

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



«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ  
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ» РҚБ

# Х А Б А Р Л А Р Ы

---

---

**ИЗВЕСТИЯ**

РОО «НАЦИОНАЛЬНОЙ  
АКАДЕМИИ НАУК РЕСПУБЛИКИ  
КАЗАХСТАН»

**N E W S**

OF THE ACADEMY OF SCIENCES  
OF THE REPUBLIC OF  
KAZAKHSTAN

**PHYSICO-MATHEMATICAL SERIES**

**4 (352)**

**OCTOBER – DECEMBER 2024**

PUBLISHED SINCE JANUARY 1963

PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

## **БАС РЕДАКТОР:**

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

## **БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ:**

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

## **РЕДАКЦИЯ АЛҚАСЫ:**

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

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

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

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

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

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

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

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

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

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

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

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

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

**«ҚР ҰҒА Хабарлары. Физика және информатика сериясы».**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

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

Тақырыптық бағыты: *физика және ақпараттық коммуникациялық технологиялар сериясы*. Қазіргі уақытта: *«ақпараттық технологиялар» бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*

Мерзімділігі: *жылына 4 рет.*

Тиражы: *300 дана.*

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

## ГЛАВНЫЙ РЕДАКТОР:

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

## ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА:

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

## РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

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

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

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

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

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

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

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

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

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

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

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

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

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

## «Известия НАН РК. Серия физика и информатики».

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

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

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

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

Периодичность: *4 раз в год.*

Тираж: *300 экземпляров.*

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

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

#### **EDITOR IN CHIEF:**

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

#### **DEPUTY EDITOR-IN-CHIEF**

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

#### **EDITORIAL BOARD:**

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

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

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

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

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

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

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

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

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

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

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

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

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

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

**Series of physics and informatics.**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

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

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

Periodicity: *4 times a year.*

Circulation: *300 copies.*

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

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

УДК 28.23.29

©A.Kydyrbekova<sup>1\*</sup>, D. Oralbekova<sup>2</sup>, 2024.

<sup>1</sup>SKU named after M.O. Auezov, Shymkent, Kazakhstan;

<sup>2</sup> Institute of Information and Computational Technology, Almaty, Kazakhstan.

E-mail: [kas.aizat@mail.ru](mailto:kas.aizat@mail.ru).

## **SPEAKER IDENTIFICATION USING DISTRIBUTION-PRESERVING X-VECTOR GENERATION**

**Aizat Kydyrbekova** – SKU named after M.O. Auezov, Shymkent, Kazakhstan, E-mail: [kas.aizat@mail.ru](mailto:kas.aizat@mail.ru). ORCID ID: 0000-0001-5740-4100;

**Dina Oralbekova** - Institute of Information and Computational Technology, Almaty, Kazakhstan, PhD, Email: [dinaoral@mail.ru](mailto:dinaoral@mail.ru). ORCID ID: 0000-0003-4975-6493.

**Abstract.** With the increasing use of voice assistants and conversational language interfaces, serious concerns have arisen regarding the privacy of voice data. In our work, we propose an x-vector-based identification and authentication system to mitigate the risk of attacks on voice data. This method modifies the speaker's pitch and accent information from the original speech signal. In this work, we present a voice recognition system that better supports the natural diversity of voices than previous approaches. By maintaining this diversity and using a generative model to learn and select properties of the x-vector space, we show that this method better captures the distribution of similarities between pseudo-vectors. In our work, we also propose to use a forced inequality that allows the speaker to ensure that the anonymous voice they produce is not too similar to their own voice. The proposed method allows to obtain a natural-sounding anonymous voice in addition to the unidentified voice. However, it provides a relative EER improvement of up to 19.30% for identified anonymous registration-test pairs. We observed that anonymous words have adequate intelligibility and natural speech in addition to good speaker identification. Our method can be easily integrated with others as a matching component of the system and eliminates the need for voice separation for use during matching.

**Keywords:** voice identification, voice privacy, x-vector

*Acknowledgments* - This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP19174298)

©А.С. Кыдырбекова<sup>1\*</sup>, Д.О. Оралбекова<sup>2</sup>, 2024.<sup>1</sup>

М.О. Әуезов атындағы ОҚУ, Шымкент, Қазақстан;

<sup>2</sup> Ақпараттық және есептеу технологиялары институты, Алматы, Қазақстан.

E-mail: [kas.aizat@mail.ru](mailto:kas.aizat@mail.ru).

## ТАРАТУДЫ САҚТАЙТЫН Х-ВЕКТОРЛАР ГЕНЕРАЦИЯСЫН ПАЙДАЛАНЫП ДАУЫСТЫ ИДЕНТИФИКАЦИЯЛАУ

Айзат Кыдырбекова – М.О. Әуезов атындағы ОҚУ, Шымкент, Қазақстан,  
E-mail: [kas.aizat@mail.ru](mailto:kas.aizat@mail.ru). ORCID ID: 0000-0001-5740-4100;

Дина Оралбекова – PhD, Ақпараттық және есептеу технологиялары институты, Алматы,  
Қазақстан, E-mail: [dinaoral@mail.ru](mailto:dinaoral@mail.ru). ORCID ID: 0000-0003-4975-6493.

**Аннотация.** Дауыстық көмекшілер мен сөйлесу тілінің интерфейстерін қолданудың артуы дауыстық деректердің құпиялылығына қатысты үлкен алаңдаушылық тудырды. Мақалада дауыстық деректерге шабуыл жасау қаупін азайту үшін х векторына негізделген сәйкестендіру және аутентификация жүйесін ұсынамыз. Бұл әдіс сөйлеушінің дыбыс деңгейі мен екпін ақпаратын бастапқы сөйлеу сигналынан өзгертеді. Зерттеуде біз алдыңғы тәсілдерге қарағанда дауыстардың табиғи әртүрлілігін жақсырақ қолдайтын дауысты тану жүйесін ұсынамыз. Осы әртүрлілікті сақтай отырып және х-векторлық кеңістіктің қасиеттерін зерттеу және таңдау үшін генеративті модельді қолдана отырып, бұл әдіс жалған векторлар арасындағы ұқсастықтардың таралуын төмендететініне көз жеткіздік. Сондай-ақ динамикке олар шығаратын анонимді дауыстың өз дауысына тым ұқсас болмауын қамтамасыз етуге мүмкіндік беретін мәжбүрлі теңсіздікті пайдалануды ұсынамыз. Ұсынылған әдіс белгісіз дауысқа қосымша табиғи дыбысты анонимді дауысты алуға мүмкіндік береді. Алайда, бұл анықталған анонимді тіркеу-тестілеу жұптары үшін салыстырмалы түрде 19,30% - ға дейін жақсартуды қамтамасыз етеді. Анонимді сөздердің сөйлеушіні жақсы сәйкестендіруден басқа, түсінікті және табиғи сөйлейтінін байқадық. Біздің әдіс жүйенің сәйкес құрамдас бөлігі ретінде басқалармен оңай біріктірілуі мүмкін және сәйкестік кезінде пайдалану үшін дауысты бөлу қажеттілігін болдырмайды.

