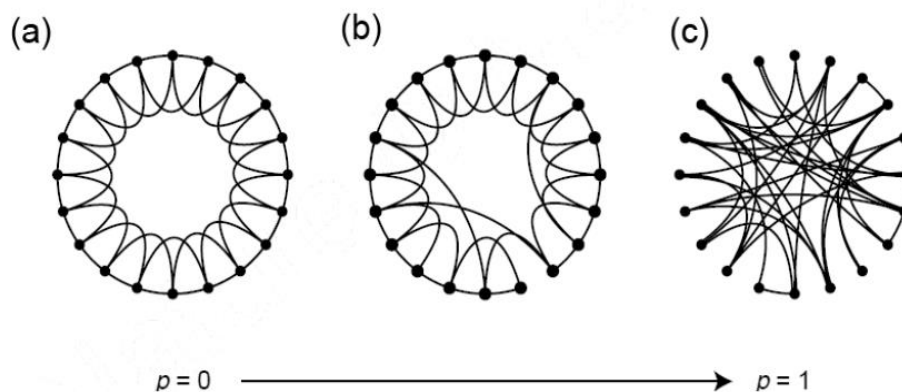
Кең көлемді автономды жүйелерде жаһандық үйлестірусіз келісімге қол жеткізу маңызды қызығушылық тудыратыны анық. Әлеуметтік жүйелер жағдайында осындай құбылыстарды түсіну және болжау басты мақсаттың бірі болып табылады. Осы мақсатқа жету үшін «Naming Game» моделін қолдануға болады.

**Зерттеу әдісі.** N көлемді агенттер қоғамын қарастырайық. «Naming Game» моделі бойынша қауымдастық мүшелері бір-бірімен ақпарат алмаса алады. Қоғамның әр мүшесінің пікірі бар, ол белгілі ережелерге сәйкес көршілерімен пікір алмасады. «Naming Game» моделінде , қауымдастықтың әр мүшесі (инноватор немесе агент) әркіммен байланыса алады және келісімге қол жеткізу үшін жұптасып әрекет етеді. «Naming Game» моделінің негізгі алгоритмдік ережелері: көрші түйіндер таңдалды, яғни, инноватор және агент. Инноватор өз пікірін айтады, егер агент инноватордың пікірімен келіссе, онда екі қатысушы да осы пікірді ұстанады, әйтпесе агент инноватордың пікірін елемейді және өз пікірінде қалады. Келесіде, қарапайымдылық үшін біз «AAA, BBB» т.с.с әріптерді пікірді білдіру үшін қолданамыз. Жоғарыда айтылған ережелер 1 суретте келтірілген.



Сурет 1. Naming Game модельінің ережесі

Алғашыда қауымдастықтың барлық мүшелерінің «сөздік қоры» шартты түрде бос болуы мүмкін. Әр қадам сайын өз тізімінен бір сөзді айтатын инноватор таңдалады. Егер агенттің «сөздік қорында» бұл термин болса, онда қалған терминдерді жойып, бұл терминді сөздікке қосады. Егер тізімде кем дегенде бір агентте осы термин кездессе, онда инноватор оны өз тізімінде қалдырады. Алда айтылғандай, бастапқыда сәтте сөздіктер бос болады. Содан кейін әр агенттің сөздігінде кемінде бір сөз пайда болғанға дейін сөздіктер кеңейе бастайды. Уақыт өте келе келіспеушіліктер жойылады, өйткені қауымдастық мүшелері бір пікірді ұстанатын болады.



Сурет 2. Кездейсоқ шыңды тандау және оны басқа шыңмен қосылуы

Әлеуметтік және ақпараттық желі ретінде кездейсоқ графты қолдануға болады:  $N$  шыңдары бар желі үшін екі шың  $p$  ықтималдығымен қосылады. Кездейсоқ графтарда қысқа жолдың орташа ұзындығы және шағын кластерлеу коэффициенті бар. Кездейсоқ графтардың желілік құрылымы көбінесе шыңдардың қосылу ықтималдығына байланысты. Егер қосылу ықтималдығы төмен болса, онда граф кішкентай окшауланған компоненттерге бөлінеді, ал жоғары болса, компоненттер түпнұсқалық графтың мөлшеріне сәйкес бір желіге біріктіріледі.

**Қорытынды.** Әлеуметтік жүйеде инновацияның таралуын зерттеу нәтижесі бойынша жаңа идея пайда болған кезде, қоғам оны бірдей қабылдамай, жаңа идеяны қабылдайтын және қабылдамайтын топтарға түсе алатындығын көрсетеді. Алғашқыда жаңа идеяны санаулы адамдар ғана қабылдайды, содан кейін инновацияны негізгі көпшілік қабылдайды, соңына келгенде, қабылдау қарқыны бәсеңдейді.

#### Әдебиеттер

1. Rogers E.M. Diffusion of Innovations, 5th ed.; Free Press: London, UK, 2003. 576 p.
2. Rogers E., Shoemaker F. Communication of Innovations: A Cross Cultural Approach, 2nd ed.; Free Press: London, UK, 1971, 476 p.
3. Lu Q., Korniss G., Szymanski B. K. Naming games in two-dimensional and small-world connected random geometric networks // Phys. Rev. E., vol. 77, Iss. 1, 2008.
4. Baronchelli A. Role of feedback and broadcasting in the naming game // Phys. Rev. E., vol.83, Iss.4, 2011.
5. Minayev V. A., Ovchinsky A. S., Skryl S. V., Trostyansky S. N. How to manage the mass consciousness: current models. Moscow, 2013, 200 p. (in Russ.)
6. Fedotov A.M., Murzakhmetov A.N., Dyusembaev A.E. Expansion of ideas and processes in social and biological communities // Eurasian journal of mathematical and computer applications, ISSN 2306–6172, vol. 6, Issue 4. pp. 7-28, 2018.
7. Ganglmair-Wooliscroft A., Wooliscroft B., Diffusion of innovation: The case of ethical tourism behavior // Journal of Business Research, vol.69, Iss.8, pp. 2711-2720, 2016.
8. Iglesias J.A., Tiemblo A., Ledezma A., Sanchis A. Web News Mining in an Evolving // Framework, Information Fusion, vol. 28. pp. 90-96, 2015.
9. Coletti C. F., Karin B.E., Rodriguez P. A stochastic two-stage innovation diffusion model on a lattice // Journal of Applied Probability, vol.53, Iss.4. pp. 1019-1030, 2016.
10. Fedotov A.M., Murzakhmetov A.N., Dyusembaev A.E. Expansion of ideas and processes in social and biological communities // Eurasian Journal of Mathematical and Computer Applications. – 2018. – Vol. 6, Issue 4. – P. 17-28.
11. Baronchelli A., Gong T., Puglisi A. et al. Modeling the emergence of universality in color naming patterns // Proc. Natl. Acad. Sci. – 2010. – Vol. 107. – P. 2403-2407.

12. Baronchelli A., Loreto V., Steels L. In-depth analysis of the naming game dynamics: The homogeneous mixing case // *Int. J. Mod. Phys.* – 2008. – Vol. C19. – P. 785-812.
13. Puglisi A., Baronchelli A., Loreto V. Cultural route to the emergence of linguistic categories // *Proc. Natl. Acad. Sci.* – 2008. – Vol. 105. – P. 7936-7940.
14. Murzakhmetov A.N., Infante Moro A., Dyusembayev A.E. Social media and networks as a tool to analysis of tourists preferences and requirements // *Vestnik KazNRTU.* – 2020. – Vol. 2(138). – P. 227-231.
15. Steels L. A self-organizing spatial vocabulary // *Artificial Life.* – 1995. – Vol. 2, Issue 3. – P. 319-332.
16. Baronchelli A., Felici M., Loreto V. et al. Sharp transition towards shared vocabularies in multi-agent systems // *J. Stat. Mech.* – 2006. – Vol. 2006. – P. 06014.
17. Lu Q., Korniss G., Szymanski B.K. The naming game in social networks: community formation and consensus engineering // *J. Economic Interaction Coordination.* – 2009. – Vol. 4, 2009. – P. 221-235.
18. Zhang W., Lim C.C. Noise in naming games, partial synchronization and community detection in social networks // <https://arxiv.org/pdf/1008.4115.pdf>. 22.02.2019.
19. Dall'Asta L., Baronchelli A., Barrat A. et al. Agreement dynamics on small-world networks // *Euro. Phys. Lett.* – 2006. – Vol. 73. – P. 969.
20. Liu R.R., Jia C.X., Yang H.X. et al. Naming game on small-world networks with geographical effects // *Physica A.* – 2009. – Vol. 388. – P. 3615-3620.
21. Barrat A., Baronchelli A., Dall'Asta L. et al. Agreement dynamics on interaction networks with diverse topologies // *Chaos.* – 2007 – Vol. 17. – P. 026111.
22. Dall'Asta L., Baronchelli A., Barrat A. et al. Nonequilibrium dynamics of language games on complex networks // *Phys. Rev.* – 2006. – Vol. E74. – P.036105.
23. Baronchelli A., Dall'Asta L., Barrat A. et al. The role of topology on the dynamics of the Naming Game // *Eur. Phys. J. Special Topics.* – 2007. – Vol. 143. – P. 233-235.
24. Wang W.X., Lin B.Y., Tang C.L. et al. Agreement dynamics of finite-memory language games on networks // *Eur. Phys. J.* – 2007. – Vol. B60. – P. 529-536.
25. Li B., Chen G., Chow T.W.S. Naming game with multiple hearers // *Commun. Nonlinear Sci.* – 2013. – Vol. 18. – P. 1214-1228.
26. Murzakhmetov A., Dyusembaev A., Umbetov U., Abdimomynova M., Shekeyeva K. Study of the innovations diffusion on the base of naming game mathematical model, vol. 9(1), COMPUSOFT, An International Journal of Advanced Computer Technology. pp. 3547-3551, 2020.

# HEALTH AND SOCIAL CARE INNOVATION ASSISTIVE TECHNOLOGIES

*Prof Salah Al-Majeed*  
*Professor of Systems Engineering*  
*School of Computer Science, University of Lincoln, UK*

## **1. Introduction**

In a society in which technology is consistently becoming more abundant and desired in the over-55s group - with the number of people aged 60 or over set to double by 2050, accounting for 1 in 5 people, making population aging one of the most difficult challenges in the coming years [1, pp. 1] - it's understandable that those with impairments can often feel left behind. Adoption of technology continues to climb in the over-55 group, with 90% of respondents to a 2019 survey reporting that they have WiFi in their home, and 80% reporting that they own a smartphone [2, pp. 7]. Along with the recent surge in uptake, technological advancements for those individuals with physical or mental impairments are always being made, allowing those who had lost the ability to perform certain tasks safely (if at all) to regain some sense of personal freedom. This sense of freedom is especially important, as it serves as a key aspect of well-being alongside personal sense of achievement [3].

This literature review aims to outline existing technologies and upcoming trends in social care technology in a variety of forms, including Internet of Things (IoT), Cloud Computing, Machine Learning, and Artificial Intelligence (AI) devices and solutions. These technologies can be thought of as falling under the categories of either Assistive Technology - bespoke/specialised devices tailored for specific needs and often privately manufactured - or Everyday Technology - typically low-cost, off-the-shelf, consumer-grade technology modified to perform care-oriented functions. The devices focused on will primarily consist Everyday Technology as these devices are typically able to provide the most beneficial solutions within the shortest timescales (and the majority are IoT-enabled), although Assistive Technology devices can have benefits despite their pitfalls and will also be discussed.

