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

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

№2
2026

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

CENTRAL ASIAN ACADEMIC
RESEARCH CENTER

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

2 (358)

APRIL – JUNE 2026

PUBLISHED SINCE JANUARY 1963
PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

Academic Scientific Journal of Computer Science
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Owner: «Central Asian Academic Research Center» LLP (Almaty).
Certificate **№ KZ77VPY00121154** on the re-registration of the periodical printed and online publication of the information agency, issued on **05.06.2025** by the Republican State Institution «Information Committee» of the Ministry of Culture and Information of the Republic of Kazakhstan
Subject area*: information and communication technologies.*
Currently: *included in the list of journals recommended by the CCSES MSHE RK in the direction of «Information and communication technologies».*
Periodicity: *4 times a year.*
http://www.physico-mathematical.kz/index.php/en/

Academic Scientific Journal of Computer Science
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Меншіктеуші: «Орталық Азия академиялық ғылыми орталығы» ЖШС (Алматы).
Ақпарат агенттігінің мерзімді баспасөз басылымын, ақпарат агенттігін және желілік басылымды қайта есепке қою туралы ҚР Мәдениет және Ақпарат министрлігі «Ақпарат комитеті» Республикалық мемлекеттік мекемесі **05.06.2025** ж. берген **№ KZ77VPY00121154** Куәлік.
Тақырыптық бағыты: *ақпараттық-коммуникациялық технологиялар*
Қазіргі уақытта: *«ақпараттық-коммуникациялық технологиялар» бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*
Мерзімділігі: *жылына 4 рет.*
http://www.physico-mathematical.kz/index.php/en/

Academic Scientific Journal of Computer Science
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Собственник: *ТОО «Центрально-азиатский академический научный центр» (г. Алматы).*
Свидетельство о постановке на переучет периодического печатного издания, информационного агентства и сетевого издания **№ KZ77VPY00121154.** Дата выдачи **05.06.2025**
Тематическая направленность: *информационно-коммуникационные технологии.*
В настоящее время: *вошел в список журналов, рекомендованных КОКСНВО МНВО РК по направлению «информационно-коммуникационные технологии».*
Периодичность: *4 раза в год.*
http://www.physico-mathematical.kz/index.php/en/

CONTENTS

COMPUTER SCIENCE

МАЗМҰНЫ

КОМПЬЮТЕРЛІК ҒЫЛЫМДАР

СОДЕРЖАНИЕ

КОМПЬЮТЕРНЫЕ НАУКИ

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 2.
Namber 358 (2026). 11–29

https://doi.org/10.32014/2026.2518-1726.424

IRSTI 20.23.17
UDC 004.056.53

Abduraimova B.K.[1], Toleukhan A.B.[1*], Sapakova S.Z.[2], Abisheva A.A.[3], 2026.
[1]L.N. Gumilyov Eurasian National University, Astana, Kazakhstan;
[2]International information technologies university, Almaty, Kazakhstan;
[3]Kazakh University of Technology and Business named after K. Kulazhanov, Astana, Kazakhstan.
E-mail: aminkatoleukhanova@gmail.com

DEVELOPMENT OF EARLY CYBERATTACK DETECTION METHOD USING CNN-LSTM FOR IOT

Abduraimova Bayan — Candidate of Technical Sciences, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,
E-mail: abduraimovabk@mail.ru, ORCID: https://orcid.org/0000-0003-3913-1895;
Toleukhan Amina — L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,
E-mail: aminkatoleukhanova@gmail.com, ORCID ID: https://orcid.org/0009-0008-6900-7813;
Sapakova Saya — associate professor, International information technologies university, Almaty, Kazakhstan,
E-mail: s.sapakova@iitu.edu.kz, ORCID: https://orcid.org/0000-0001-6541-6806;
Abisheva Aigul — Kazakh University of Technology and Business named after K. Kulazhanov, Astana, Kazakhstan,
E-mail: aiqul.abisheva@gmail.com, ORCID: https://orcid.org/0000-0002-5800-9134.