**Түйін сөздер:** дауысты сәйкестендіру, дауыстық құпиялылық, х-вектор.

©А.С. Кыдырбекова<sup>1\*</sup>, Д.О. Оралбекова<sup>2</sup>, 2024.

<sup>1</sup>ЮКУ имени М.О. Ауезова, Шымкент, Казахстан;

<sup>2</sup>Институт информационных и вычислительных технологий, Алматы, Казахстан.

E-mail: [kas.aizat@mail.ru](mailto:kas.aizat@mail.ru).

## ИДЕНТИФИКАЦИЯ ГОВОРЯЩЕГО С ИСПОЛЬЗОВАНИЕМ ГЕНЕРАЦИИ Х-ВЕКТОРОВ С СОХРАНЕНИЕМ РАСПРЕДЕЛЕНИЯ

Айзат Кыдырбекова – ЮКУ имени М.О.Ауезова, Шымкент, Казахстан, E-mail: [kas.aizat@mail.ru](mailto:kas.aizat@mail.ru), ORCID ID: 0000-0001-5740-4100;



**Дина Оралбекова** – PhD, Институт информационных и вычислительных технологий, Алматы, Казахстан, E-mail: dinaoral@mail.ru, ORCID ID: 0000-0003-4975-6493.

**Аннотация.** С ростом использования голосовых помощников и разговорных языковых интерфейсов возникли серьезные опасения относительно конфиденциальности голосовых данных. В нашей работе мы предлагаем систему идентификации и аутентификации с использованием X-векторов, чтобы снизить риск атак на голосовые данные. Этот метод изменяет информацию о поле и акценте говорящего из исходного речевого сигнала. В этой работе мы представляем систему распознавания голоса, которая лучше поддерживает естественное разнообразие голосов, чем предыдущие подходы. Поддерживая это разнообразие и используя генеративную модель для изучения и выбора свойств пространства X-векторов, показывая, что этот метод лучше отражает распределение сходств между псевдовекторами. В нашей работе также предлагаем использовать принудительное неравенство, которое позволяет говорящему гарантировать, что производимый им анонимный голос не слишком похож на его собственный голос.

Предлагаемый метод позволяет получить естественно звучащий анонимный голос в дополнение к неопознанному голосу. Тем не менее, это обеспечивает относительное улучшение EER до 19,30% для идентифицированных анонимных пар регистрация-тест. Мы заметили, что анонимные слова обладают адекватной разборчивостью и естественностью речи в дополнение к хорошей идентификации говорящего. Наш метод можно легко интегрировать с другими в качестве согласующего компонента системы и устраняет необходимость разделения голосов для использования во время согласования.

**Ключевые слова:** голосовая идентификация, конфиденциальность голоса, X-вектор.

**Introduction.** Speech is widely used as a powerful form of communication between humans and several automated systems. With the advancement and ease of voice biometrics and voice assistants, people are using them for online banking, security, meeting transcription, online shopping and more is used for (Chen, et al, 2018). However, voice data may include passwords, age, gender, health status, geographic origin and more, for example, the speaker has sensitive and personal information that may threaten the privacy of the speaker (Liu, et al, 2018). Therefore, the General Data Protection Regulation pays special attention to the protection of personal data, including speech data (Arik, et al, 2018). There are several ways to protect the speaker data in the speech signal: cryptographic methods, de-identification, pseudonymization and anonymization of speech (Arik, et al, 2018) A system that includes this method is called a voice privacy system (Tomashenko, et al, 2020). Speaker identification is the process of hiding or changing the identity of the speaker so that the speech sounds as if it is spoken by

another speaker (i.e., a pseudo-speaker) without affecting the linguistic content. The unpublished voice of the user is fed to the system and the resulting speech signal is called the test speech of the fake speaker. Some of these studies include methods based on voice transformation, voice mask, voice path length normalization based on voice transformation, and noise addition (i.e., pink noise). J.Qian et al. proposed a de-identification approach using a dual-frequency distortion feature, voice mask. Moreover, it was observed that the age and gender of the speaker could be manipulated to anonymize the speech signal. To achieve this, J.Piribill et al. varied the fundamental frequency (i.e.,  $F_0$ ), the first four formant frequencies (i.e.,  $F_1$  to  $F_4$ ), and the corresponding -3 dB bandwidth (i.e.,  $B_1$  to  $B_4$ ), which carries speaker-specific information, especially the higher formant frequencies. Fang et al. proposed that the speaker embedding (x-vector) is modified after it is separated from the language content (i.e.,  $F_0$ ). The modified x-vector is then used with the source language features to generate anonymized voice using the Neural Source Filter (NSF) model. This approach is presented as the basis of the Voice Privacy Challenge organized during INTERSPEECH 2020. Recently, the study by Mawalim et al. further improved this x-vector and NSF-based approach by modifying the singular value of the x-vector. Both approaches require a pool of x-vectors to obtain an anonymized x-vector. Generative Adversarial Networks (GANs) are mainly proposed to estimate the probability density function of the underlying data. They have many variations that have proven effective in several key areas. One such system is the DNN x&i vector, which is popular in voice transformation. The translatability of DNN features from one domain to another makes it a suitable candidate for the anonymization approach in our proposed study. Therefore, we used the x-vector to adapt the features of male and female speakers. We also studied the effect of combining the proposed approach with the underlying system - identification and authentication.

**Voice Identification.** Speaker privacy is not a new concept, work on protecting and encrypting voices has been around for decades, dating back to the analog processing era (Cox, et al,1987). This physical anonymization has its uses, but approaches that work at this level either do not mask the voice itself, such as by adding a signal to existing audio (Мамырбаев, et al, 2021), or make the audio unintelligible without a decryption key, preventing other legitimate uses.