## **2. Existing research**

Technology utilised for in-home care can be thought of as falling into one of two categories: Everyday Technology (ET), or Assistive Technology (AT). Everyday Technology (ET) refers to electronic/digital devices - such as app-enabled computers, smartphones, tablets, etc. - that already exist in many homes and/or are typically readily accessible (or accessible with small modifications) to those with physical or mental impairment. The usage of ET devices by those with impairments (both mental and physical) can have both significant social benefits [4] and meaningful lifestyle benefits by providing the user with the ability of taking part in independent interaction with hobbies and other pleasurable activities [5]. ET's also have the benefit of generally being comparatively more affordable than bespoke/specialised devices - even for those with very specialised or specific needs - increasing the likelihood of them being affordable enough for use in home settings. Such a dramatic difference in cost also results in beneficial solutions/devices being able to be deployed in comparatively significantly shorter timescales than more specialised or professional-grade devices. A large portion of ET devices also happen to be Internet of Things devices simply due to their affordable, off-the-shelf, interconnected nature, but they can also be non-interconnected devices; everyday objects with no inherent connectivity - washing machines, coffee machines, record players, etc. - also fall under the umbrella term of ET [6, pp. 145] as they perform a function or fulfil a need for the user while also being available off-the-shelf. Assistive Technology (AT), on the other hand, typically refers to devices that're bespoke or fit-for-purpose, such as prosthetics/orthotics, walkers, scooters, assistant robots, etc. ATs are generally not available off-the-shelf and are very much tailored to a specific use case, be that a specific disease, impairment, or task. ATs also tend to be comparatively much less adaptable than ETs, as ATs are typically designed to serve a singular purpose, whereas ETs are typically either multipurpose or highly modifiable.

### 3. Assistive Technology

Assistive technology can be thought of as falling into 4 categories: physical, visual, hearing/speaking, and wearable. AT devices can, however, fall under multiple categories. Physical AT refers to technology that can aid a user with physical tasks (for example, an electric wheelchair for movement or a mouth stick for pointing/typing/etc). For users with visual impairments, visual ATs are available; the usage of adaptations/technologies like the usage of large fonts (for example on keyboards) or screen reading/magnifying technologies allow for high degrees of quality-of-life retention in the visually impaired. Devices or technologies tailored for aiding those with hearing or speaking impairment are also available, taking the form of hearing aids, automated transcription/closed captioning, or other such systems. Wearable devices can go a long way toward improving the quality of life of users by helping to ensure personal safety and confidence via monitoring of conditions. Some examples of wearable ATs would be:

#### A. Physical Assistive Technology

Physical AT is extremely beneficial for those with physical impairments, allowing them to perform physical tasks they otherwise couldn't by more typical means. Some prime examples of physical AT either already commonly used in at-home care settings or being actively researched are as follows:

- Mobility aids
- Head pointers
- Mouth sticks
- Sip-and-puff systems
- Voice recognition
- Caregiver/social robots

B. Visual Assistive Technology. Visual AT is especially vital to the field of home care given the large numbers of people declaring they are suffering from visual impairment - 2.2 billion according to the WHO [9]. AT devices tailored to aiding those with visual impairments are therefore a fruitful field of research [10]. A wide variety of these solutions are simple changes to existing methods of interaction, for example simply printing a book in a larger print, but there is also a plethora of technology that can aid those with visual impairments:

- Braille displays
- Braille watches
- Speech recognition
- Speech synthesis
- Magnification systems

C. Hearing/Speaking Assistive Technology. Hearing/speaking AT is an interesting area, as some of the devices' capabilities are more difficult to replace by more readily available ET solutions due to their specialised nature. One example of this would be hearing aids, which are most often the main clinical intervention for mild to moderate hearing loss [12] and perform a very specialised task. There are various other examples of hearing/ speaking AT commonly used in at-home care, all fulfilling different needs for the user, of a less specialised nature:

- Automatic transcription/closed captioning
- Eye-gaze tracking
- Audio listening devices

D. Wearable Assistive Technology. Wearable AT is quite common in the area of in-home AT, as it allows for the device to be on the user at all times, often performing active monitoring, which is beneficial for many application areas. These devices are also unintrusive, as they need to be small enough to be wearable, which means they can often be hidden for those users who fear societal stigma for utilisation of such devices. Common wearable AT devices in use within in-home care are as follows:

- Fall detection devices
- Health/Fitness monitoring devices
- Reminder devices (e.g. for medication)

- Head-mounted devices

A. **Everyday Technology. Internet of Things.** Internet of Things (IoT) devices are especially common within the realm of ET devices. As wireless connectivity is becoming commonplace in the majority of households, inter-connectivity between devices and systems is also increasing, allowing IoT-enabled devices to provide limitless prospects to augment the standard of living for the differently abled [16, pp. 1]. IoT-enabled devices are a prime example of cognitive computing, which are defined as smart systems that naturally interact with humans and other smart systems [17] - this native inter-connectivity allows for the sharing of information between IoT devices in a network, allowing for an even wider range of use cases. As an example, an Alzheimer's patient may forget to turn off an appliance, but an IoT-enabled appliance or detector can simply communicate this information to a device capable of communicating to the patient, avoiding potential disaster while also increasing user independence and confidence. A strong benefit of IoT devices resides in their dynamic and adaptable nature, which allows for an IoT device (or network of devices) set up in a patient home to be tailored to follow a patient's clinical pathway (diagnostic and therapeutic procedures tailored to the treatment of that specific patient) [18].

B. **Edge/Fog and Cloud Computing.** One issue that can be encountered when using devices in at-home care settings (particularly off-the-shelf IoT devices with limited hardware capabilities) is resource scarcity - the devices themselves not having sufficient hardware for the proposed task. The current common method for dealing with this issue, particularly in tasks where Machine Learning (or any other resource-intensive operation) is involved, is the usage of edge computing (also known as fog computing) and/or cloud computing to offload tasks to more capable systems [26, pp. 19]. An example of this would be the framework for object recognition with audible return outlined in [27], where a Raspberry Pi (the edge computing component) [28] communicates with the IoT device (an EPS32-CAM camera) to provide object detection via the utilisation of Google Colab (the cloud computing component) for machine learning algorithm application. The small footprint of these devices can allow for the devices to be carried by the patient, which can result in increased confidence/decreased anxiety (via both concealment of the devices and the increase in personal freedom), as well as high levels of portability.

C. **Machine Learning.** Breakthroughs and ongoing active research within the field of machine learning (ML) focused on tasks with high healthcare relevancy - such as face recognition [29] (even if face coverings are involved [30]), medical diagnosis [31], and natural language processing [32] - alongside high levels of potential and promising performance make it a prime candidate for usage within in-home care systems. However, with high levels of performance and potential comes the issue of very large datasets, massive storage utilisation, and demanding computing power requirements. These issues can be lessened significantly via the leveraging of cloud computing (for computing power and storage requirements), edge computing (for offloading of computing power demands to other devices) and edge ML (the usage of edge devices to learn tasks on local data to provide low latency inference). These technologies are often combined for offloading of individual tasks to devices best suited for those tasks e.g. IoT devices in-home for patient interfacing with machine learning performed on the cloud/edge [33].

#### 4. Conclusion

To conclude, high levels of adaptability, comparatively very low cost of entry, and native ability to form a network and interface with other devices makes IoT devices an ideal choice as compared to more bespoke, privately manufactured devices. The adaptability IoT devices permit allows for each device to be tailored specifically to the needs of an individual. The low barrier of entry in terms of price should result in more uptake, which should further proliferate the number of in-home patients receiving assistance tailored to their everyday needs. IoT devices' ability to interface with one another allows for the spreading of hardware workload - this capability combined with the potential integration of edge/fog computing systems and cloud computing systems allows for in-home assistance capabilities that would normally be impossible (or far more expensive) due to the hardware limitations of typical IoT devices. This low barrier of entry

and inherent adaptability, when coupled with the capabilities of ML algorithms (and adaptability via methods such as domain adaptation and the use of local training data), allows for a lot of in-home automation of care via the creation of an ambient assisted living (AAL) ecosystem. Sensors (be they camera devices, temperature sensors, accelerometers, or any other) can be used to persistently monitor not only the patient, but also their environment and its visitors, care staff or not. While privacy concerns are valid, much research exists (and is still ongoing) on privacy preservation within in-home monitoring systems [42, 43, 44].

AI and robotics systems/devices have the capability of serving a wide variety of patients' needs, whether that be guidance (route planning), reminders (medication alerts), company (socially assistive robotics), or physical help (patient transport/lifting, bathing, and feeding. AI and robotics have been tested within research in a wide variety of in-home care environments and has certainly shown itself capable of performing important safety tasks (fall detection, medical emergency detection, etc) to a reliable degree via live analysis of multi-modal sensors. Despite these benefits, AI and especially robotics have a very high financial barrier of entry, with some robotic assistants costing tens of thousands of pounds [54]. Societal stigmas and ethical challenges surrounding Human-Robot Interaction (HRI) serve as an additional barrier of entry, with some patients likely being uncomfortable with the idea of a live-in robot assistant. HRI and the associated societal and ethical challenges is, however, a very hot topic and advances are likely to be made in this field in the future [55].

## References

- [1] R. de Podest a Gaspar, R. Bonacin, and V. P. Gon calves, \Designing iot solutions for elderly home care: a systematic study of participatory design, personas and semiotics," in International Conference on Universal Access in Human-Computer Interaction. Springer, 2018, pp. 226{245.
- [2] Link-age Connect, \2019 technology survey older adults age 55-100," Linkage Connect, 2019.
- [3] M. E. Seligman, Flourish: A visionary new understanding of happiness and well-being. Simon and Schuster, 2012.
- [4] A. Capstick, \Travels with a flipcam: bringing the community to people with dementia in a day care setting through visual technology," Visual Studies, vol. 26, no. 2, pp. 142{147, 2011.
- [5] F. S. Lim, T. Wallace, M. A. Luszcz, and K. J. Reynolds, \Usability of tablet computers by people with early-stage dementia," Gerontology, vol. 59, no. 2, pp. 174{182, 2013.
- [6] L. Nyg ard and S. Starkhammar, \The use of everyday technology by people with dementia living alone: Mapping out the di culties," Aging & mental health, vol. 11, no. 2, pp. 144{155, 2007.
- [7] D. Purwanto, R. Mardiyanto, and K. Arai, \Electric wheelchair control with gaze direction and eye blinking," Arti cial Life and Robotics, vol. 14, no. 3, pp. 397{400, 2009.
- [8] M. Ferretti, G. Morgavi, and G. Veruggio, \The acceptability of caregiver robots in elderly people." in ICT4AWE, 2018, pp. 111{118.
- [9] W. H. Organization, \Vision impairment and blindness." [Online]. Available: <https://www.who.int/en/news-room/fact-sheets/detail/blindnessand-visual-impairment>
- [10] S. S. Senjam, \Assistive technology for people with visual loss," The Official Scientific Journal of Delhi Ophthalmological Society, vol. 30, no. 2, pp. 7{12, 2020.
- [11] X. Qu, X. Ma, B. Shi, H. Li, L. Zheng, C. Wang, Z. Liu, Y. Fan, X. Chen, Z. Li et al., \Refreshable braille display system based on triboelectric nanogenerator and dielectric elastomer," Advanced Functional Materials, vol. 31, no. 5, p. 2006612, 2021.
- [12] M. A. Ferguson, P. T. Kitterick, L. Y. Chong, M. Edmondson-Jones, F. Barker, and D. J. Hoare, \Hearing aids for mild to moderate hearing loss in adults," Cochrane Database of Systematic Reviews, no. 9, 2017.