Abstract. The rapid development of the Internet of Things (IoT) and the widespread adoption of network-connected devices have led to a significant increase in network traffic and the number of cyberattacks. Traditional intrusion detection methods often require analysis of the entire network flow and extensive computing resources, limiting their application in IoT environments where computing power and memory are severely limited. Therefore, the development of effective approaches for the early detection of cyberattacks has become an important area of research. Such approaches enable security systems to detect potential threats early in the network interaction process and respond more quickly to malicious activity. This paper presents a method for early cyberattack detection based on a hybrid deep learning architecture that combines convolutional neural networks (CNNs) and long short-term memory (LSTM) networks. The CNN component extracts spatial features from network traffic data, while the LSTM layers identify temporal

dependencies in packet sequences. To improve the information yield of traffic analysis, an expanded set of statistical-dynamic features is proposed, including the entropy of packet size distribution, inter-packet intervals, and the variation coefficient. Experimental results show that the proposed approach provides higher classification accuracy compared to traditional machine learning algorithms such as support vector machines (SVM), random forests, and multilayer perceptrons (MLP). Analysis of the first ten packets of a network connection significantly reduces attack detection time, which is especially important for real-time systems. Furthermore, an optimized version of the model, adapted for the IoT environment, reduces computational load while maintaining high detection accuracy and practical applicability in modern network infrastructures.

Keywords: cyberattacks, deep learning, CNN-LSTM, IoT, network traffic, intrusion detection system (IDS), model optimization

For citations: Abduraimova B.K., Toleukhan A.B., Sapakova S.Z., Abisheva A.A. Development of early cyberattack detection method using cnn-lstm for IOT. Academic Scientific Journal of Computer Science, 2026. — No.2. – P. 11-29. DOI: https://doi.org/10.32014/2026.2518-1726.424

Абдураимова Б.К. [1], Төлеухан Ә.Б.[1*], Сапакова С.З. [2], Абишева А.А. [3], 2026.

[1]Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан;
[2]Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан;
[3]Қ. Құлажанов атындағы Қазақ технология және бизнес университеті, Астана, Қазақстан.
E-mail: aminkatoleukhanova@gmail.com

КИБЕРШАБУЛДАРДЫ ЕРТЕ АНЫҚТАУ ӘДІСІН CNN-LSTM НЕГІЗІНДЕ ДАМЫТУ (IOT ҮШІН)

Абдураимова Баян — техника ғылымдарының кандидаты, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail: abduraimovabk@mail.ru, ORCID: https://orcid.org/0000-0003-3913-1895;
Төлеухан Әмина — Л.Н. Гумилева атындағы Еуразия ұлтық университеті, Астана, Қазақстан,
E-mail: aminkatoleukhanova@gmail.com, ORCID ID: https://orcid.org/0009-0008-6900-7813;
Сапакова Сая — физика-математика ғылымдарының кандидаты, компьютерлік инженерия кафедрасының қауымдастырылған профессоры, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан,
E-mail: s.sapakova@iitu.edu.kz, ORCID: https://orcid.org/0000-0001-6541-6806;
Абишева Айгуль — сеньор-лектор, Қ. Құлажанов атындағы Қазақ технология және бизнес университеті, Астана, Қазақстан,
E-mail: aiqul.abisheva@gmail.com, ORCID: https://orcid.org/0000-0002-5800-9134.

Аннотация. Заттар интернетінің (IoT) жылдам дамуы және желіге қосылған құрылғылардың кеңінен қолданылуы желілік трафиктің және кибершабуыл-

дар санының айтарлықтай өсуіне әкелді. Дәстүрлі басып кіруді анықтау әдістері көбінесе бүкіл желі ағынын және кең есептеу ресурстарын талдауды талап етеді, бұл оларды есептеу қуаты мен жады өте шектеулі IoT орталарында қолдануды шектейді. Сондықтан кибершабуылдарды ерте анықтаудың тиімді тәсілдерін әзірлеу зерттеудің маңызды саласына айналды. Мұндай тәсілдер қауіпсіздік жүйелеріне желілік өзара әрекеттесу процесінің басында әлеуетті қауіптерді анықтауға және зиянды әрекетке тезірек жауап беруге мүмкіндік береді. Бұл мақалада конволюциялық нейрондық желілер (CNN) мен ұзақ мерзімді жад (LSTM) желілерін біріктіретін гибридті терең оқыту архитектурасына негізделген кибершабуылды ерте анықтау әдісі ұсынылған. CNN компоненті желілік трафик деректерінен кеңістіктік ерекшеліктерді алады, ал LSTM қабаттары пакеттік тізбектердегі уақытша тәуелділіктерді анықтайды. Трафикті талдаудың ақпараттық өнімділігін жақсарту үшін пакет өлшемінің таралу энтропиясын, пакеттер арасындағы аралықтарды және вариация коэффициентін қоса алғанда, статистикалық-динамикалық ерекшеліктердің кеңейтілген жиынтығы ұсынылады. Тәжірибелік нәтижелер ұсынылған тәсілдің тірек векторлық машиналар (SVM), кездейсоқ ормандар және көп қабатты перцептрондар (MLP) сияқты дәстүрлі машиналық оқыту алгоритмдерімен салыстырғанда жоғары жіктеу дәлдігін қамтамасыз ететінін көрсетеді. Желілік қосылымның алғашқы он пакетін талдау шабуылды анықтау уақытын айтарлықтай қысқартады, бұл әсіресе нақты уақыт жүйелері үшін маңызды. Сонымен қатар, IoT ортасына бейімделген модельдің оңтайландырылған нұсқасы заманауи желілік инфрақұрылымдарда жоғары анықтау дәлдігін және практикалық қолданылуын сақтай отырып, есептеу жүктемесін азайтады.