This work focuses specifically on identification, which suppresses personally identifiable attributes of the speech signal but leaves all other aspects intact. Past work in this area includes using voice mapping to transform voices into a specific speaker identity (Mamyrbayev, т.б., 2021), or using a convolutional neural network (CNN) to transform each speaker into a new anonymized voice created as a function of a set of mapping functions between the original voice and a database of voices (Kalimoldayev, et al, 2020). The level of anonymity offered by previous work is not immediately clear, so the Voice Identification and Authentication Framework was created to evaluate systems with common data sets, protocols, and metrics (Hashimoto, et al, 2016).



The Voice Identification and Authentication Framework provides the foundation for this work and defines a specific goal, data set, and metrics for evaluating and comparing voice anonymization systems. The problem seeks a solution to a scenario in which “speakers want to hide their identity while simultaneously achieving all other subgoals” (Kalimoldayev, et al, 2020). This is done by turning the speaker into a fake speaker, a new identity for the original speaker.

The following system requirements are imposed on the task to achieve the lower goals: (a) the shape of the output speech waveform, (b) maximal concealment of the speaker’s identity, (c) as little distortion of other speech characteristics as possible, possibly (d) all test words from a given speaker spoken by a single pseudo-speaker are guaranteed to appear, and test words from different speakers appear to be spoken by different pseudo-speakers. The task provides a common set of permitted data sets.

**Attack Framework.** The problem assumes that attackers have access to one or more anonymous test words and possibly to the original or hidden registration words for each speaker. The threat model states that the attacker cannot access the identification system used by the user. Although our proposal works within this threat model, we do not believe that this assumption is necessarily the most appropriate for a voice recognition system. In fact, security assumes that the attacker knows the details of the system (Kerckhoff’s principle).

### **Architecture of the x-vector model**

**Rationale.** The design of our system is based on the same approach as in (Jin, et al, 2009). F. Bahmaninezhad et al. proposed three methods for generating pseudo x-vectors: nearest speakers, random sampling, and range sampling. The basic identification and authentication system uses a variant of the last of these methods, choosing the outermost x-vector and then averaging their random selection. The rationale behind this work is that the above methods for generating pseudo x-vectors introduce a bias that causes their distribution to differ from that of the original x-vectors. In particular, we note that the cross-similarity properties of the x-vector distribution should be preserved, i.e., the similarities between the forgeries should have the same behavior.

We replace the base module identification x-vector with the method of the new generation (shown in orange color in the diagram).

The similarity of the fake voice is higher than that of the original voice. Such behavior is a consequence of averaging a set of x-vectors during the period of pseudo-voting and leads to underutilization of the global space of x-vectors, which leads to the following disadvantages:

- low entropy in the space of pseudo x-vectors: anonymous votes are more similar to each other than a pair of original votes,
- Confidentiality is reduced, as it is possible to easily distinguish anonymous voices from real ones.

- The system requires a pool of x-vectors to receive at the anonymization stage, which can lead to the leakage of confidential information, since this information is sent together with the system. Next, we explain how our system improves the generation of x-vectors to preserve the desired similarity properties (Figure 1).

We have a voice identification and authentication system similar to the one described in (Tomashenko, et al, 2020). The generator 2-1-2D, described in (Nautsch, et al, 2020), is used with some modifications to restore signs as close as possible to real ones. Stepwise 2D-layer convolution with sample normalization and valve linear unit (GLU) as an activation function were used for sampling the two lower layers (F.Fang, et al, 2019). Six residual layers of 1D CNN with position normalization and GLU were used for feature transformation. Instead of using pixel-by-pixel interpolation, stepwise transposition layers are used as two layers of upsampling to study the own sampling. The  $1 \times 1$  1-D transformation is used to adjust the size of objects when changing the shape, performed before and after the residual layers. For voice transformation based on CycleGAN, instead of dividing the data set by speakers, we created two classes depending on the gender of the speakers, namely the female class and the male class. Using CycleGAN training, we try to change male speech to female speech and vice versa.

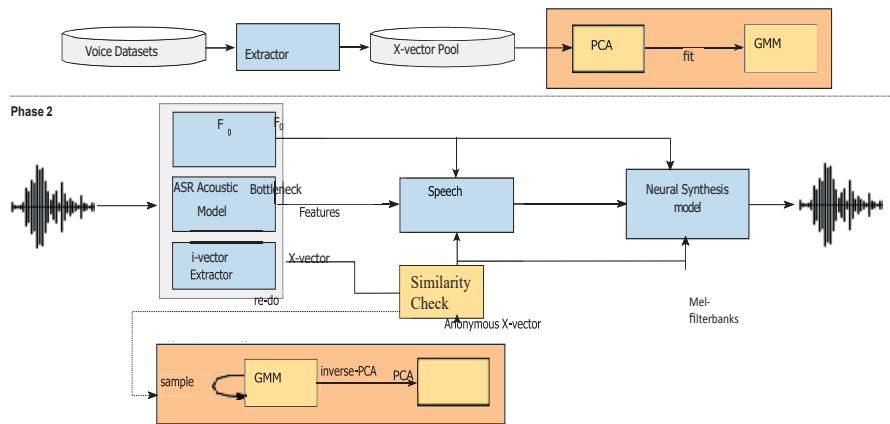


Figure 1. Architecture of voice identification and authentication system.

Let  $M \subset \mathbb{R}^D \times T$  be the space of features of utterances of male speakers, and let  $F \subset \mathbb{R}^D \times T$  be the space of features of utterances of women.  $D$  is the size of the feature vector used, and  $T$  corresponds to the number of speech frames. Our goal is to transform  $m \in M$  into  $f \in F$  and vice versa. Using the CycleGAN network, ten study functions mapping  $G_M \rightarrow F$ :  $M \rightarrow F$  and  $G_F \rightarrow M$ :  $F \rightarrow M$ . To improve the performance of mapping these functions, they are trained on the opponent using the discriminative functions  $D_M: \mathbb{R}^D \times T \rightarrow (0,1)$  and  $D_F: \mathbb{R}^D \times T \rightarrow (0,1)$ .

**Method.** In our method, we focus on improving the generation of the false x-vector base system. In the basic version, three types of characteristics are extracted

for dynamics: fundamental frequency, bottleneck characteristics and x-vector. The x-vector describes the personality of the speaker, while other features only encode the speech content (E.Richardson, et al, 2018).