- [13] H. Khachatryan and A. L. Rihn, "Eye-tracking methodology and applications in consumer research," *Electron Data Inf Source UF/IFAS Ext*, pp.1{5, 2014.
- [14] G. E. Lancioni, N. N. Singh, M. F. O'Reilly, J. Sigafos, F. D'Amico, F. Buonocunto, C. Lanzilotti, G. Alberti, and J. Navarro, "Mainstream technology to support basic communication and leisure in people with neurological disorders, motor impairment and lack of speech," *Brain Injury*, vol. 34, no. 7, pp. 921{927, 2020.
- [15] A. Ramachandran and A. Karuppiah, "A survey on recent advances in wearable fall detection systems," *BioMed research international*, vol. 2020, 2020.
- [16] M. Bansal and S. Garg, "Internet of things (iot) based assistive devices," in *2021 6th International Conference on Inventive Computation Technologies (ICICT)*. IEEE, 2021, pp. 1006{1009.
- [17] H. Demirkan, S. Earley, and R. R. Harmon, "Cognitive computing," *IT professional*, vol. 19, no. 4, pp. 16{20, 2017.
- [18] C. Ardito, T. Di Noia, E. Di Sciascio, D. Lof u, G. Mallardi, C. Pomo, and F. Vitulano, "Towards a trustworthy patient home-care thanks to an edge-node infrastructure," in *International Conference on Human-Centred Software Engineering*. Springer, 2020, pp. 181{189.
- [19] "All-new echo show 5: 2nd generation (2021 release), smart display with alexa and 2 mp camera." [Online]. Available: <https://www.amazon.co.uk/echo-show-5-2nd-generation-2021-release/dp/B08KGTW3CV>
- [20] [Online]. Available: <https://www.google.co.uk/>
- [21] H. Wandke, M. Sengpiel, and M. Sönksen, "Myths about older people's use of information and communication technology," *Gerontology*, vol. 58, no. 6, pp. 564{570, 2012.
- [22] A. Astell, S. Smith, and P. Jodrell, *Using technology in dementia care: a guide to technology solutions for everyday living*. Jessica Kingsley Publishers, 2019.
- [23] A. F. Klaib, N. O. Alsrehin, W. Y. Melhem, and H. O. Bashtawi, "Iot smart home using eye tracking and voice interfaces for elderly and special needs people." *J. Commun.*, vol. 14, no. 7, pp. 614{621, 2019.
- [24] [Online]. Available: <https://thinksmartbox.com/product/grid-3/>
- [25] P. Nawandar and V. V. Gohokar, "Design and development of multisensory smart assistive technology for blind persons," in *2018 International Conference on Research in Intelligent and Computing in Engineering (RICE)*. IEEE, 2018, pp. 1{4.
- [26] M. Hartmann, U. S. Hashmi, and A. Imran, "Edge computing in smart health care systems: Review, challenges, and research directions," *Transactions on Emerging Telecommunications Technologies*, p. e3710, 2019.
- [27] M. J. Junior, O. B. Maia, H. Oliveira, E. Souto, and R. Barreto, "Assistive technology through internet of things and edge computing," in *2019 IEEE 9th International Conference on Consumer Electronics (ICCE-Berlin)*. IEEE, 2019, pp. 330{332.
- [28] R. Pi, "Teach, learn, and make with raspberry pi." [Online]. Available: <https://www.raspberrypi.org/>
- [29] U. Aiman and V. P. Vishwakarma, "Face recognition using modified deep learning neural network," in *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*. IEEE, 2017, pp. 1{5.
- [30] Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H. Wu, P. Yi, K. Jiang, N. Wang, Y. Pei et al., "Masked face recognition dataset and application," *arXiv preprint arXiv:2003.09093*, 2020.
- [31] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *Ieee Access*, vol. 6, pp. 9375{9389, 2017.
- [32] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," *iee Computational intelligenCe magazine*, vol. 13, no. 3, pp. 55{75, 2018.

- [33] D. Mrozek, A. Koczur, and B. Ma lysiak-Mrozek, "Fall detection in older adults with mobile iot devices and machine learning in the cloud and on the edge," *Information Sciences*, vol. 537, pp. 132{147, 2020.
- [34] Arduino, "Arduino." [Online]. Available: <https://www.arduino.cc>
- [35] NodeMcu, "Nodemcu { an open-source rmware based on esp8266 wi -soc." [Online]. Available: <https://www.nodemcu.com>
- [36] S. Nooruddin, M. Milon Islam, and F. A. Sharna, "An iot based device-type invariant fall detection system," *Inter- net of Things*, vol. 9, p. 100130, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2542660519301623>
- [37] N. Nahar, M. S. Hossain, and K. Andersson, "A machine learning based fall detection for elderly people with neurodegenerative disorders," in *International Conference on Brain Informatics*. Springer, 2020, pp. 194{203.
- [38] N. Zurbuchen, P. Bruegger, and A. Wilde, "A comparison of machine learning algorithms for fall detection using wearable sensors," in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*. IEEE, 2020, pp. 427{431.
- [39] H. Gjoreski, S. Stankoski, I. Kiprijanovska, A. Nikolovska, N. Mladenovska, M. Trajanoska, B. Velichkovska, M. Gjoreski, M. Lu strek, and M. Gams, "Wearable sensors data-fusion and machine-learning method for fall detection and activity recognition," in *Challenges and Trends in Multimodal Fall Detection for Healthcare*. Springer, 2020, pp. 81{96.
- [40] K. Adhikari, H. Bouchachia, and H. Nait-Charif, "Long short-term memory networks based fall detection using uni ed pose estimation," in *Twelfth International Conference on Machine Vision (ICMV 2019)*, vol. 11433. International Society for Optics and Photonics, 2020, p. 114330H.
- [41] M. M. Rahman, M. M. Islam, S. Ahmmed, and S. A. Khan, "Obstacle and fall detection to guide the visually impaired people with real time monitoring," *SN Computer Science*, vol. 1, pp. 1{10, 2020.
- [42] U. Asif, B. Mashford, S. Von Cavallar, S. Yohanandan, S. Roy, J. Tang, and S. Harrer, "Privacy preserving human fall detection using video data," in *Machine Learning for Health Workshop*. PMLR, 2020, pp. 39{51.
- [43] J. Liu, Y. Xia, and Z. Tang, "Privacy-preserving video fall detection using visual shielding information," *The Visual Computer*, vol. 37, no. 2, pp. 359{370, 2021.
- [44] S. A. Shah, J. Ahmad, F. Masood, S. Y. Shah, H. Pervaiz, W. Taylor, M. A. Imran, and Q. H. Abbasi, "Privacy-preserving wandering behavior sensing in dementia patients using modi ed logistic and dynamic newton leipnik maps," *IEEE Sensors Journal*, vol. 21, no. 3, pp. 3669{3679, 2020.
- [45] A. Moorcroft, N. Scarinci, and C. Meyer, "A systematic review of the barriers and facilitators to the provision and use of low-tech and unaided aac systems for people with complex communication needs and their families," *Disability and Rehabilitation: Assistive Technology*, vol. 14, no. 7, pp. 710{731, 2019.
- [46] A. Thirumanickam, P. Raghavendra, and C. Olsson, "Participation and social networks of school-age children with complex communication needs: A descriptive study," *Augmentative and Alternative Communication*, vol. 27, no. 3, pp. 195{204, 2011.
- [47] I. Kostavelis, D. Giakoumis, S. Malasiotis, and D. Tzovaras, "Ramcip: towards a robotic assistant to support elderly with mild cognitive impairments at home," in *International Symposium on Pervasive Computing Paradigms for Mental Health*. Springer, 2015, pp. 186{195.
- [48] S. G. R. d IRCCS, "Interacting with dementia: the mario approach," *Har- nassing the Power of Technology to Improve Lives*, vol. 242, p. 38, 2017. [49] J. Prip , T. K□prtner, D. Batko-Klein, D. Hebesberger, M. Weninger, C. Gisinger, S. Frennert, H. Efring, M. Antona, I. Adami et al., "Results of a real world trial with a mobile social service robot for older adults," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2016, pp. 497{498.

- [50] S. Co-sar, M. Fernandez-Carmona, R. Agrigoroaie, J. Pages, F. Ferland, F. Zhao, S. Yue, N. Bellotto, and A. Tapus, "Enrichme: Perception and interaction of an assistive robot for the elderly at home," *International Journal of Social Robotics*, vol. 12, no. 3, pp. 779-805, 2020.
- [51] R. Soyama, "The development of meal-assistance robot 'my spoon'," in *Proceedings of the 8th International Conference on Rehabilitation Robotics*, 2003, 2003, pp. 88-91.
- [52] RIKEN-TRI, "World's first robot that can lift up a human in its arms." [Online]. Available: <http://rtc.nagoya.riken.jp/RIBA/index-e.html>
- [53] A. Vercelli, I. Rainero, L. Ciferri, M. Boido, and F. Pirri, "Robots in elderly care," *DigitCult-Scientific Journal on Digital Cultures*, vol. 2, no. 2, pp. 37-50, 2018.
- [54] R. Booth, "Robots to be used in uk care homes to help reduce loneliness," *The Guardian*, 2020.
- [55] R. Wullenkord and F. Eyssel, "Societal and ethical issues in hri," *Current Robotics Reports*, pp. 1-12, 2020.

## ЖОО-НЫҢ БАСҚАРУ ШЕШІМДЕРІН ҚАБЫЛДАУ ҮШІН БІЛІМ САПАСЫН ҚАБЫЛДАУ

*Зүнімова Гүлназ Дүйсенәліқызы*

*С.Сейфуллин атындағы қазақ агро-техникалық университеті, Астана қаласы,  
[zunimova.g.d@mail.ru](mailto:zunimova.g.d@mail.ru)*

### **Кіріспе**

Қазіргі заман жағдайында жоғары оқу орындарын басқару жүйесін реформалау, білім беруді инновациялау, тұрақты жұмыс істейтін орта жағдайында қалыптасқан дәстүрлі басқару стратегияларын өзгертіп, сыртқы және ішкі жағдайларға икемделуге қабілетті, неғұрлым икемді, жаңа стратегияларға біртіндеп ауыстыруға алып келеді. Сонымен бірге, тәжірибеде қолданылатын жоғары оқу орындары қызметінің сапасын мониторингтеу жүйелері негізінен істердің ағымдағы үдеріс диагностикалауға бағытталған және директивті басқару жағдайында қалыптастырылды.

Мониторинг зерттеу мен зерттеудің дәстүрлі әдістерін жоққа шығарды. Мұның себебі, ол жаңа тұжырымдама емес. Мониторинг жеке тәуелсіз бағытқа айналды, оның жеткілікті ықтималдығын өлшеу, зерттеу, эксперимент және басқару арқылы анықтауға болады. Қазіргі уақытта мониторинг, университетті басқару тәжірибесінде екі негізгі аспект бойынша қолданылады. Бір жағынан, мониторинг - білім беру процесін ұйымдастырудың өзекті мәселелерін шешуге ықпал ететін технология ретінде қарастырылады. Екінші жағынан, мониторинг - басқару жүйелерін зерттеу процесінде ақпарат алу құралы. Сонымен бірге, мониторинг нәтижесінде алынған ақпарат білім беру қызметінің әртүрлі түрлерін жетілдірудің шарты ретінде қарастырылады. Бірқатар зерттеушілер мониторингті жоғары оқу орнының ақпараттық ортасын ұйымдастыру әдісі ретінде қолданады.

### **Негізгі бөлім**

Арнайы зерттеулердің нәтижелері білім беру процесі сапасының мониторингі негізінен, білім беру қызметін жетілдіру мақсатында одан әрі пайдалану үшін ақпарат алу тәсілі ретінде қарастырылатындығын көрсетеді. Мемлекеттік білім беру саясатының мазмұнына сәйкес қазіргі заман жоғары кәсіптік білім берудің мақсаттары мен міндеттеріне қатысты ХҚКК-нің жалпы мақсатын ел мен белгілі бір аймақты бәсекеге қабілетті кадрлармен қамтамасыз ету міндетін қаншалықты шеше алатынын және білім беру қызметтері мен еңбек нарығының сәйкестігін қамтамасыз ететін ақпаратты жүйелі түрде жинайды әрі талдау жүргізеді. [1]

*Нәтижелер.* Бұл жұмыс университет деңгейіндегі жоғары білім беру жүйесіндегі сапа мониторингін қарастырады. Бақылау, басқару және оңтайландыру бойынша университеттің дербес қабылданатын әрекеттері сапа, ішкі сапа мониторингі ретінде сипатталады. Мониторингтің бұл түрі үшін кейде «өзін-өзі бағалау, өзін-өзі сертификаттау» термині қолданылады. Сыртқы бағалауға (үшінші тараптың бағалауы) негізделген білім беру нәтижелерін бағалау бойынша университеттің қызметін сыртқы сапа мониторингі деп атайды. Мониторингтің бұл түрі университеттің басқару органдары алдындағы есептілігі арқылы жиі талданады.