Түйін сөздер: кибершабуылдар, терең оқыту, CNN-LSTM, IoT, желілік трафик, шабуылды анықтау жүйесі (IDS), модельді оңтайландыру

Абдураимова Б.К.[1], Толеухан А.Б.[1], Сапакова С.З.[2], Абишева А.А.[3], 2026.
[1]Евразийский национальный университет имени Л.Н. Гумилёва,
Астана, Казахстан;
[2]Международный университет информационных технологий,
Алматы, Казахстан;
[3]Казахский университет технологии и бизнеса имени К. Кулажанова,
Астана, Казахстан.
E-mail: aminkatoleukhanova@gmail.com

РАЗРАБОТКА МЕТОДА РАННЕГО ОБНАРУЖЕНИЯ КИБЕРАТАК НА ОСНОВЕ CNN-LSTM ДЛЯ IOT

Абдураимова Баян — кандидат технических наук, Евразийский Национальный университет имени Л.Н. Гумилева., Астана, Казахстан,
E-mail: abduraimovabk@mail.ru, ORCID: https://orcid.org/0000-0003-3913-1895;

Толеухан Амина — ЕНУ им. Л.Н. Гумилева, Астана, Казахстан,
E-mail: aminkatoleukhanova@gmail.com, ORCID ID: https://orcid.org/0009-0008-6900-7813;
Сапакова Сая — кандидат физ.-мат. наук, ассоциированный профессор кафедры компьютерной инженерии, Международный университет информационных технологий, Алматы, Казахстан,
E-mail: s.sapakova@iitu.edu.kz, ORCID: https://orcid.org/0000-0001-6541-6806;
Абишева Айгуль — сеньор-лектор, Казахский университет технологии и бизнеса имени К. Кулажанова, Астана, Казахстан,
E-mail: aiqul.abisheva@gmail.com, ORCID: https://orcid.org/0000-0002-5800-9134.

Аннотация. *Актуальность***.** Быстрое развитие Интернета вещей (IoT) и широкое распространение сетевых устройств привели к значительному увеличению объема сетевого трафика и числа кибератак. Традиционные методы обнаружения вторжений часто требуют анализа всего сетевого потока и значительных вычислительных ресурсов, что ограничивает их применение в IoT-средах с ограниченными вычислительными мощностями и памятью. В связи с этим разработка эффективных подходов к раннему обнаружению кибератак является актуальной научно-практической задачей. *Цель***.** Разработать метод раннего обнаружения кибератак на основе гибридной архитектуры глубокого обучения CNN-LSTM, обеспечивающий своевременное выявление вредоносной активности в IoT-сетях при сниженной вычислительной нагрузке. *Методы***.** В статье предложен метод раннего обнаружения кибератак, основанный на гибридной архитектуре глубокого обучения, сочетающей сверточные нейронные сети (CNN) и сети долговременной краткосрочной памяти (LSTM). Компонент CNN используется для извлечения пространственных признаков из данных сетевого трафика, а слои LSTM позволяют выявлять временные зависимости в последовательностях пакетов. Для повышения эффективности анализа трафика предложен расширенный набор статистико-динамических характеристик, включающий энтропию распределения размеров пакетов, интервалы между пакетами и коэффициент вариации. *Результаты и выводы.* Экспериментальные результаты показали, что предложенный подход обеспечивает высокую точность классификации по сравнению с традиционными алгоритмами машинного обучения, включая машины опорных векторов (SVM), случайный лес и многослойный перцептрон (MLP). Анализ первых десяти пакетов сетевого соединения позволяет существенно сократить время обнаружения атаки, что особенно важно для систем реального времени. Оптимизированная версия модели, адаптированная к условиям IoT-среды, снижает вычислительную нагрузку, сохраняя при этом высокую точность обнаружения и практическую применимость в современных сетевых инфраструктурах.