Taking into account the shortcomings, we improve the generation of the x-vector in two stages. First, we study the properties of the 512-dimensional x-vector space using principal component analysis (PCA) on a large set of x-vector data. Secondly, we fit the generative model to the space reduced by PCA, and for sampling from it we use the Gaussian mixture model (GMM). Using a generative model, we avoid the systematic error introduced by the basic generation of false x-vectors, which generates them by averaging subgroups of population vectors. Whenever a voice needs to be identified, a vector of reduced dimension is randomly selected from the GMM and then returned to the 512-dimensional x-vector space by applying the inverse PCA transformation. We note that the generation of the x-vector can be solved by training a generative-adversarial network, however, it has been shown that GMMs better generalize the captured distributions (R.Shokri, et al, 2017) and do not suffer from membership inference (which can damage system confidentiality guarantees) (A.Nagrani, et al, 2017). Loss of cycle consistency. Since the job of generators is to transform an object from one class to another, we need to recreate the original object using another generator. In other words, the two mapping functions  $G_{M \rightarrow F}$  and  $G_{F \rightarrow M}$  must be inverse functions. This ensures that the two comparisons are mutually exclusive (also included).

$$GF \rightarrow M(G_{M \rightarrow F}(m)) \approx m. \tag{1}$$

$$L_{cyc}(G_{M \rightarrow F}, G_{F \rightarrow M}) = E_{m \sim PM(m)}(\|G_{M \rightarrow F}(G_{M \rightarrow F}(m)) - m\|) + E_{f \sim PF}(f)(G_{M \rightarrow F}(G_{F \rightarrow M}(f)) - f), \tag{2}$$

where  $\|\cdot\|_1$  represents  $L_1$ -norm, and  $E(\cdot)$  represents expectation operator.

**Experiments. Determination of optimal parameters**

In order to evaluate the performance of generating pseudo x-vectors, we analyze the cross-similarity distribution of the generated vectors, varying the number of PCA and GMM components. We use the Kolmogorov-Smirnov test between two distributions to check how close the cross-similarity distributions are between the fake and original x-vectors. The Kolmogorov-Smirnov test quantifies the distance between two empirical cumulative distribution functions, and low scores indicate that these two distributions are similar. For PCA, we focus on three values of the total amount of variance extracted, namely 90%, 95% and 99%.

We will set up the evaluation as follows: we extracted all x-vectors from VoxCeleb1, VoxCeleb2 (20 for men and 30 for women), performed a 50% split of the training test for each gender and trained PCA+GMM. Then we sample from the GMM and apply the inverse PCA transformation to obtain 512 dummy x-vectors. Then we calculate the similarity distribution between pseudo x-vectors

and cosines for the remaining 50% of the test distributions and calculate the statistic KS between the distributions. For GMM, we study the diagonal covariance matrix, set the maximum number of EM iterations to 1000 and the convergence resistance to 10-16.

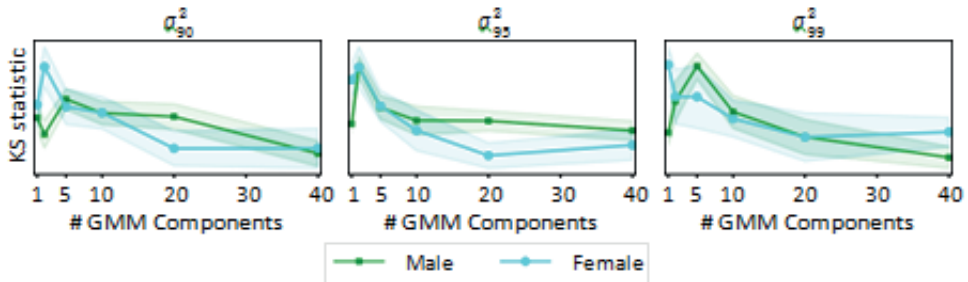


Figure 2. KS statistics between distributions of cross-similarity of fake x-vectors and x-vectors of the system for different stored variances of PCA and numbers of GMM. component floor.

### Results

Figure 2 shows the results of three PCA models with an increase in the number of components used to fit the GMM, and some examples of the obtained eCDF are shown in Figure 4. The use of one or two components of the GMM provides a relatively good fit with a tendency to decrease less pronounced for men, but not for women. The components of GMM increase. Figure 2 shows how close our mock x-vectors are to the cross-similarity distribution in the Vox-Celeb data compared to the original mock x-vectors. Although increasing the number of components usually leads to greater similarity between the distributions, in order to avoid overestimation of the data, VoxCelebs decided to use 95% of the variance retained by the components of PCA and 20 GMM. This allows us to better approximate the 512-dimensional space of x-vectors without requiring too complex a model.

Table 1: Results of agreement for the pre-trained ASVeval model. Results of our anonymization method with 20-component GMM and  $\sigma_2$  PCA without forced distance.

Dataset	Gender	Identification		Development				Test				
		Register	Trial	EER (%)	$\alpha_{min}$	C <sub>lr</sub>	EER (%)	$\alpha_{min}$	C <sub>lr</sub>	EER (%)	$\alpha_{min}$	C <sub>lr</sub>
LibriSpeech	Female	Original	Original	9,1	0,31	41,7	7,8	0,17	25,4			
		Anonymous	Anonymized	42,3(-6,9)	0,98(-0,02)	134,8(-11,2)	41,5(-6,5)	0,98(-0,13)	146,5(-5,8)			
	Male	Original	Original	122	0,03	14,8	1,2	0,05	15,4			
		Anonymous	Anonymized	50,3(-8,2)	0,98(-0,02)	147,8(-20,8)	49,5(-3,8)	0,98(-0,02)	174,1(+6,8)			
VCTK (dfl)	Female	Original	Original	2,8	0,10	1,1	4,8	0,18	1,6			
		Anonymous	Anonymized	46,5(-3,5)	0,95(-0,04)	167,0(+5,2)	43,5(-5,5)	0,97(-0,02)	147,3(+5,8)			
	Male	Original	Original	1,5	0,05	1,3	2,2	0,08	1,8			
		Anonymous	Anonymized	31,1(+1,5)	0,85(+0,02)	13,5(+3,3)	33,1(-1,2)	0,87(-0,02)	17,9(+5,4)			
VCTK (dfl)	Anonymous	Original	Anonymized	45,2(-11,2)	0,98(-0,01)	154,8(-11,8)	46,8(-7,1)	0,97(-0,001)	163,5(-3,2)			
		Anonymous	Anonymized	38,1(+13,1)	0,95(+0,17)	10,5(-8,5)	42,7(+16,9)	0,98(+0,24)	14,6(-1,9)			