*Сыртқы мониторинг.* Сыртқы мониторингтің мақсаты білім беру сапасының ішкі бағалауға қаншалықты сәйкес келетінін анықтау болып табылады.

Білім беру сапасының сыртқы мониторинг жүйесі:

- 1) Нақты білім беру бағдарламасының қоғам тарапынан қойылатын талаптарға әлеуметтік-экономикалық сәйкестігін анықтайды;
- 2) білім беру тиімділігін бағалау критерийлері мен қағидаттарын әзірлеу есебінен жоғары білім берудегі прогресті қамтамасыз етеді;
- 3) үздіксіз өзін-өзі басқару және университет қызметін жоспарлау арқылы білім беру бағдарламаларының дамуы мен жетілдірілуін ынталандырады;

4) оқу орнының білім беру қызметінде оның академиялық еркіндігін дамытуға ықпал етеді.[2]

Мемлекеттік деңгейде жоғарғы білім беру сапасын бағалау құрылымындағы мониторинг маңызы 1-кестеде келтірілген:

*1-кесте.* Университеттегі білім беру деңгейлері бойынша білім беру сапасын бағалау нысандары

Жоғарғы білім деңгейлері			
	Бакалавриат	Магистратура	Докторантура
Сапаны бағалаудың мемлекеттік рәсімдері	Қорытынды мемлекеттік аттестаттау		
	Лицензиялау		
	Лицензиялық бақылау		
	Мемлекеттік аккредиттеу		
	Сапаны бақылау		
Білім беру сапасын тәуелсіз бағалау	Білім беру сапасын қоғамдық бағалау		
	Кәсіби-қоғамдық аккредиттеу		
	Кәсіби және қолданбалы біліктіліктерді сертификаттау		
Деректерді жинау механизмі	Мониторинг		

*Ішкі мониторинг* кез келген ЖОО-да білім беру бағдарламаларын іске асыру сапасын анықтау, рейтингтер және басқа да бағалау рәсімдерін кейіннен жетілдіруге бағытталған іс-шаралар кешенін әзірлеу және іске асыру мақсатында жүзеге асырылады. Университет ішкі мониторинг моделін дербес жасайды, бірақ кез келген жағдайда ол келесі құрамбөліктерді қамтитын интегралды жүйе болып табылады. Құрамдас бөліктер:

*Мониторингтің мақсаты*-ЖОО-дағы білім беру жай-күйі және оның қызметінің негізгі көрсеткіштері туралы ақпаратты талдау үшін база қалыптастыру;

*Мониторинг міндеттері:*

- даму процесінің шарттары, ұйымдастырылуы, мазмұны және нәтижелері туралы мәліметтер жинау;
- білім беру процесі сапасының жай-күйін бағалау, ЖОО кіші жүйелерінің жұмыс істеуін бағалау (түпкілікті нәтижелерге қол жеткізу бағытындағы өзгерістерді анықтау – "өте жақсы", "жақсырақ", "нашар", "тең");
- білім беру үрдісін түзету (нәтижелілігін арттыру) және ЖОО даму үрдісін болжау бойынша басқарушылық қызмет;

*Мониторинг талаптары.* Ол болуы керек: үнемді; объективті; үздіксіз, болжамды, білім беру процесінің барлық қатысушыларын өз қызметінің сапасын арттыруға ынталандыру;

*Мониторинг объектілері* мүдделі тараптардың қажеттіліктерін талдаудан бастап және тұтынушылардың қанағаттануын бағалауға дейін университетішілік жүйенің кез келген құрамдас бөліктері болуы мүмкін;

*Мониторинг субъектілері* – мониторингтік зерттеулер мен бағалауларды жүргізетіндердің барлығы: университеттің функционалдық қызметтері, топ-менеджмент, мұғалімдер, студенттер. Білім беру үдерісіне тікелей қатысы жоқ, бірақ оны жүзеге асыруға қатысатын басқа да тұлғалар (ғалымдар, білім беру органдарының өкілдері, жұмыс берушілер және т.б.) субъектілер бола алады;

*Мониторинг әдістері:* сараптамалық бағалау әдісі; құрылымдық бөлімшелердің, оқытушылар құрамының құжаттары мен жұмыс тәжірибесін талдау; социологиялық әдістер: бақылау, сауалнама;

*Мониторингті университет басшылығы басқарады,* оның орындалуын тиісті қызметтер ұйымдастырады. Мониторинг нәтижелері бойынша нақты зерттеулердің мақсаттары мен міндеттеріне сәйкес келетін аналитикалық материалдар дайындалады;

*Мониторинг критерийлері:*

- стандарттарға (әртүрлі деңгейдегі құжаттарда белгіленген), ең алдымен мемлекеттік білім беру стандартында бекітілген жоғары білім сапасының стандарттарына сәйкестік дәрежесі;
- мүдделі тараптар қоятын талаптарға сәйкестік дәрежесі;
- жағдайдың өзгеруі ( " + " , " - "тенденциялары).

Білім беру сапасын Мемлекеттік бағалау рәсімдерінен өтуге ықпал ететін аккредиттеу, сондай-ақ жоғары оқу орны ішіндегі сапа менеджменті жүйесін растау кезіндегі инспекциялық және ресертификациялық сыртқы аудиттер. Жалпы түрде аккредиттеу тобының критерийлері мен көрсеткіштері 2-кестеде келтірілген.

Осыдан тұжырымдайтын болсақ: жүйелік мониторингтің жеке ерекшеліктерінің мазмұны әр түрлі болуы мүмкін, өйткені әрбір университет өз қажеттіліктеріне сәйкес мониторингті қалыптастырады, сондықтан пайдаланылған көрсеткіштер ішкі бақылау кезінде (кестеде көрсетілген), егжей-тегжейлі көрсетуді талап етеді және университеттің өзі анықтайды.

*2-кесте*

№	Критерийлер	Көрсеткіштер
1	Оқу жоспарлары мен бағдарламаларының анализі	Мемлекеттік білім беру стандарттарына оқу жоспарларының сәйкестігі
2	Профессор-оқытушылар құрамының анализі	ПОҚ үлесі, яғни ғылым докторлары, штаттық оқытушылар құрамының үлесі Ғылыми қызмет Әдістемелік іс-әрекет Біліктілікті арттыру
3	Логистиканы оқыту бағдарламаларын қамтамасыз ету	Ғимаратпен қамтамасыз етілуі Зертханалармен қамтамасыз етілуі, Арнайы аудиториялар Шығыс материалдарымен қамтамасыз ету
4	Білім беру процессін электрондық-ақпараттық қамтамасыз ету	Кітаппен қамтамасыз ету Электрондық білім беру ресурстарымен қамтамасыз ету Интернетте лицензиялық БҚ-ның болуы, қолжетімділігі
5	ПОҚ ғылыми-зерттеу жұмысы	Ғылыми мектептер Ғылыми бағдарламалар Ғылыми гранттар және экономикалық келісім-шарттар тақырыптары
6	Тәжірибелерді ұйымдастыру	Практика базаларымен қамтамасыз етілуі Мемлекеттік білім беру стандарттары практикасының еңбек сыйымдылығымен сәйкестігі
7	Студенттер анализі (формализацияланған бағалау білімі)	Студенттер үлгерімі Барлық оқу мерзімінде студенттерді оқудан шығару пайызы

8	Студенттердің ғылыми жұмысы	Студенттік басылымдар Студенттік ғылыми гранттар
9	Білімді бақылау жүйесін ұйымдастыру	Ағымдағы және аралық бақылау нысандары Оқу үлгерімі Балдық-рейтингтік жүйенің болуы
10	Түлектер сұранысы	Түлектердің жұмысқа орналастырылуы

Осы талаптар мен көрсеткіштерге байланысты студенттер мен түлектерге қандай құзыреттер жетіспейтінін, түлектердің қандай талаптары бар екенін және білім беру бағдарламасын қалай өзгерту керектігін, жұмыс берушілердің түлектерге қандай талаптары бар екенін, олардың мансабына осы жұмыс орнын анықтайтындардың кедергі келтіретінін және т. б. анықтау болуы мүмкін. [3,4,5]

### **Қорытынды**

Қорытындылай келе, ЖОО ішілік мониторингі –білім сапасының жағдайын қадағалау процесі болып табылады. Яғни, оқу үдерісін кез келген уақытта қадағалау арқылы университет жағдайына жылдам жауап беруге және оны басқаруға мүмкіндік беретін “кері байланыс” функциясы ретінде қолданылады.

### **Қолданылған әдебиеттер**

- 1 Крылова Е.В. Внутренний мониторинг качества профессиональной подготовки студентов инженерного вуза: автореф. дис. ... канд. пед. наук. – М., 2009 [Электронный ресурс]. URL: <http://aeer.cctpu.edu.ru/winn/magazine/m6/IO-2010.pdf> (дата обращения: 12.05.2014).
- 2 Berlin Principles on Ranking of Higher Education Institutions [Электронный ресурс]. URL: [http://www.che.de/downloads/Berlin\\_Principles\\_IREG\\_534.pdf](http://www.che.de/downloads/Berlin_Principles_IREG_534.pdf) (дата обращения: 20.11.2014).
- 3 Aguiar, E., Chawla, N., Brockman, J., Ambrose, G., & Goodrich, V. (2014). Engagement vs performance: Using electronic portfolios to predict first semester engineering student retention. In I. Molenaar, X. Ochoa, & S. Dawson (Eds.), Proceedings of the fourth international learning analytics & knowledge conference (pp. 103–112). New York, NY: ACM.
4. Круглов В.И. Качество профессионального образования и ресурсы. [Электронный ресурс]. URL: [http://yandex.ru/clck/jsredir?from=yandex.ru%3Byand\\_search%3Bweb%3B%3B&text](http://yandex.ru/clck/jsredir?from=yandex.ru%3Byand_search%3Bweb%3B%3B&text) (дата обращения: 29.08.2014).
5. Капелюк З.А., Донецкая С.С., Струминская Л.М. Потребительский мониторинг удовлетворенности качеством образовательных услуг в вузе // Стандарты и качество. – 2006. – № 1. – С. 62–66.

# METHODS FOR DETECTING AND HIGHLIGHTING HOMOGENEOUS AREAS ON TEXTURE IMAGES

**Gulzira Abdikerimova, Jamalbek Tussupov**

*L.N. Gumilyov Eurasian National University  
2 Satpayev str., Astana, Republic of Kazakhstan  
E-mail: [gulzira1981@mail.ru](mailto:gulzira1981@mail.ru), [tussupov@mail.ru](mailto:tussupov@mail.ru)*

## **1 Abstract**

The article discusses methods for analyzing texture images. Methods for processing texture images in various branches of science are considered. The main goal of the study is the development and implementation of algorithms that allow identifying and segmenting areas of interest in the image, which will be associated in the field of chemistry with porosity, chemical activities, in the field of agriculture with pests and weed foci, in medicine with pathology, etc. To solve such problems, methods of texture features, clustering, R/S analysis, orthogonal transformations, and wavelet analysis are used. The main task in each study is the application of mathematical methods and algorithms that describe texture differences to separate texture regions into sub-regions. That is the question of the applicability of sets of textural features and other parameters for the analysis of experimental data to identify objects of interest in different areas of science.

## **2 Introduction**

Due to the rapid development of imaging technologies in various industries, their automatic object recognition, which in turn is based on mathematical algorithms, has received particular importance. The issue of creating computer algorithms capable of dividing a set of images into classes depending on the type of information they illustrate (to solve the classification problem) often arises when solving problems from completely different applied fields of knowledge, such as the analysis of aerial photographs of the area, the flaw detection of materials from images, the identification of pathologies on medical imaging. The need for automation in such tasks is due to the extremely large number of images analyzed, so large that manual analysis of them would take months and years. Therefore, the important qualities of the algorithms used should be both high classification accuracy and high speed. In addition, image analysis and classification algorithms are always problem-oriented.