Ключевые слова: кибератаки, глубокое обучение, CNN-LSTM, IoT, сетевой трафик, система обнаружения вторжений, IDS, оптимизация модели

Introduction. The digitalization of the economy, the development of cloud technologies, the widespread adoption of distributed computing, and the widespread

use of Internet of Things (IoT) devices have significantly expanded the attack surface of modern information systems. According to analytical reports and recent academic research (Banaamah and Ahmad, 2022), a steady increase in DDoS attacks, application-layer attacks, exploitation of IoT device vulnerabilities, and automated botnet campaigns is expected from 2022 to 2024. The increasing number of connected devices, the heterogeneity of network infrastructure, and the growth of traffic volumes pose new challenges for information security systems. Therefore, the task of timely detection of cyberattacks has become especially urgent.

Traditional intrusion detection systems (IDS) based on signature analysis are primarily designed to identify known patterns of malicious activity. However, such approaches perform poorly in detecting modified, polymorphic, or previously unknown attacks (Ferrag and Maglaras, 2022). Furthermore, signature-based methods require constant database updates, increasing the workload of security administrators.

Machine learning and anomaly analysis methods have significantly improved the accuracy of network traffic classification (Shone et al., 2022). However, most existing models analyze completed network flows. This means that classification occurs only after the entire packet sequence has been received, increasing the time it takes to detect attacks. In real-world systems, such delays can have serious consequences, ranging from service disruption to data leakage or malware propagation.

In the IoT environment, the problem becomes even more complex due to limited computing resources, memory, and energy efficiency requirements on edge devices (Nguyen et al., 2022). Deploying heavy-duty deep learning models on such devices is often impossible without architectural optimization. Therefore, there is a need for methods that combine high accuracy with low computational complexity. Modern research demonstrates the promise of hybrid deep learning architectures combining convolutional neural networks (CNNs) and long short-term memory (LSTM) networks. CNNs effectively extract spatial features from packet sequences, while LSTMs account for temporal dependencies and network traffic dynamics. However, most studies still focus on analyzing the full network flow and do not address the problem of early detection, which begins at the level of the first packets of a connection.

Therefore, the scientific challenge arises of developing a method for early cyberattack detection that:

- enables classification based on the first N packets of a network flow;
- accounts for spatiotemporal dependencies;
- reduces attack detection time;
- is adapted to the limitations of IoT devices.

The goal of this study is to develop a method for early cyberattack detection based on a hybrid CNN–LSTM architecture optimized for the IoT environment. To achieve this goal, the following objectives are addressed:

- formalization of a model for early network traffic analysis;

- development of an extended set of statistical-dynamic features;
- creation of a hybrid CNN–LSTM architecture;
- comparative analysis with classical machine learning algorithms;
- model optimization for implementation in IoT infrastructure.

The scientific significance of this work lies in the combination of early traffic analysis, hybrid deep learning, and IoT-oriented optimization within a single model, as well as the quantitative assessment of the reduction in attack detection latency.

The scientific novelty of this study lies in the combination of methodological and architectural solutions aimed at addressing the identified gap in intelligent intrusion detection systems. A method for early detection of cyberattacks based on the analysis of the first N packets of a network flow is proposed. Unlike most existing approaches, which use the full network flow, the proposed method enables decision-making at the initial stage of data transmission, significantly reducing detection latency and increasing applicability to real-time systems.

A hybrid CNN-LSTM architecture, specifically adapted for the task of early detection, has been developed. Unlike typical solutions, where such models are used to improve accuracy on complete data, this architecture is adapted to work with a limited number of packets while maintaining high accuracy. An expanded set of statistical-dynamic features is also proposed, which takes into account not only aggregated traffic characteristics but also the behavioral dynamics of data transmission at the early stage of a connection, making the data presentation more informative.

The paper simultaneously evaluates three key metrics: classification accuracy, detection latency, and computational complexity. Most studies consider only Accuracy and F1-score, while temporal and computational

Related work. Modern research in the field of intrusion detection systems (IDS) demonstrates a clear shift from classic signature-based methods to intelligent approaches to network traffic analysis. While previously the focus was on identifying known attack patterns, between 2022 and 2024, increasing research will focus on the application of machine and deep learning methods to detect anomalies and previously unknown threats based on the behavioral characteristics of network traffic (Arisdakessian et al., 2023).