## Conclusion

In this paper, we propose a voice recognition system that better supports the natural diversity of voices than previous approaches. Using a generative model to study and select the properties of the space of x-vectors, we support this diversity, showing that this method better reflects the distribution of similarities between fictitious vectors. Such an increase in the diversity of anonymous votes distinguishes them from each other, which is evidenced by the improved results in both registration scenarios and anonymous feedback scenarios. In our work, we also propose to use a forced inequality, which allows the speaker to guarantee that the anonymous voice produced by him is not too similar to his own voice.

We experimentally confirm that the proposed system gives more votes, and evaluate our system using the base level. The results of our tests show a slight decrease in the quality of anonymous voices compared to the original recorded voices, but show a significant improvement when comparing two versions of the same anonymous voice. Our results also show that women perform worse than men, which is a result of the unbalanced data set used for training, and suggests the potential to eliminate this systematic error.

## References

- Chen Z., Zhang Y., Wang Y., Skerrv-Ryan R. et al., “Natural tts synthesis by conditioning wavenet on mel spectrogram pre- dictions,” in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2018, pp. 4779– 4783.
- Liu L.J., Ling Z.H., Yuan-Jiang, Ming-Zhou, “Wavenet vocoder with limited training data for voice conver- sion,” Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, vol. 2018- Septe, no. September, pp. 1983–1987, 2018.
- Arik S., Chen J., Peng K., Ping W., and Zhou Y., “Neural voice cloning with a few samples,” in Advances in Neural Information Processing Systems, 2018, pp. 10 019–10 029.
- Turner H., Lovisotto G., Martinovic I., “Attacking speaker recognition systems with phoneme morphing,” in European Sym- posium on Research in Computer Security. Springer, 2019, pp. 471–492.
- Tomashenko N., Srivastava B., Wang X., Vin-cent E., Nautsch A. , “The VoicePrivacy 2020 Challenge evaluation plan,” 2020.(On-line). Available.
- Cox R.V., Bock D.E., Bauer K.B., Johnston J.D., Synder J. H., “Analog Voice Privacy System.” AT&T Technical Jour- nal, vol. 66, no. 1, pp. 119–131, 1987.
- Kydyrbekova A. S., Mamyrbayev O. Zh., Osman M., Akhmediyarova A. T. Identification and authentication of the user’s voice using DNN capabilities and pages 1-21 of i-vector Cogent Engineering 2020 No. 7(1751557).
- Mamyrbayev O., Kydyrbekova A., Alimkhan K., Oralbekova D., Zhumazhanov B., Nuranbayeva B., (2021). Development of security systems using DNN and I & x-vector classifiers. East European Journal of enterprise technologies, 4/9 (112) 2021, 32-45. doi: <https://doi.org/10.15587/1729-4061.2021.239186> (Scopus, percentage 43);
- Mamyrbayev O., Kydyrbekov A., Oralbekova D., Turdalykyzy T. and A. Bekaristanovna, «complex model based on RNN-T for Speech Recognition in the Kazakh language», 3rd International Conference on Computer Communication and the Internet (ICCCI) (June 25-27, 2021), Tokyo).
- Kalimoldayev M.N., Mamyrbayev O.Zh., Kydyrbekova A.S., Mekebayev N.O., **Algorithms for Detection Gender Using Neural Networks**, International journal of circuits, systems and signal processing, ISSN: 1998-4464, Volume 14, 2020
- Hashimoto K., Yamagishi J., and Echizen I., “Privacy-preserving sound to degrade automatic



speaker verification performance,” ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, vol. 2016-May, pp. 5500–5504, 2016.

Jin Q., Toth A. R., Schultz T., Black A.W., “Speaker de-identification via voice transformation,” Proceedings of the 2009 IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU 2009, pp. 529–533, 2009.

F. Bahmaninezhad, C. Zhang, and J. Hansen, “Convolutional Neural Network Based Speaker De-Identification,” vol. 2016, no. June, pp. 255–260, 2018.

Tomashenko B. M., Srivastava L., Wang X., Vincent E., “Introducing the VoicePrivacy initiative,” 2020.

Fang F., Wang X., Yamagishi J., Echizen I., Todisco M., Evans N., and Bonastre J.-F., “Speaker Anonymization Using X-vector and Neural Waveform Models,” pp. 3–8, 2019. (Online). Available: <http://arxiv.org/abs/1905.13561>

Richardson E. and Weiss Y., “On GANs and GMMs,” in Advances in Neural Information Processing Systems, no. NeurIPS, 2018, pp. 5847–5858.

Shokri R., Stronati M., Song C., and Shmatikov V., “Membership inference attacks against machine learning models,” in 2017 IEEE Symposium on Security and Privacy (SP). IEEE, 2017, pp. 3–18.

Nagrani A., Chung J.S., Zisserman A., “Voxceleb: A large-scale speaker identification dataset,” in Proc. Interspeech 2017, 2017, pp. 2616–2620. (Online). Available: <http://dx.doi.org/10.21437/Interspeech.2017-950>