Recognition operations on images of certain objects, as a rule, are preceded by image processing to create conditions that increase the efficiency and quality of selection and recognition of the desired or studied objects. Image processing methods depend on the objectives of the research, are quite diverse, and may include, for example, highlighting the most informative fragments, enlarging them, increasing contrast resolution, improving image quality, studying textural features, etc. In this regard, the development of such algorithms and software tools that implement them is an urgent task.

## **3 Methods for analyzing texture images**

Despite the ubiquitous presence of textures in images, there is no single and formal approach to the description of texture and its strict definition at the moment. Texture analysis methods are usually developed for each case. In [1], the texture is understood as “the spatial organization of elements within a certain area of the surface”. It also explains that this organization is due to a certain statistical distribution of the intensity of gray tones or tones of different colors. An area can be considered a textured area if the number of intensity drops or color changes noted on it is large enough. In [2], a texture is called “a surface area organized in some way”. In [3], the texture is defined as a matrix or fragment of spatial properties of image areas with homogeneous statistical characteristics. Textures can be divided into several classes as follows:



1) by origin: artificial - for example, graphic patterns, and natural - for example, grass, forest, earth;

2) according to the surface structure: structural, consisting of geometrically regular repeating elements, and stochastic, formed by a sequence of random elements; according to the relative sizes of texture elements: fine-grained and coarse-grained;

3) according to the shape of the texture elements: wavy, spotted, irregular, lined, and so on [1]. From the above definitions and characteristics, it follows that a texture is a certain area of the image, one that has uniform statistical characteristics. This means that each texture of a given class can be described using a characteristic property common to all textures of a given class [1]. Such properties are called textural features [4]. Textural features play an important role in dividing the image into separate areas. As such features, one can use the statistical characteristics of spatial distributions calculated as homogeneity measures for a one-dimensional histogram of signal values (characteristics of the 1st order) and two-dimensional histograms of signal values (characteristics of the 2nd order). The second class is signs that take into account the relative position.

When extracting informative features, we consider two methods of calculation. The standard approach for computing texture features is as follows. It is necessary to select the so-called running window with an odd side: 3, 5, and 7 pixels. The attribute is calculated inside the running window. The size of the local fragment is the bearer of texture properties. The feature value is written into a new matrix of the same size as the original one. In the new matrix, the value is written to a point with coordinates equal to the coordinates of the center of the running window. The elements of the new matrix are obtained in some intervals [A, B]. Further, this interval is usually linearly mapped into the interval [0, 255]. After that, it is possible to visualize the result of calculating the texture feature.

When calculating texture features using a non-standard approach, large windows (including non-square ones) are selected, which the user can specify by choosing an area that may be of interest to him. That is, we are talking about the calculation of numerical characteristics relating to vast areas, including various kinds of artifacts. In each of the presented tasks, important textural features were considered, all of which are implemented in the program. Table 1 below shows some of them.

Table 1. Textural features

<p><b>Energy</b></p> $T_1 = \sum_{i=1}^N \sum_{j=1}^N [p(i, j)]^2$
<p><b>Homogeneity</b></p> $T_3 = \sum_{i=1}^N \sum_{j=1}^N \frac{p(i, j)}{1 +  i - j };$
<p><b>Entropy</b></p> $T_2 = - \sum_{i=1}^N \sum_{j=1}^N p(i, j) \cdot \log_2(p(i, j));$
<p><b>Dispersion</b></p> $T_8 = \sum_{i=1}^N \sum_{j=1}^N (i - \mu)^2 p(i, j);$

Image segmentation of the cell wall on microscopic images. As an example, we took a micrograph of the wall of wheat straw obtained after mechanical activation at a temperature of 10°C (Fig. 1). It is carried out in a standard way, that is, the local fragment size of 9x9 is the bearer of the texture properties.

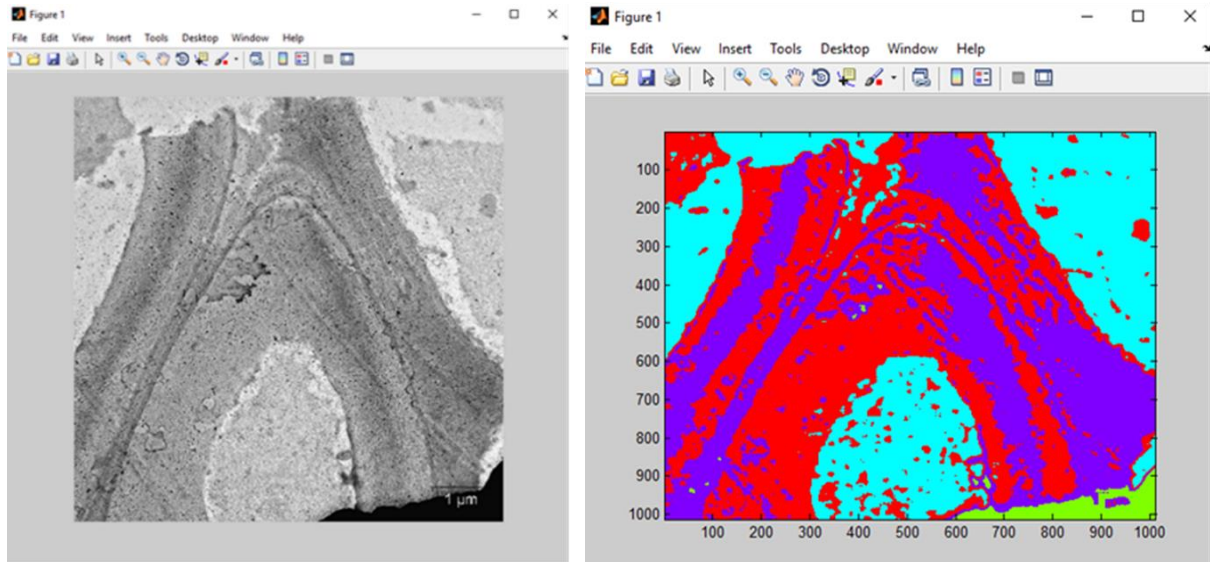


Figure 1. The result of clustering after applying informative textural features

From the obtained results, the trace elements of micrographs and the contour of the cell wall were distinguished.

Wavelet is a decomposition of images. The second approach is the wave decomposition [5]. It can be said that the wavelet decomposition [6] is performed using blocks of two-channel filters associated with a tree. At the top of this tree, whose height is  $d$ , there should be a signal length of  $2^d$ . Accordingly, the image size should be  $2^d \times 2^d$ . If this is violated, zero values are usually added to the missing values. At each transformation step, the image is divided into 4 matrices. One is an image similar to the original, but half the size horizontally and vertically. We can say that this is a "rough" version of the original image. In the other three matrices, differences in the brightness of the original image are "encoded". Thus, there is a "detailing" of the information contained in the image. Element values may not fall within the interval  $\{0, \dots, 255\}$ . For each of the matrices, we find the values of the minimum and maximum elements and display the corresponding interval in the interval  $\{0, \dots, 255\}$ . This allows you to visualize the results as grayscale or color images. The simplest forms of wavelet transform for images are the Haar and Daubechies-4 transform. Below are the results of such an experiment on the Haar transform (Fig. 2)

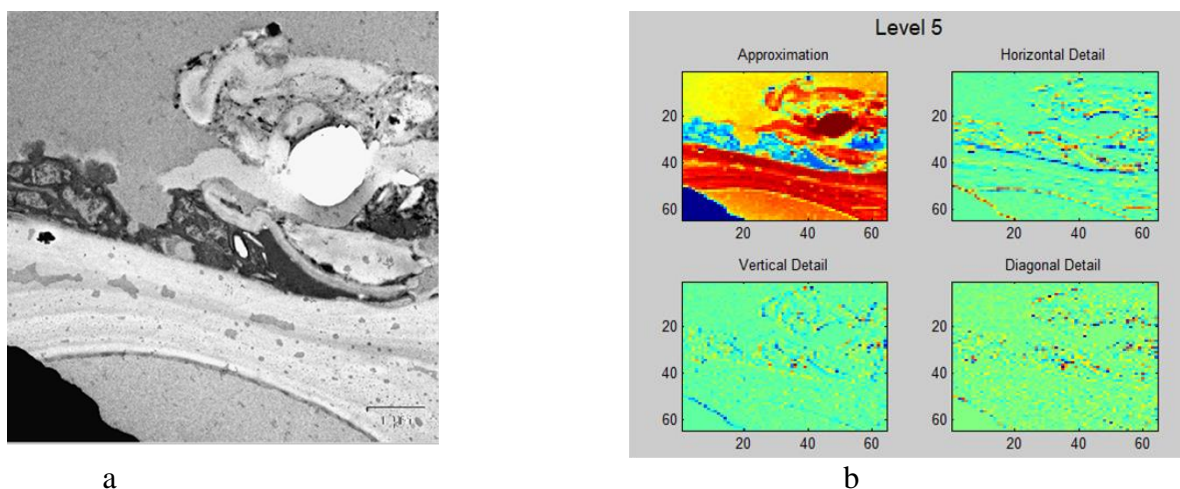


Figure 2. a - Original image; b - fifth level of decomposition

By experimenting with the results obtained by the decomposition method, we left the result of the fifth level informative. As shown in figure 2, we can see microelements, i.e. information signs, of microscopic images vertically, diagonally, and horizontally.

Various methods and technologies of remote information processing are used to solve various thematic problems of video decoding [7]. Methods for processing aerial photographs include pre-correction procedures for working with images of several regions, transformation, brightness transformation, image classification, and complex analysis of spatial information. Aerial photographs were taken to identify crops and weeds in them. An example of the identification and cultivation of plant species is shown in Figure 3.

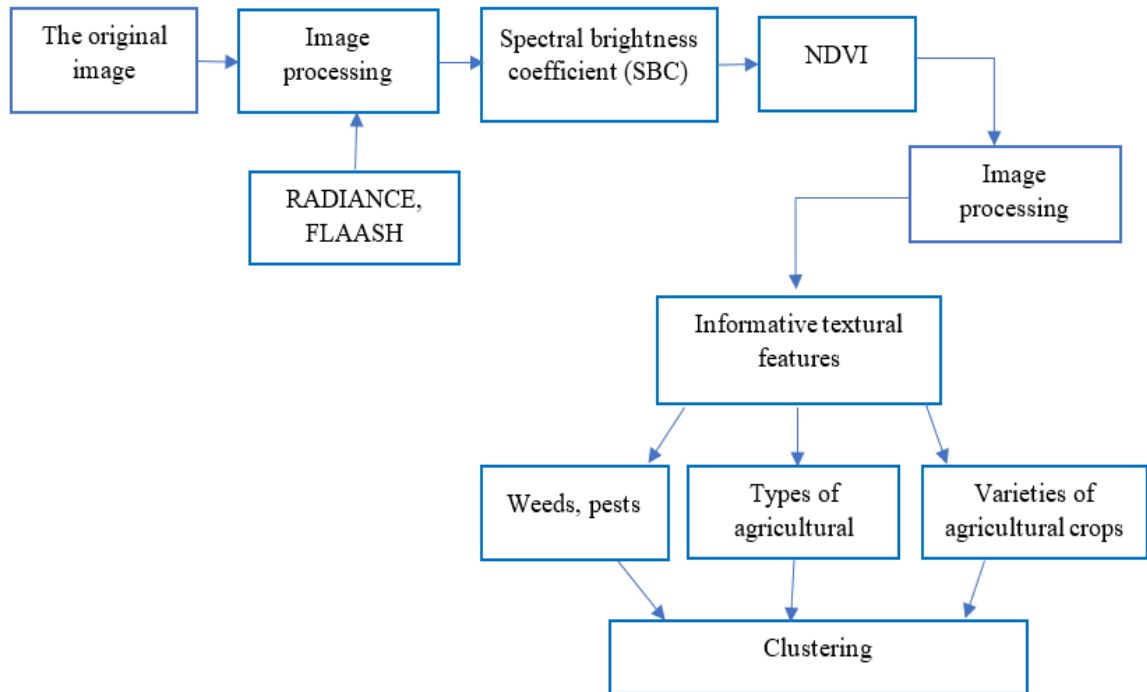


Figure 3. A method for processing satellite images with the joint use of SBC and NDVI

- The original multispectral image was taken from [8]. In the ENVI environment, it is possible to carry out all stages of the "image-to-map" process, analyze raw data of many formats, provide pre-processing of images, offer tools for image enhancement, automatic classification according to a user-defined benchmark or non-benchmark scheme, measure and analyze intermediate and final results, a wide range of tools that allow the use of spatial modeling operations include the presentation of the final result in various formats and the creation of printed maps and three-dimensional models [9].