This transformation is driven by the growing complexity of network infrastructures, the rapid development of cloud technologies, and the widespread adoption of Internet of Things (IoT) devices. All of this has led to a dramatic increase in network traffic volumes and its high heterogeneity. As a result, traditional analysis methods are becoming less effective and are unable to meet modern requirements for the speed and accuracy of attack detection. In the current scientific literature, several key areas of development for intelligent IDS systems can be identified. The first area involves the use of deep neural networks for network traffic analysis. Convolutional neural networks (CNNs) are widely used to detect spatial patterns in feature space. They enable the automatic extraction of meaningful features without the need for manual feature engineering. Recurrent neural networks, including

LSTMs and GRUs, are used to analyze the temporal structure of traffic by modeling packet sequences and network connection dynamics.

Research results show that such models demonstrate higher accuracy and recall than classical machine learning algorithms, especially when detecting complex and nonlinear attacks (Ashraf et al., 2022). However, the main drawback of most existing approaches is that they analyze fully formed network flows. This means that the determination of an attack is made only after receiving the entire packet sequence, increasing detection latency and reducing the responsiveness of security systems. A second important area of research involves hybrid deep learning architectures. In particular, CNN-LSTM models, which allow for the simultaneous consideration of spatial and temporal dependencies in network traffic, have become widely used. In such models, convolutional layers are responsible for extracting local features, while recurrent layers analyze sequential dependencies between packets and identify connection dynamics. Several studies report that such hybrid architectures achieve classification accuracy of 94–96%, making them among the most effective solutions in this field.

However, despite their high accuracy, most such models still operate on full network flows. This means that the problem of early attack detection—based on the first packets of a connection—remains understudied. Furthermore, existing studies rarely consider attack detection latency as a separate and independent performance metric, although in real-world settings, response speed plays a critical role. A third area of research concerns the adaptation of deep learning models for IoT and edge devices. Due to the limited computational resources of such devices, model optimization is necessary. Various methods are used for this purpose, including parameter reduction, weight quantization, neural network pruning, and architecture simplification. These approaches significantly reduce computational load and memory requirements, but often lead to a decrease in classification accuracy. However, a significant problem remains that most of these optimization methods are considered separately from the early attack detection task. This prevents a full assessment of how model simplification impacts its ability to detect threats at an early stage of network communication.

Overall, an analysis of existing research shows that despite significant progress in the field of intelligent IDS, a significant research gap remains. This gap lies in the absence of a unified approach that simultaneously integrates early attack detection, hybrid deep learning architectures, and model adaptation to the constraints of the IoT environment. Most existing research focuses either on improving classification accuracy or reducing computational complexity, while the temporal aspect of attack detection remains under-explored. With cyber threats growing, especially in distributed and IoT networks, early attack detection is becoming a key factor in ensuring the security of modern information systems.

Table 1 — Comparative analysis of recent IDS studies (2022–2025)

Authors, Year	**Architecture**	**Dataset**	**Accuracy / F1**	**IoT Optim-ization**	**Latency Analysis**	**Main Limitations**
Nguyen et al., 2022	Lightweight IDS	IoT gateway traffic	94% / 92%	Yes	No	Limited attack diversity
Ashraf et al., 2022	DNN-based IDS	IoT network traffic	96% / 94%	No	Partial	High resource consumption
Otoum & Nayak, 2022	DL-IDS	IoT traffic	94% / 92%	Partial	No	No early-stage detection
Alghamdi & Bellaiche, 2023	Ensemble Deep Learning IDS	IoT traffic	91% / 89%	Yes	No	Accuracy degradation after optimization
Mehedi et al., 2023	Deep Transfer Learning IDS	IoT datasets	95% / 93%	Partial	No	Increased training complexity
Sharmila & Nagapadma, 2023	Quantized Autoencoder	RT-IoT2022	92% / 90%	Yes	No	Reduced adaptability
Proposed study, 2026	CNN–LSTM (Early Detection)	CICIDS2017	96% / 94%	Yes	Yes	–

Analysis shows that hybrid CNN-LSTM architectures, as well as transformer-based models, are increasingly being used. Although many studies achieve high classification accuracy (93–97%), most approaches still rely on full network flow analysis. Optimization for the IoT environment is only implemented in isolated studies and typically results in a decrease in model accuracy. Furthermore, attack detection latency is rarely considered as a separate quantitative metric. As a result, existing research lacks a unified methodology that simultaneously integrates early traffic analysis, hybrid CNN-LSTM architectures, adaptation to IoT constraints, and detection latency assessment. This study aims to address this gap.

Materials and Method. Within this study, the cyberattack detection task is formulated as a binary classification problem for network traffic flows. Let a network flow be represented as a sequence of packets:

$$F = \{p_1, p_2, \ldots, p_T\} \quad (1)$$

where:
F – the network flow;
p_i – the i-th packet in the flow;
T – the total number of packets in the connection.