## CONTENTS

### INFORMATION AND COMMUNICATION TECHNOLOGIES

<b>M. Aitimov, R.U Almenayeva, K.K. Makulov, A.B. Ostayeva, R. Muratkhan</b> APPLICATION OF MACHINE LEARNING METHOD TO ANALYZE AND EXTRACT SEMANTIC STRUCTURES FROM SCIENTIFIC TEXTS.....	5
<b>A.K. Aitim, G.K. Sembina</b> MODELING OF HUMAN BEHAVIOR FOR SMARTPHONE WITH USING MACHINE LEARNING ALGORITHM.....	17
<b>G. Aksholak, A. Bedelbayev, R. Magazov</b> ANALYSIS AND COMPARISON OF MACHINE LEARNING METHODS FOR MALWARE DETECTION.....	29
<b>A.L. Alexeyeva</b> SUBSONIC VIBROTRANSPORT SOLUTIONS OF THE WAVE EQUATION IN SPACES OF DIMENSION $N=1,2,3$ .....	42
<b>K. Bagitova, Sh. Mussiraliyeva, K. Azanbai</b> ANALYSIS OF SYSTEMS FOR RECOGNIZING POLITICAL EXTREMISM IN ONLINE SOCIAL NETWORKS.....	60
<b>A.S. Baegizova, G.I. Mukhamedrakhimova, I. Bapiyev, M.Zh. Bazarova, U.M. Smailova</b> EVALUATING THE EFFECTIVENESS OF MACHINE LEARNING METHODS FOR KEYWORD COVERAGE.....	73
<b>G. Bekmanova, B. Yergesh, G. Yelibayeva, A. Omarbekova, M. Strecker</b> MODELING THE RULES AND CONDITIONS FOR CONDUCTING PRE-ELECTION DEBATES.....	89
<b>M. Bolatbek, M. Sagynay, Sh. Mussiraliyeva</b> USING MACHINE LEARNING METHODS FOR DETECTING DESTRUCTIVE WEB CONTENT IN KAZAKH LANGUAGE.....	99
<b>Y. Golenko, A. Ismailova, K. Kadirkulov, R. Kalendar</b> DEVELOPMENT OF AN ONLINE PLATFORM FOR SEARCHING FOR TANDEM REPEATS USING WHOLE GENOME SEQUENCING.....	112

<b>T. Zhukabayeva, L. Zholshiyeva, N. Karabayev, Sh. Akhmetzhanova</b> A BIBLIOMETRIC ANALYSIS OF EDGE COMPUTING IN INDUSTRIAL INTERNET OF THINGS (IIoT) CYBER-PHYSICAL SYSTEMS.....	123
<b>S.S. Koishybay, N. Meirambekuly, A.E. Kulakaeva, B.A. Kozhakhmetova, A.A. Bulin</b> DEVELOPMENT OF THE DESIGN OF A MULTI-BAND DISCONE ANTENNA.....	138
<b>A. Kydyrbekova, D. Oralbekova</b> SPEAKER IDENTIFICATION USING DISTRIBUTION-PRESERVING X-VECTOR GENERATION.....	152
<b>B. Medetov, A. Nurlankyzy, A. Akhmediyarova, A. Zhetpisbayeva, D. Zhexebay</b> COMPARATIVE ANALYSIS OF THE EFFECTIVENESS OF NEURAL NETWORKS WITHIN THE LOW SNR.....	163
<b>A.A Myrzatay, L.G. Rzaeva, B. Zhumadilla, A.A. Mukhanova, G.A. Uskenbayeva</b> DOUBLE EXPONENTIAL SMOOTHING AND TIME WINDOW METHODS FOR PREDICTIVE LAN MONITORING: ANALYSIS, COMPARISON AND APPLICATION.....	174
<b>L. Naizabayeva, M.N. Satymbekov</b> PREDICTING URBAN SOIL POLLUTION USING MACHINE LEARNING ALGORITHMS.....	194
<b>A.U. Mukhiyadin, U.T. Makhazhanova, A.Z. Alimagambetova, A.A. Mukhanova, A.I. Akmoldina</b> PREDICTING STUDENT LEARNING ENGAGEMENT USING MACHINE LEARNING TECHNIQUES: ANALYSIS OF EDUCATION DATA IN KAZAKHSTAN.....	204
<b>Zh. Tashenova, Zh. Abdugulova, Sh. Amanzholova, E. Nurlybaeva</b> PENETRATION TESTING APPROACHES EMPLOYING THE OPENVAS VULNERABILITY MANAGEMENT UTILITY.....	218
<b>D.B. Tyulemissova, A.K. Shaikhanova, V. Martsenyuk, G.A. Uskenbayeva</b> MODERN APPROACHES TO STUDYING THE DYNAMICS OF INFORMATION FLOW IN SOCIAL MEDIA BASED ON MACHINE LEARNING METHODS.....	231

## МАЗМҰНЫ

### АҚПАРАТТЫҚ-КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР

**М. Айтимов, Р.У Альменаева, К.К. Макулов, А.Б. Остаева, Р. Муратхан**  
ҒЫЛЫМИ МӘТІНДЕРДЕН СЕМАНТИКАЛЫҚ ҚҰРЫЛЫМДАРДЫ  
ТАЛДАУ ЖӘНЕ АЛУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ ӘДІСІН  
ҚОЛДАНУ.....5

**Ә.Қ. Әйтiм, Г.К. Сембина**  
МАШИНАЛЫҚ ОҚУ АЛГОРИТМІН ПАЙДАЛАНЫП СМАРТФОН  
ҮШІН АДАМ МІНЕЗІН МОДЕЛДЕУ.....17

**Г.И. Ақшолақ, А.А. Бедельбаев, Р.С. Мағазов**  
ЗИЯНДЫ БАҒДАРЛАМАЛАРДЫ АНЫҚТАУҒА АРНАЛҒАН  
МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН ТАЛДАУ ЖӘНЕ САЛЫСТЫРУ.....29

**А.Л. Алексеева**  
N=1,2,3 ӨЛШЕМДІ КЕҢІСТІГІНДЕГІ ТОЛҚЫНДЫҚ ТЕҢДЕУДІҢ  
ДЫБЫСҚА ДЕЙІНГІ ДІРІЛКӨЛІКТІК ШЕШІМДЕРІ.....42

**Қ.Б. Бағитова, Ш.Ж. Мусиралиева, Қ. Азанбай**  
ӘЛЕУМЕТТІК ЖЕЛІЛЕРДЕГІ САЯСИ ЭКСТРЕМИЗМДІ ОНЛАЙН ТАҢУ  
ЖҮЙЕЛЕРІН ТАЛДАУ.....60

**А.С. Баегизова, Г.И. Мухамедрахимова, И.М. Бапиев, М.Ж. Базарова,  
У.М. Смайлова**  
ТҮЙІН СӨЗДЕРДІ ҚАМТУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІНІҢ  
ТИІМДІЛІГІН БАҒАЛАУ.....73

**Г.Т. Бекманова, Б.Ж. Ергеш, Г.К. Елибаева, А.С. Омарбекова,  
M. Strecker**  
САЙЛАУ АЛДЫНДАҒЫ ПІКІРТАЛАСТАРДЫ ӨТКІЗУ ЕРЕЖЕЛЕРІ  
МЕН ШАРТТАРЫН МОДЕЛЬДЕУ.....89