- Image processing includes transformations aimed at improving the quality of images, increasing their visual properties, and bringing them to the most optimal form possible for further thematic interpretation. The main preprocessing procedures are radiometric and geometric correction and image enhancement operations.

The original digital image values (DN) are numbers without physical meaning associated with the values of spectral brightness coefficients according to parameters that are individual for different sensors [10]. First-level radiometric calibration consists of converting pixel values to values of reflected radiation characteristics. For example, for images from the Landsat 8 satellite, this recalculation is performed using the formula [11]:

$$r_{\lambda} = M_{\rho} \times DN + A_{\rho} \quad (1)$$

here:

- $r_{\lambda}$  - spectral coefficient of light;
- $M_{\rho}$  - strengthening of calibration;
- $A_{\rho}$  - offset calibration;
- $DN$  - initial numerical values.

Radiometric calibration of the second level is aimed at eliminating atmospheric interference and device malfunctions, correcting the angle of the ascension of the Sun and terrain [12].

According to this model, you can find any type of crop and its weeds. For example, the model below shows informative textural features of flax and weeds (figure 4).

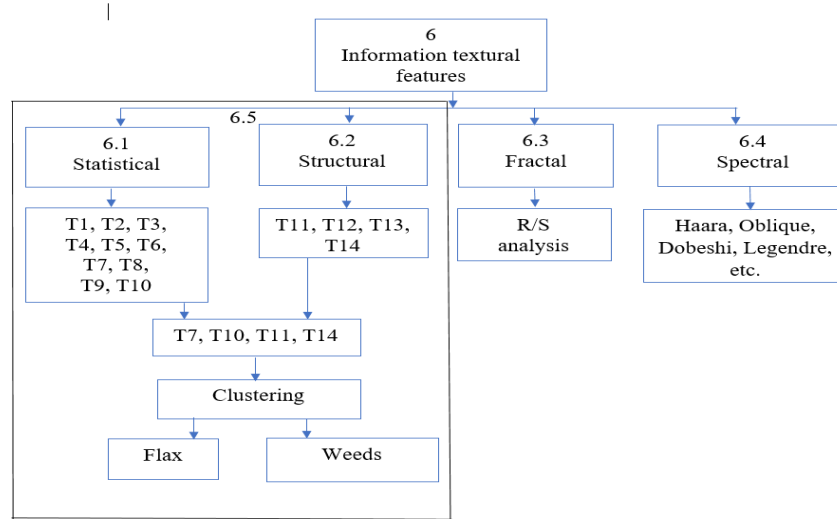


Figure 4. Informative textural features that determine the growth of weeds (flax) by growing seasons

Textural features that define weeds on flax:

The overall average value in an image closely related to the adjacency matrix is determined by the sum histogram of brightness values over element pairs [2]:

$$T_7 = \sum_{i=2}^{2N} i \cdot p_{x+y}(i);$$

The inverse moment is a measure of local similarity that determines the clarity of short runs:

$$T_{10} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{p(i, j)}{j^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j)}$$

Moments are the weight of a line that has a constant optical density. This feature is characterized by the amount of increase in the weight of each line for any gray level:

$$T_{11} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j^2 p(i, j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j)}$$

Image share in series:

$$T_{14} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j p(i, j)}$$

In the field of agriculture, cereal species have different informative vectors of textural features to identify weeds, pests, and diseases in them.

### Conclusion

The work is devoted to the study of texture images, texture segmentation, the selection of homogeneous areas in aerial photographs, and the demonstration of the effectiveness of texture marking methods. A large set of texture research methods was analyzed, and the most informative ones were selected from among the considered methods. In addition, pests, diseases,

and types of crops were identified on aerial photographs based on a model for solving problems in the field of agriculture. As a result, the following software tools have been implemented:

1. A prototype has been created for calculating various texture features.
  2. Selection of informative textural features by standard and non-standard methods.
  3. Hidden features in microphotographs were revealed by the wavelet decomposition method.
- In the future, an automated application will be developed that will allow using the machine learning method to solve problems in the main scientific field with a set of the methods mentioned, develop a system of actions to increase productivity, eliminate factors that negatively affect obtaining a high yield without harm to the environment.

## References

- [1]. Андреев Г. А. и др. Анализ и синтез случайных пространственных текстур //Зарубежная радиоэлектроника. – 1984. – №. 2. – С. 3-33.
- [2]. Харалик Р. М. Статистический и структурный подходы к описанию текстур, ТИИЭР, 1979 //Т. – 1979. – Т. 67. – С. 98.
- [3]. Потапов А. А. Новые информационные технологии на основе вероятностных текстурных и фрактальных признаков в радиолокационном обнаружении малоконтрастных целей //Радиотехника и электроника. – 2003. – Т. 48. – №. 9. – С. 1101-1119.
- [4]. Колодникова Н. В. Обзор текстурных признаков для задач распознавания образов //Доклады Томского государственного университета систем управления и радиоэлектроники. – 2004. – №. 1 (9). – С. 113-123.
- [5]. Карманов А. П. Лигнин. Структурная организация и самоорганизация //Химия растительного сырья. – 1999. – №. 1. – С. 65-74.
- [6]. Воробьев В. И., Грибунин В. Г. Теория и практика вейвлет-преобразования. – 1999.
- [7]. Лабутина И.А., Балдина Е.А. Практикум по курсу «Дешифрирование аэрокосмических снимков»: Учебное пособие. М.: Географический факультет МГУ, 2013. 168 с.
- [8]. EarthExplorer, USGS [Электронный ресурс]. – Режим доступа <http://earthexplorer.usgs.gov/>, доступ свободный
- [9]. ERDAS Imagine. Практическое руководство Tour Guide. Norcross, Georgia, USA, 2005 г.
- [10]. Шовенгердт Р.А. Дистанционное зондирование. Модели и методы обработки изображений. Москва: Техносфера, 2010, - 560 с., 32 с. цв. вкл.
- [11]. Using the USGS Landsat 8 Product [Электронный ресурс]. – Режим доступа <https://landsat.usgs.gov/using-usgs-landsat-8-product>, доступ свободный (дата обращения 11.03.2017 г.).
- [12]. Книжников Ю. Ф., Кравцова В.И., Тутубалина О.В. Аэрокосмические методы географических исследований: учебник для студ. Учреждений высш. Проф. Образования – 2-е изд., перераб. и доп. – М.: Издательский центр «Академия», 2011 – 416 с.

## БІЛІМ БЕРУ ҚҰЖАТТАРЫНЫҢ ҰҚСАСТЫҒЫН АНЫҚТАУДА КЛАСТЕРЛІК ТАЛДАУДЫҢ ҚОЛДАНЫЛУЫ

*Кайбасова Динара Женисбековна*  
КеАҚ «Әбілқас Сағынов атындағы Қарағанды техникалық университеті»  
(Қарағанды, Қазақстан), e-mail: [dindgin@mail.ru](mailto:dindgin@mail.ru)

Бұл мақалада оқу жұмыс бағдарламаларының ұқсастығын анықтау үшін білім беру мазмұнын кластерлік талдауды қолдану арқылы жүргізілген зерттеу нәтижелері келтірілген. Эксперименттік жұмыс барысында ІТ-мамандарын даярлау саласындағы 120 құзіреттілікке сәйкес пәндердің 350 оқу жұмыс бағдарламалары талданды. Алынған косинус қашықтық матрицасы бізге ұқсас оқу жұмыс бағдарламаларын анықтауға мүмкіндік берді.

Жоғары оқу орындарының білім беру бағдарламаларында кәсіби құзыреттіліктерін қалыптастыру мақсатында білім беру мазмұнының сапасын арттыру үшін процесті интеллектуалдық қолдау бойынша зерттеулер жүргізілді. Бұл ретте, субъектілерді автоматты түрде шығарып алу және олардың арасындағы қатынастар бар құжаттардағы субъектілердің мәнмәтінін ескере отырып, пәндердің тиісті жұмыс оқу бағдарламаларын анықтауда білім базаларын көп еңбекті қажет етпей қайта өңдеусіз және бейімдеусіз іске асыруға мүмкіндік берді. Зерттеу пәні - оқу нәтижелері мен пән мазмұнын сипаттайтын мәліметтер жиынтығы ретінде анықталған жұмыс оқу бағдарламаларының мазмұны (syllabus). Жұмыс нәтижесінде [1], автор "Ақпараттық жүйелер" білім беру бағдарламасы пәндері бойынша жұмыс оқу бағдарламаларының құжаттарынан мәтіндер корпусын құрды.

Білім мен ережелерге негізделген жүйелердің онтологиялық модельдері [2-4] негізінде білім беру бағдарламаларын қалыптастыруды интеллектуалды қолдаудың қолданыстағы тәсілдері, автоматтандырылған оқу жоспарларын құруға арналған эвристикалық алгоритмдер, сараптамалық бағалау әдістері және когнитивтік карталар еңбек нарығында да, білім беру мазмұны кеңістігінде де өзгерістерді тиімді есепке алуға және жедел бақылауға мүмкіндік бермейді. Өз кезегінде, білім беру бағдарламаларын дайындау бағыттарының барлық қолданыстағы пәндік салалары үшін сарапшылардың күшімен онтологиялық модельдерді, ережелер мен прецеденттер жүйесін қалыптастыру және өзектендіру қажетті дәлдікті қамтамасыз ету үшін пәндік салалардың әрқайсысында сарапшылардың өкілдік құрамын тартуды талап ететін өте көп еңбекті қажет ететін процесс болып табылады.

Кластерлік талдау - бұл объектілерді іріктеу туралы ақпаратты қамтитын деректерді жинауды жүзеге асыратын, содан кейін объектілерді салыстырмалы түрде біртекті топтарға бөлетін көп өлшемді статистикалық процедура [5]. Кластерлік талдауда топтастыру белгілері кейбір "метрика" арқылы біріктіріледі. Олардың кейбіреулері автордың алдыңғы жұмысында қарастырылған [6]. Бұл ретте бір белгі бойынша топтастыру және бірқатар белгілер бойынша кластерлеу бір-біріне келтіріледі. Кластерлеу кезінде қосылатын белгілердің саны бірлікке тең болуы мүмкін. Бұл кластерлеуге бір белгі бойынша топтастыру тапсырмасын береді. Сонымен қатар, белгілерді біріктіру кезінде қолданылатын метрика оларды бір белгіге дейін азайтады, сәйкесінше кластерлерге бөлу осы белгі бойынша топтастыруға тең.

Үлгідегі объектілерді бағалайтын көптеген айнымалыларды анықтау келесідей жасалынды: силлабустағы әрбір термин үшін TF-IDF индексі [7,8] есептелді, олардың жиынтығы құжат векторы болды. Әрі қарай, векторлардан матрица жасалды, онда әр жол жеке құжатты сипаттады. Бұл жағдайда әрбір векторда құжаттар корпусында кездесетін әрбір термин үшін tf-idf көрсеткіштері құрылды. Матрицаны ақпараттық негізде алу үшін кейбір сүзу операциялары жүргізілді, мысалы, ақпараты аз бағандарды алып тастау, яғни

тек бір силлабуста кездесетін терминдер жалпыға ортақ емес деп саналып, олар алынып тасталды.

Енді айнымалыларды (белгілерді) және объектілерді (силлабустарды) таңдау міндеті орындалғаннан кейін, силлабустар арасындағы ұқсастық өлшемінің мәндерін есептеуге көшуге болады.