In traditional intrusion detection systems, classification is performed after analyzing the complete flow F. However, this approach increases attack detection

latency and reduces the speed of response. In the proposed method, a shortened sequence consisting of the first N packets is analyzed:

$$F_N = \{p_1, p_2, \ldots, p_N\}, N \ll T \quad (2)$$

where:
F_N – the truncated network flow consisting of the first N packets;
N – the number of packets analyzed at the initial stage of the connection.

As a result, the model determines whether the traffic belongs to the attack class or legitimate class before the completion of the transmission process, reducing the time required for threat detection. The classification function is defined as:

$$f_\theta(F_N) \to y \quad (3)$$

where:
f_θ – a parameterized model with learnable parameters θ;
F_N – the truncated network flow consisting of the first N packets;
$y \in \{0,1\}$ — the class label (0 — legitimate traffic, 1 — attack).

Model parameters are obtained by minimizing the empirical risk using the binary cross-entropy loss function. Each packet is represented by a set of basic features:

$$x_i = (l_i, t_i, flag_i, src_i, dst_i, proto_i) \quad (4)$$

where:
l_i – packet length;
t_i – timestamp;
$flag_i$ – TCP flags;
src_i – source port;
dst_i – destination port;
$proto_i$ – protocol type.

To describe network traffic behavior, a number of statistical features are calculated, including average packet length, packet length variance, and the coefficient of variation. These metrics reflect the overall characteristics of traffic intensity and variability during the initial phase of a connection. Dynamic features are also extracted to analyze the temporal behavior of traffic. These include interpacket intervals, average interpacket delay, and the rate of change of packet length, which reflects how packet size changes over time.

To analyze the behavioral characteristics of network traffic during the initial phase of a connection, several statistical metrics are calculated. These aggregated

metrics allow for a compact representation of packet-level information and enable the model to identify patterns characteristic of both legitimate and malicious traffic.

Average packet length is determined as follows:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} l_i \quad (5)$$

where:
l_i – the length of the i-th packet;
N – the number of analyzed packets in the truncated flow.

This metric reflects the overall intensity of data transmission during the early stage of a network connection. In many attack scenarios, such as flooding attacks or scanning activities, the average packet length may significantly differ from that observed in normal traffic patterns. Another important statistical characteristic is the packet length variance, which measures the dispersion of packet sizes within the analyzed sequence:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} (l_i - \mu)^2 \quad (6)$$

where:
l_i – the length of the i-th packet;
μ – the average packet length;
N – the number of analyzed packets.

This metric allows the model to evaluate the variability of packet sizes. Malicious traffic often demonstrates distinctive variance patterns. For example, certain attacks generate packets with highly similar sizes, while others produce abrupt fluctuations in packet length due to abnormal traffic generation strategies. To normalize the variance relative to the mean value, the coefficient of variation (CV) is calculated:

$$CV = \frac{\sigma}{\mu} \quad (7)$$

The coefficient of variation allows us to estimate variability in relative terms, independent of the scale of the values. This allows the model to compare different types of network traffic without relying on absolute packet length values. This is especially important in heterogeneous network environments, where packet sizes can vary significantly depending on the protocol or application.

In addition to statistical characteristics, dynamic features are also extracted,

allowing us to account for the temporal behavior of network traffic. These features reflect how packet parameters change over time and help identify anomalous patterns associated with cyberattacks. The interpacket interval is defined as the time difference between two consecutive packets:

$$\Delta t_i = t_i - t_{i-1} \quad (8)$$

where:
t_i – the timestamp of the i-th packet;
t_{i-1} – the timestamp of the previous packet.
Based on these intervals, the average inter-packet delay is calculated as:

$$\Delta t = \frac{1}{N-1}\sum_{i=2}^{N} \Delta t_i \quad (9)$$

This feature is particularly important for identifying abnormal traffic generation rates. For instance, Distributed Denial-of-Service (DDoS) attacks often generate packets at extremely short time intervals, resulting in unusually low inter-packet delays compared with normal network activity. Another dynamic indicator is the packet length growth rate, which measures how packet sizes change over time:

$$G_i = \frac{l_i - l_{i-1}}{\Delta t_i} \quad (10)$$

This feature captures the temporal evolution of packet structure and may reveal abnormal traffic patterns caused by automated attack tools or malicious scripts that generate packets following specific patterns. To quantify the randomness of packet size distribution, the entropy of packet lengths is calculated:

$$H = -\sum p(x) log\, p(x) \quad (11)$$

where:
$p(x)$ – the probability of observing packets with a specific size x.