**М.А. Болатбек, М.Сағынай, Ш.Ж. Мусиралиева**  
ҚАЗАҚ ТІЛІНДЕГІ ДЕСТРУКТИВТІ ВЕБ-КОНТЕНТТІ АНЫҚТАУ ҮШІН  
МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН ҚОЛДАНУ.....99

**Е.С. Голенко, А.А. Исмаилова, К.К. Кадиркулов, Р.Н. Календарь**  
ТОЛЫҚ ГЕНОМДЫҚ СЕКВЕНИРЛЕУДЕ ТАНДЕМДІК  
ҚАЙТАЛАНУЛАРДЫ ІЗДЕУ ҮШІН ОНЛАЙН ПЛАТФОРМАСЫН  
ӘЗІРЛЕУ.....112

- Т. Жукабаева, Л. Жолшиева, Н. Карабаев, Ш. Ахметжанова**  
ӨНДІРІСТІК ЗАТТАР ИНТЕРНЕТІ (IoT) КИБЕРФИЗИКАЛЫҚ  
ЖҮЙЕЛЕРІНДЕ ШЕТКІ ЕСЕПТЕУЛЕРДІ ҚОЛДАНУҒА  
БИБЛИОМЕТРИЯЛЫҚ ТАЛДАУ.....123
- С.С. Қойшыбай, Н. Мейрамбекұлы, А.Е. Кулакаева, Б.А. Кожаметова,  
А.А. Булин**  
КӨПДИАПАЗОНДЫДИСКОНУСТЫҚАНТЕННАКОНСТРУКЦИЯСЫН  
ӘЗІРЛЕУ.....138
- А.С. Кыдырбекова, Д.О. Оралбекова**  
ТАРАТУДЫ САҚТАЙТЫН Х-ВЕКТОРЛАР ГЕНЕРАЦИЯСЫН  
ПАЙДАЛАНЫП ДАУЫСТЫ ИДЕНТИФИКАЦИЯЛАУ.....152
- Б. Медетов, А. Нурланқызы, А. Ахмедиярова, А. Жетписбаева, Д. Жексебай**  
СИГНАЛШУЫЛ ҚАТЫНАСЫ ТӨМЕН ЖАҒДАЙДА НЕЙРОНДЫҚ  
ЖЕЛЛЕРДІҢ ТИІМДІЛІГІНЕ САЛЫСТЫРМАЛЫ ТАЛДАУ ЖАСАУ.....163
- А.А. Мырзатай, Л.Г. Рзаева, Б. Жұмаділла, А.А. Муханова,  
Г.А. Ускенбаева**  
ЖЕРГІЛІКТІ ЖЕЛІНІ БОЛЖАМДЫ БАҚЫЛАУҒА АРНАЛҒАН ҚОС  
ЭКСПОНЕНЦИАЛДЫ ТЕГІСТЕУ ЖӘНЕ УАҚЫТ ТЕРЕЗЕЛЕРІНІҢ  
ӘДІСТЕРІ: ТАЛДАУ, САЛЫСТЫРУ ЖӘНЕ ҚОЛДАНУ.....174
- Л. Найзабаева, М.Н. Сатымбеков**  
МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМДЕРІН ПАЙДАЛАҢУ АРҚЫЛЫ  
ҚАЛА ТОПЫРАҒЫНЫҢ ЛАСТАНУЫН БОЛЖАУ.....194
- А.Ұ. Мұхиядин, У.Т. Махажанова, А.З. Алимагамбетова, А.А.Муханова,  
А.И. Акмолдина**  
МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН ПАЙДАЛАНА ОТЫРЫП,  
ОҚУШЫЛАРДЫҢ БІЛІМ АЛУҒА ЫНТАСЫН БОЛЖАУ:  
ҚАЗАҚСТАНДАҒЫ БІЛІМ БЕРУ ДЕРЕКТЕРІН ТАЛДАУ.....204
- Ж.М. Ташенова, Ж.К. Абдугулова, Ш.А. Аманжолова, Э. Нурлыбаева**  
OPENVAS ОСАЛДЫҒЫН БАСҚАРУ УТИЛИТАСЫН ҚОЛДАНА  
ОТЫРЫП, ЕНУДІ ТЕСТІЛЕУ ТӘСІЛДЕРІ.....218
- Д.Б. Тюлемисова, А.К. Шайханова, В.П. Мартценюк, Г.А. Ускенбаева,  
Г.В. Бекешева**  
МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІНЕ НЕГІЗДЕЛГЕН ӘЛЕУМЕТТІК  
ЖЕЛЛЕРДЕГІ АҚПАРАТ АҒЫНЫНЫҢ ДИНАМИКАСЫН ЗЕРТТЕУДІҢ  
ЗАМАНАУИ ТӘСІЛДЕРІ.....231