*Анықтама.*  $s(X_i, X_j) = s_{ij}$  теріс емес нақты функция ұқсастық өлшемі деп аталады, егер:

- 1)  $0 \leq s(X_i, X_j) < 1$  для  $X_i \neq X_j$ ;
- 2)  $s(X_i, X_j) = 1$ ;
- 3)  $s(X_i, X_j) = s(X_j, X_i)$ .

мұндағы  $X_1, X_2, \dots, X_n - p \times n$  өлшемді деректер матрицасы ретінде ұсынылған:

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{p1} & x_{p2} & \dots & x_{pn} \end{pmatrix} = (X_1, X_2, \dots, X_n).$$

Сонымен қатар,  $d(X_i, X_j)$  векторлар жұбы арасындағы қашықтық симметриялық қашықтық матрицасы ретінде ұсынылуы мүмкін:

$$D = \begin{pmatrix} 0 & d_{12} & \dots & d_{1n} \\ d_{21} & 0 & \dots & d_{2n} \\ \dots & \dots & \dots & \dots \\ d_{n1} & d_{n2} & \dots & 0 \end{pmatrix}.$$

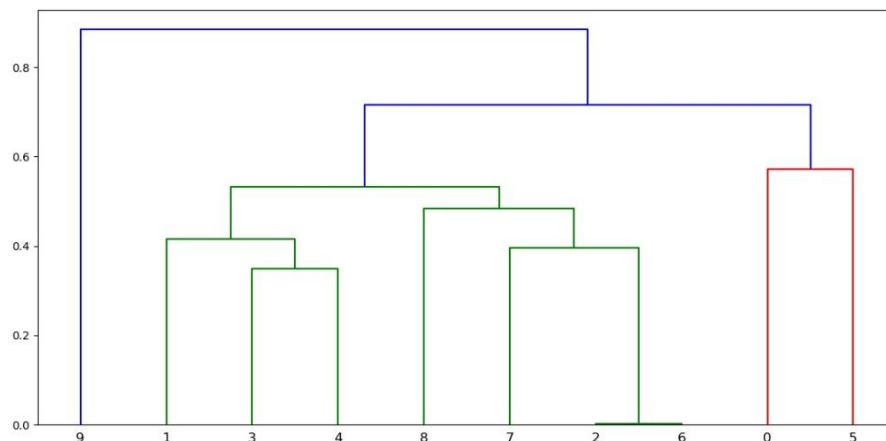
Ұқсастық өлшемдерінің мәндерінің жұптарын төмендегідей ұқсастық матрицасына біріктіруге болады:

$$S = \begin{pmatrix} 1 & s_{12} & \dots & s_{1n} \\ s_{21} & 1 & \dots & s_{2n} \\ \dots & \dots & \dots & \dots \\ s_{n1} & s_{n2} & \dots & 1 \end{pmatrix}.$$

Кластерлеудің дәйекті процесі  $N$  объектілерді қарастырудан басталады; содан кейін ең аз қашықтағы (жақын) екі объект бір кластерге біріктіріліп, кластерлер саны  $n-1$  болады. Процесс барлық  $N$  объектілері барлық объектілерді қамтитын бір кластерге жеткенше қайталанады.

Ағаш диаграммасында нысандар сол жақта тігінен, ал кластерлеу нәтижелері оң жақта орналасқан. Жаңа кластерлердің құрылысына сәйкес келетін қашықтық немесе ұқсастық мәндері дендограмманың жоғарғы жағында көлденең түзу сызықпен бейнеленген.  $N$  нысандарымен берілген кластерлеу процедурасына сәйкес келетін көптеген ағаш диаграммаларын салуға болады, бірақ белгілі бір қашықтық немесе ұқсастық матрицасы үшін бір ғана диаграмма бар-ағаш болады.

Төмендегі 1-суретте көрсетілгендей дендограмма он объектінің ( $n=10$ ) және  $p$  сипаттамаларының (белгілерінің) жағдайына сәйкес келеді. Мұнда 2 және 6 нысандары ең жақын (бір-бірінен ең аз қашықтықта), сондықтан 0,009 жақындық деңгейінде бір кластерге біріктіріледі. 3 және 4 нысандар 0,3703 деңгейінде біріктіріледі. Бұл қадамда 8 кластер бақыланады: (2, 6), (0), (5), (9), (1), (3, 4), (8), (7). Процестің үшінші қадамында келесі кластерлер пайда болды: (7, 2, 6), (1, 3, 4), (9), (8), (0), (5) процестің төртінші қадамында (2, 6, 7, 8), (1, 3, 4), (9), (0), (5) кластерлері пайда болады, олар 0,5001 жақындық деңгейіне сәйкес келеді. Соңында барлық нысандар 0,923 деңгейінде бір кластерге топтастырылады.



Сурет 1– Силлабусты кластерлеу дендрограммасы

Нәтижесінде құжаттардың ұқсастығын сандық түрде анықтайтын әдістер бар екендігі анықталды, олар үлкен корпустарды тез және тиімді өңдей алады және қызықты және өзекті ақпаратты ұсынады. Мәтіндік құжаттарды өңдеудің қолданыстағы әдістерін талдау білім беру бағдарламаларын қалыптастыру мәселесін шешуге қолданылатын бірқатар тәсілдер бар екенін көрсетті.

#### Қолданылған әдебиеттер тізімі

1. Кайбасова Д.Ж. «Предварительная обработка коллекции рабочих учебных программ дисциплин для формирования корпуса текстов» - Вестник КазНИТУ, ISSN 1680-9211, № 6 (136) декабрь, 2019, стр. 541-546.
2. Bakanova A., Letov N.E., Kaibassova D., Kuzmin K.S., Loginov K.V., Shikov A.N. The use of Ontologies in the Development of a Mobile E-Learning Application in the Process of Staff Adaptation, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-2S10, September 2019
3. Chung H., Kim J. An Ontological Approach for Ssemantic Modelling of Curriculum and Syllabus in Higher Education // International Journal of Information and Education Technology. Vol. 6, no. 5. 2016. - pp. 365–369.
4. Oprea M. On the Use of Educational Ontologies as Support Tools for Didactical Activities // *Proceedings of the International Conference on Virtual Learning (ICVL2012)*, Nov. 2012. - pp. 67–73.
5. <https://ru.wikipedia.org>
6. Сокэл Р.Р. Кластер-анализ и классификация: предпосылки и основные направления. Классификация и кластер // Под ред. Дж.Вэн Райзина М: Мир, 1980, с. 57-79
7. Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С. Автоматическая обработка текстов на естественном языке и анализ данных: уч.пособие. М.: НИУ ВШЭ, 2017. – 269 с.
8. D. Kaibassova, L. La, A. Smagulova, L. Lisitsyna, A. Shikov, M. Nurtay/ Methods and algorithms of analyzing syllabuses for educational programs forming intellectual system/ Journal of theoretical and applied information technology, ISSN: 1992-8645, vol.98. no 05, 2020, pp. 876-888



# METHODS FOR DETECTING PATHOLOGY OF THE CHEST ON RADIOGRAPHIC IMAGES

**Shekerbek Ainur, Abdikerimova Gulzira**

*Eurasian National University named after L.N. Gumilyov, Astana, [Shekerbek80@mail.ru](mailto:Shekerbek80@mail.ru)*

*Eurasian National University named after L.N. Gumilyov, Astana, [Gulzira1981@mail.ru](mailto:Gulzira1981@mail.ru)*

## **Annotation**

With the spread of radiation diagnostic methods, the volume of workload on radiologists has increased significantly. The constant process of digitalization in healthcare allows the use of more and more new technologies in various areas of medicine. Chest X-ray is the most commonly used and effective diagnostic method for investigations of serious lung diseases such as cancer, pneumonia, and tuberculosis. The lungs are the largest object of the x-ray. The correct separation of the shape and size of the contours of the lungs is important information for the diagnosis, based on which an intelligent information environment can be created. The article proposes a solution to the problem of clustering pneumonia diseases by using an autocorrelation function to obtain the most accurate result. Thus, a reliable auxiliary tool will be formed for the doctor in the diagnosis of pulmonary radiographs. Image pre-processing and database generation play an important role in creating a correctly functioning automated system.

## **Introduction**

With the spread of radiation diagnostic methods, the volume of workload on radiologists has increased significantly. The constant process of digitalization in healthcare allows the use of more and more new technologies in various areas of medicine. The data taken from the <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia> database was checked by mathematical methods to detect pathologies to further assess the possibility of introducing artificial intelligence into the working practice of a radiologist to optimize work with fluorograms to reduce workload and resource costs. Health-related digital data is expanding from the more obvious and traditional, such as records in a medical record, to sometimes less obvious information about our daily lives, as well as a wide range of data describing the environment in which we live. The constant process of digitalization in healthcare allows the use of more and more new technologies in various areas of medicine [1]. With the spread of radiation diagnostic methods, the volume of workload on radiologists has increased significantly. In 2019, it was estimated that the average radiologist must interpret a fluorogram, radiograph, or one CT or MRI image every 3–4 seconds during an 8-hour workday to meet demand. This increasing number of images requiring interpretation means that the amount of work has increased significantly. With technological advances, radiologists are processing more and more images in a single exam. Reducing the time of work is of great interest not only to reduce the burden on the radiologist, but also to reduce resource costs, thereby improving the economic situation in healthcare [2]. Chest fluorography is one of the most commonly used X-ray methods in the world, currently available to researchers for the digitalization of pathology.

## **Research method**

In this approach, the texture is related to the spatial size of the tone non-derivative elements of the image (the tone non-derivative element is an area of the image with certain "anomalous" features). The value of the autocorrelation function is just the feature that characterizes the size of homogeneous non-derivative elements. The spatial arrangement is characterized by a correlation coefficient, which is a measure of the linear dependence of the brightness of one image element on the brightness of another [3]. Mathematically, the process of calculating the autocorrelation function can be described by the following formula:

$$\rho(x, y) = \frac{\frac{1}{(L_x - |x|) \times (L_y - |y|)} \iint S(u, v) \times S(u + x, v + y) du dv}{\frac{1}{L_x \times L_y} \iint S^2(u, v) du dv},$$

where  $S(u, v)$  is the transparency of the transparencies at the point with coordinates  $(u, v)$ ,  $(x, y)$  is the amount of shift in the  $x$  and  $y$  directions. We also assume that outside the rectangle  $0 \leq u \leq L_x$ ,  $0 \leq v \leq L_y$  the transparency is equal to zero.

### Results and Discussion

During the study, the following results were obtained: the program implemented in Python marked 15 (100%) fluorograms as “normal”, 17 (99%) fluorograms as “pathology” and 1 (1%) as “alarming”. When comparing the processing data of the autocorrelation function and the data of the conclusions of radiologists, it was found that the results of the program and doctors completely coincide. The results obtained indicate the high diagnostic accuracy of the method used, as well as the possibility of using the method to automate the work of a radiologist. During the calculation, images from the database [4] were considered. All images were processed according to figure (1).

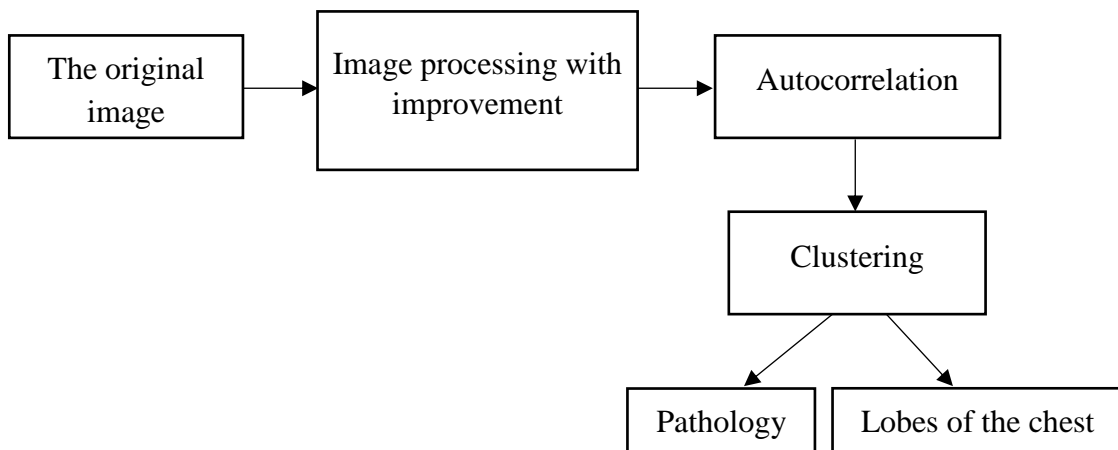


Fig. 1. Stages of determining the pathology of the chest

The program was implemented in Python. The research work carried out according to the algorithm in Figure 1 is aimed at determining the pathology of database images. During the work, the original images were processed, that is, only the images of the lung cavity were selected for the study, and the contrast was enhanced (Fig. 2).