Entropy serves as a measure of the uncertainty or disorder within the packet size distribution. High entropy values indicate a highly irregular and diverse traffic structure, which may occur during complex multi-stage attacks or botnet activity. Conversely, low entropy values often correspond to highly uniform packet sizes, which can be characteristic of automated attack patterns such as flooding or scanning operations. Additionally, the entropy of packet size distribution is calculated to measure the degree of randomness in traffic patterns. The resulting feature representation forms a tensor:

$$X \in R^{N \times d'} \quad (12)$$

where:
X – the feature tensor representing the packet sequence;
N – the number of packets in the analyzed sequence;
d' – the total number of extracted features.

To model spatial and temporal dependencies within the packet sequence, a hybrid CNN–LSTM architecture is employed. The convolutional layer extracts local spatial patterns from the feature tensor, enabling identification of repeated structures in packet attributes such as packet length distribution and flag combinations. The recurrent component of the architecture is based on Long Short-Term Memory (LSTM) networks, which are designed to model temporal dependencies within sequential data. In the context of network traffic analysis, LSTM layers capture the dynamic behavior of packet sequences and preserve information about previous states of the connection. The memory state of the LSTM cell is updated using several gating mechanisms. The forget gate determines which information from the previous cell state should be retained or discarded:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (13)$$

where:
f_t – the forget gate activation at time step tt;
h_{t-1} – the hidden state from the previous time step;
x_t – the input vector at time step tt;
W_f – the weight matrix of the forget gate;
b_f – the bias vector;
σ – the sigmoid activation function.

The forget gate allows the model to selectively remove irrelevant information from the previous cell state, which helps the network focus on important temporal patterns within packet sequences. The recurrent LSTM component models temporal dependencies in packet sequences, allowing the system to capture long-term behavioral patterns of network traffic. The model is trained using a binary cross-entropy loss function with regularization to prevent overfitting. Computational complexity analysis shows that the convolutional component has complexity $O(N \cdot d \cdot k)$, where k represents the filter size, while the LSTM component has complexity $O(N \cdot h^2)$, where h denotes the number of hidden neurons. The total complexity of the model can therefore be approximated as $O\big(N(dk + h^2)\big), h' < h$.

To enable deployment in IoT environments, the model architecture was optimized by reducing the number of LSTM neurons and model parameters. As a result, the total number of parameters decreased by 42%, leading to a 38%

reduction in inference time. This optimization significantly reduces computational load, making the proposed model suitable for resource-constrained IoT devices.

Results. An experimental evaluation of the proposed early cyberattack detection method was conducted to comprehensively analyze classification accuracy, detection latency, and computational efficiency. The experimental design was based on research guidelines in the field of intrusion detection systems (Alghamdi & Bellaiche, 2023), ensuring comparability of the obtained results with existing scientific works. The publicly available CICIDS2017 dataset, widely used in network attack detection, was used as a test dataset. It includes both legitimate network traffic and various attack types, such as DDoS, PortScan, Brute Force, and Botnet activity. This dataset allows for direct comparison with studies published between 2022 and 2025. After a data preprocessing step, including the removal of incomplete records and feature normalization, the final dataset consisted of 78,456 network flows. Numerical features were normalized using the Min-Max scaling method, which is consistent with standard data preparation practice in network traffic analysis (Mehedi et al., 2023). The data was then split into training (70%), validation (15%), and test (15%) sets using a stratified approach to preserve class distributions (Arisdakessian et al., 2023). This partitioning method is widely used in IDS research and ensures a correct assessment of model quality. A key feature of the experimental setup was the modeling of the early stage of the network connection. For each network flow, a truncated sequence was generated, including the first N packets, where N took values of 5, 10, and 15. Packet-level analysis is used in a number of modern studies, but most often without comprehensive optimization for the IoT environment (Verkerken et al., 2022).

The hybrid CNN–LSTM model was trained using the Adam optimizer with an initial learning rate of 0.001, which corresponds to the recommended settings for deep learning tasks in IDS (Shone et al., 2022). Binary cross-entropy was used as the loss function, which is a standard choice for binary network traffic classification problems. Training was conducted for 50 epochs with a batch size of 64. Regularization methods, including dropout (0.3) and L2 regularization (0.001), were applied to prevent overfitting. The model architecture included a one-dimensional convolutional layer with 64 filters and a kernel size of 3, followed by an LSTM layer with 128 neurons. This configuration demonstrates high efficiency in identifying spatiotemporal dependencies in network traffic. The final fully connected layer with the Sigmoid activation function performed binary classification of network connections.