## СОДЕРЖАНИЕ

### ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

<b>М. Айтимов, Р.У Альменаева, К.К. Макулов, А.Б. Остаева, Р. Муратхан</b> ПРИМЕНЕНИЕ МЕТОДА МАШИННОГО ОБУЧЕНИЯ ДЛЯ АНАЛИЗА И ИЗВЛЕЧЕНИЯ СЕМАНТИЧЕСКИХ СТРУКТУР ИЗ НАУЧНЫХ ТЕКСТОВ.....	5
<b>А.К. Айтим, Г.К. Сембина</b> МОДЕЛИРОВАНИЕ ЧЕЛОВЕЧЕСКОГО ПОВЕДЕНИЯ ДЛЯ СМАРТФОНА С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМА МАШИННОГО ОБУЧЕНИЯ.....	17
<b>Г.И. Акшолок, А.А. Бедельбаев, Р.С. Магазов</b> АНАЛИЗ И СРАВНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ ВРЕДОНОСНОГО ПО.....	29
<b>Л.А. Алексеева</b> ДОЗВУКОВЫЕ ВИБРОТРАНСПОРТНЫЕ РЕШЕНИЯ ВОЛНОВОГО УРАВНЕНИЯ В ПРОСТРАНСТВАХ РАЗМЕРНОСТИ $N=1,2,3$ .....	42
<b>К.Б. Багитова, Ш.Ж. Мусиралиева, К. Азанбай</b> АНАЛИЗ СИСТЕМ РАСПОЗНАВАНИЯ ПОЛИТИЧЕСКОГО ЭКСТРЕМИЗМА В СОЦИАЛЬНЫХ СЕТЯХ ОНЛАЙН.....	60
<b>А.С. Баегизова, Г.И. Мухамедрахимова, И.М. Бапиев, М.Ж. Базарова, У.М. Смайлова</b> ОЦЕНКА ЭФФЕКТИВНОСТИ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОХВАТА КЛЮЧЕВЫХ СЛОВ.....	73
<b>Г.Т. Бекманова, Б.Ж. Ергеш, Г.К. Елибаева, А.С. Омарбекова, М. Strecker</b> МОДЕЛИРОВАНИЕ ПРАВИЛ И УСЛОВИЙ ПРОВЕДЕНИЯ ПРЕДВЫБОРНЫХ ДЕБАТОВ.....	89
<b>М.А. Болатбек, М. Сагынай, Ш.Ж. Мусиралиева</b> ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ ДЕСТРУКТИВНОГО ВЕБ-КОНТЕНТА НА КАЗАХСКОМ ЯЗЫКЕ.....	99
<b>Е.С. Голенко, А.А. Исмаилова, К.К. Кадиркулов, Р.Н. Календарь</b> РАЗРАБОТКА ОНЛАЙН-ПЛАТФОРМЫ ДЛЯ ПОИСКА ТАНДЕМНЫХ ПОВТОРОВ ПРИ ПОЛНОГЕНОМНОМ СЕКВЕНИРОВАНИИ.....	112



<b>Т. Жукабаева, Л. Жолшиева, Н. Карабаев, Ш. Ахметжанова</b> БИБЛИОМЕТРИЧЕСКИЙ АНАЛИЗ ПРИМЕНЕНИЯ ГРАНИЧНЫХ ВЫЧИСЛЕНИЙ В КИБЕРФИЗИЧЕСКИХ СИСТЕМАХ ПРОМЫШЛЕННОГО ИНТЕРНЕТА ВЕЩЕЙ (IIoT).....	123
<b>С.С. Койшыбай, Н. Мейрамбекұлы, А.Е. Кулакаева, Б.А. Кожаметова, А.А. Булин</b> РАЗРАБОТКА КОНСТРУКЦИИ МНОГОДИАПАЗОННОЙ ДИСКОНУСНОЙ АНТЕННЫ.....	138
<b>А.С. Кыдырбекова, Д.О. Оралбекова</b> ИДЕНТИФИКАЦИЯ ГОВОРЯЩЕГО С ИСПОЛЬЗОВАНИЕМ ГЕНЕРАЦИИ X-ВЕКТОРОВ С СОХРАНЕНИЕМ РАСПРЕДЕЛЕНИЯ...152	152
<b>Б. Медетов, А. Нурланкызы, А. Ахмедиярова, А. Жетписбаева, Д. Жексебай</b> СРАВНИТЕЛЬНЫЙ АНАЛИЗ ЭФФЕКТИВНОСТИ НЕЙРОННЫХ СЕТЕЙ ПРИ НИЗКОМ ЗНАЧЕНИИ ОТНОШЕНИЯ С/Ш.....	163
<b>А.А. Мырзатай, Л.Г. Рзаева, Б. Жұмаділла, А.А. Муханова, Г.А. Ускенбаева</b> МЕТОДЫ ДВОЙНОГО ЭКСПОНЕНЦИАЛЬНОГО СГЛАЖИВАНИЯ И ВРЕМЕННЫХ ОКОН ДЛЯ ПРЕДИКТИВНОГО МОНИТОРИНГА ЛВС: АНАЛИЗ, СРАВНЕНИЕ И ПРИМЕНЕНИЕ.....	174
<b>Л. Найзабаева, М.Н. Сатымбеков</b> ПРОГНОЗИРОВАНИЕ ЗАГРЯЗНЕНИЯ ГОРОДСКОЙ ПОЧВЫ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ.....	194
<b>А.У. Мухиядин, У.Т. Махажанов, А.З. Алимагамбетова, А.А. Муханова, А.И. Акмолдина</b> ПРОГНОЗИРОВАНИЕ МОТИВАЦИИ УЧАЩИХСЯ К ОБУЧЕНИЮ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ: АНАЛИЗ ДАННЫХ ОБ ОБРАЗОВАНИИ В КАЗАХСТАНЕ.....	204
<b>Ж.М. Ташенова, Ж.К. Абдугулова, Ш.А. Аманжолова, Э. Нурлыбаева</b> ПОДХОДЫ К ТЕСТИРОВАНИЮ НА ПРОНИКНОВЕНИЕ С ИСПОЛЬЗОВАНИЕМ УТИЛИТЫ УПРАВЛЕНИЯ УЯЗВИМОСТЯМИ OPENVAS.....	218
<b>Д.Б. Тюлемисова, А.К. Шайханова, В. Мартценюк, Г.А. Ускенбаева, Г.В. Бекешева</b> СОВРЕМЕННЫЕ ПОДХОДЫ К ИЗУЧЕНИЮ ДИНАМИКИ ИНФОРМАЦИОННОГО ПОТОКА В СОЦИАЛЬНЫХ МЕДИА НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ.....	231

**Publication Ethics and Publication Malpractice  
the journals of the National Academy of Sciences of the Republic of Kazakhstan**

For information on Ethics in publishing and Ethical guidelines for journal publication see <http://www.elsevier.com/publishingethics> and <http://www.elsevier.com/journal-authors/ethics>.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see <http://www.elsevier.com/postingpolicy>), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct ([http://publicationethics.org/files/u2/New\\_Code.pdf](http://publicationethics.org/files/u2/New_Code.pdf)). To verify originality, your article may be checked by the Cross Check originality detection service <http://www.elsevier.com/editors/plagdetect>.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:

**[www.nauka-nanrk.kz](http://www.nauka-nanrk.kz)**

**<http://physics-mathematics.kz/index.php/en/archive>**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

Директор отдела издания научных журналов НАН РК *А. Ботанқызы*

Редакторы: *Д.С. Аленов, Ж.Ш. Әден*

Верстка на компьютере *Г.Д. Жадыранова*

Подписано в печать 2.12.2024.

Формат 60x881/8. Бумага офсетная. Печать – ризограф.

16,0 п.л. Тираж 300. Заказ 4.