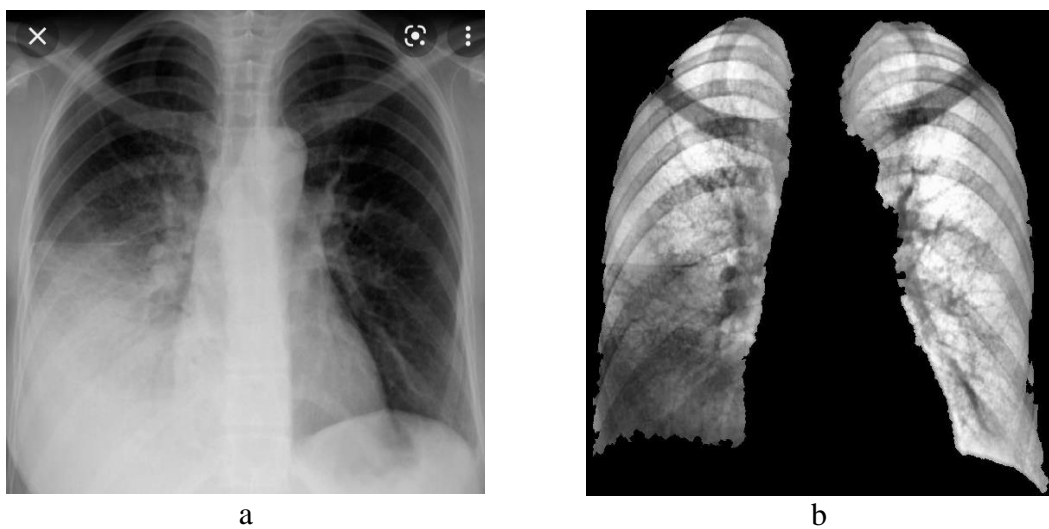


Fig. 2. Original image (a), enhancement of the original image (b)

In each image, "anomalous" textures were distinguished and their percentage was determined (Figure 3).

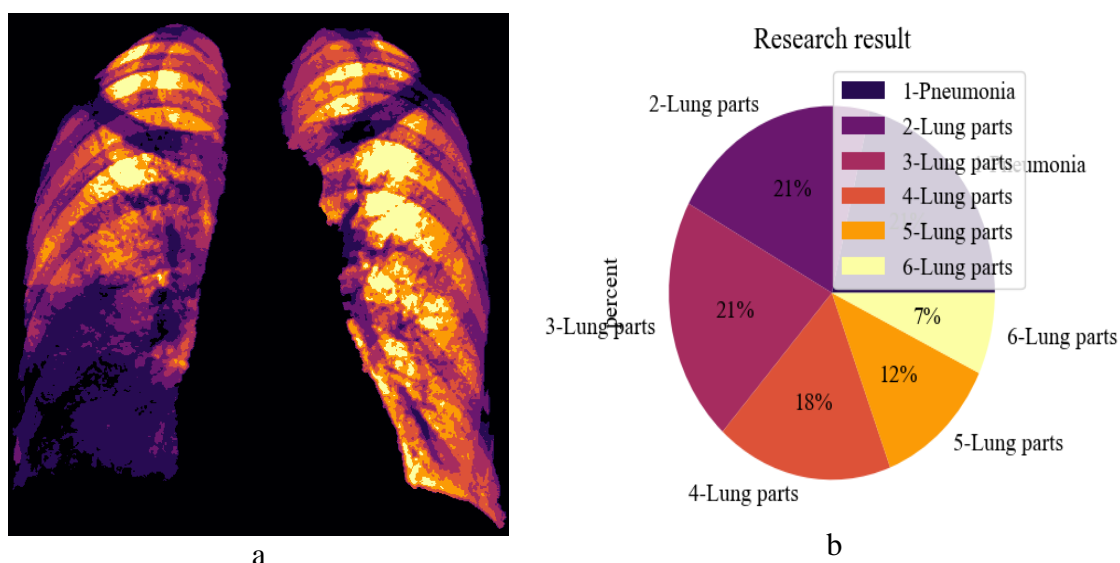


Fig. 3. Clustering result (a), percentage of lungs (b)

33 randomly selected images from the database were processed and an autocorrelation function was applied to them (Table -1).

Table - 1. Autocorrelation function and k-mean values table of values of the result of using the methods

Images title	Autocorrelation function method	K-means method
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

Below is the deviation graph of the normal state of the chest cavity and the appearance of pathology as a result of using the method of autocorrelation functions (Figure 4).

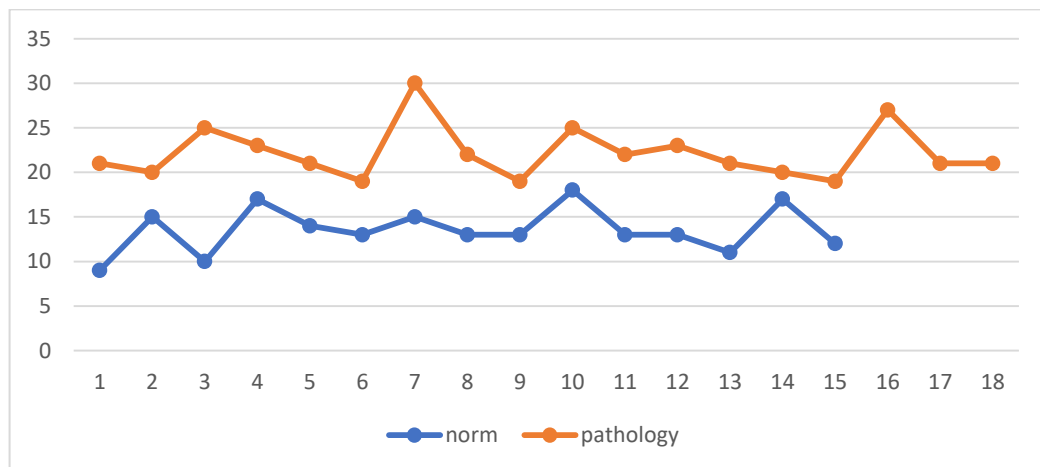


Fig. 4. Visibility graph of normal and pathological images using the autocorrelation function

To see the effectiveness of the autocorrelation function method, the k-means method was considered and its graph is shown below (Figure-5).

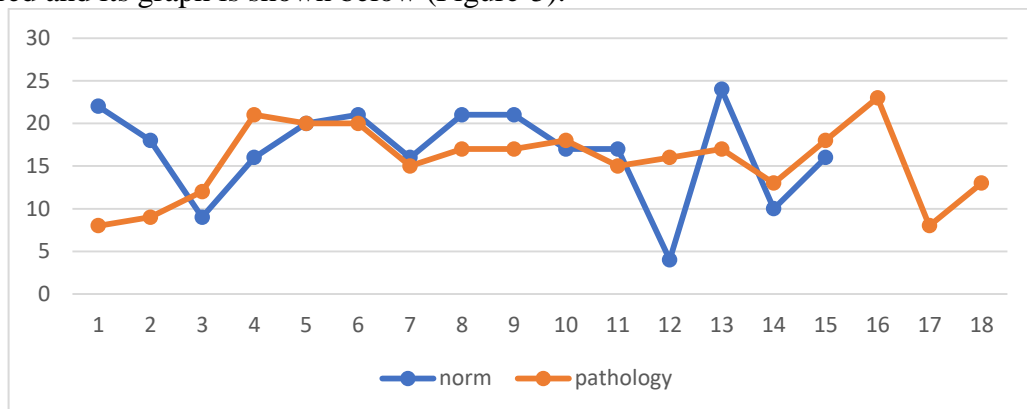


Fig. 5. The result of using the K-means method

## Conclusion

Radiologists play a key role in solving several current challenges of digitalization of health care in the Republic of Kazakhstan, such as creating high-quality data sets for training, determining the clinical problem that needs to be solved, and interpreting the results. Many studies are needed for the further implementation of artificial intelligence in the practice of a radiologist, but now we can say that an automated medicine system can take over part of the workload, facilitating the work of a doctor, as well as improve the economic situation by reducing the costs of the healthcare resource base.

## References

1. Pisaniello, H. L., & Dixon, W. G. (2020). What does digitalization hold for the creation of real-world evidence?. *Rheumatology (Oxford, England)*, 59(1), 39–45. [Электронный ресурс].
2. Driver, C. N., Bowles, B. S., Bartholmai, B. J., & Greenberg-Worisek, A. J. (2020). Artificial Intelligence in Radiology: A Call for Thoughtful Application. *Clinical and translational science*, 13(2), 216–218. [Электронный ресурс].
3. Харалик Р. М. Статистический и структурный подходы к описанию текстур //тииэр. – 1979. – Т. 67. – №. 5. – С. 98.
4. <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>
5. Shekerbek, A. et al. Application of mathematical methods and machine learning algorithms for classification of X-ray images//*Eastern-European Journal of Enterprise Technologies*. – 2022. Т. 3. – №.2. – С. 6–17, <https://doi.org/10.15587/1729-4061.2022.259710>
6. Abdikerimova G. B. et al. Software tools for cell walls segmentation in microphotography //Journal of Theoretical and Applied Information Technology. – 2018. – Т. 96. – №. 15. – С. 4783-4793.
7. Taubayev G. et al. Machine learning algorithms and classification of textures //Journal of Theoretical and Applied Information Technology. – 2020. – Т. 98. – №. 23. – С. 3854-3866.

МАТЕМАТИКАЛЫҚ ЛОГИКА ЖӘНЕ  
КОМПЬЮТЕРЛІК ҒЫЛЫМДАР  
Халықаралық ғылыми конференцияның еңбектері  
2022 жыл, 7-8 қазан, Астана, Қазақстан

MATHEMATICAL LOGIC AND  
COMPUTER SCIENCE  
Proceedings of the International scientific conference  
October 7-8, 2022, Astana, Kazakhstan

МАТЕМАТИЧЕСКАЯ ЛОГИКА И  
КОМПЬЮТЕРНЫЕ НАУКИ  
Труды Международной научной конференции  
7-8 октября 2022 г., Астана, Казахстан

*Конференция және халықаралық ғылыми конференцияның еңбектерін басып шығару Қазақстан Республикасы Ғылым және жоғары білім министрлігінің №AP08855497 «Алгебралық құрылымдардың теоретикалық, -модельдік және алгоритмикалық қасиеттері» гранты есебінен қаржыландырылды.*

*Конференция и издание трудов международной научной конференции профинансированы грантом № AP08855497 «Теоретико-модельные и алгоритмические свойства алгебраических структур» Министерство науки и высшего образования Республики Казахстан.*

*The conference and the publication of the proceedings of the international scientific conference were funded by grant No. AP08855497 “Theoretic-model and algorithmic properties of algebraic structures” of the Ministry of Science and Higher Education of the Republic of Kazakhstan.*

Оригинал-макет подготовлен в ЕНУ им. Л.Н. Гумилева

Z01C0X9, Республика Казахстан,

г. Астана, ул. Сатпаева, 2,

Подписано в печать 17.11.2022 г.

Усл. п. л. 21,31. Тираж 50 экз.

Отпечатано в типографии  
ТОО «ArtPen»

Z01C0X9, Республика Казахстан,

г. Астана, Улица Кажымукан, 18/1.