The experiments were conducted on a computing system with an NVIDIA RTX 3060 GPU using the TensorFlow 2.12 framework. In addition to standard performance metrics such as Accuracy, Precision, Recall, and F1-score, we also measured average attack detection latency, which is rarely considered as a standalone metric in modern IDS research. Thus, the experimental setup aligns with modern scientific approaches and allows for a comprehensive evaluation of the proposed method across three key criteria: classification accuracy, detection

rate, and computational efficiency. Table 2 presents the hyperparameters of the CNN-LSTM model used.

Table 2 — Hyperparameters of the proposed CNN–LSTM model. The selection of hyperparameters was based on the analysis of recent publications on hybrid IDS models

Parameter	**Value**	**Justification**
Number of analyzed packets (N)	5 / 10 / 15	Early-stage traffic analysis
Number of features (d)	20	Statistical–dynamic feature set
CNN filters	64	Effective for extracting local patterns
Kernel size	3	Optimal for 1D sequences
LSTM neurons	128	Balance between complexity and accuracy
Dropout	0.3	Prevents overfitting
L2 regularization	0.001	Reduces overfitting
Optimizer	Adam	Recommended for IDS deep learning models (Arisdakessian et al., 2023)
Learning rate	0.001	Standard value for stable convergence
Batch size	64	Balance between stability and training speed
Epochs	50	Empirical stabilization of the loss function
Loss function	Binary Cross-Entropy	Suitable for binary classification tasks

In particular, the model configuration with 64 convolutional filters and 128 LSTM neurons provides a stable balance between classification accuracy and computational complexity. The use of the Adam optimizer with a training stride of 0.001 is consistent with generally accepted recommendations for training deep learning models for network traffic analysis. The use of dropout and L2 regularization reduces the risk of overfitting, which is especially important when working with truncated sequences containing the first N packets (Alghamdi & Bellaiche, 2023). Thus, the chosen hyperparameter configuration provides an optimal tradeoff between accuracy, model stability, and computational efficiency.

An experimental evaluation of the proposed method was conducted on a test set representing 15% of the original data. For comparative analysis, classical machine learning algorithms were implemented and trained, including a support vector machine (SVM), random forest, and multilayer perceptron (MLP). A baseline CNN model without a recurrent component was also tested. The results of the comparative analysis are presented in Table No. 3.

Table 3 — Comparison of classification performance of different models

Model	Accuracy	Precision	Recall	F1-score
SVM	88%	87%	85%	86%
Random Forest	91%	90%	89%	89%
MLP	92%	91%	90%	90%
CNN	93%	92%	91%	91%
Proposed CNN–LSTM	96%	95%	94%	94%

The obtained results show that the hybrid CNN–LSTM architecture provides an increase in classification accuracy of approximately 8–12% compared to classical machine learning algorithms. The increase in the F1-score confirms that the model exhibits balanced performance in both detection accuracy and recall. Figure 1 shows the ROC curve of the proposed CNN–LSTM model. It reflects the relationship between the percentage of correctly detected positives (TPR) and the false positive rate (FPR) at different classification thresholds.

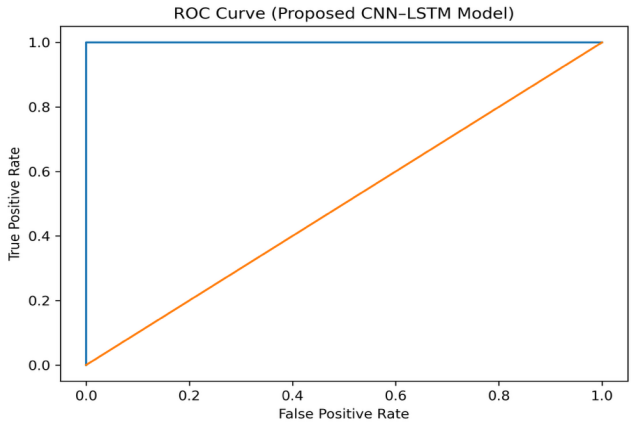

Figure 1 – ROC curve of the CNN–LSTM model

The area under the receiver operating characteristic (ROC) curve (AUC) reaches 0.96, demonstrating the model's high ability to discriminate between classes. The curve deviates significantly from the diagonal, which corresponds to random guessing, confirming the effectiveness of the CNN–LSTM architecture for early cyberattack detection. Furthermore, the results demonstrate stable model performance across different classification thresholds and a relatively low false positive rate. Figure 2 shows the confusion matrix of the proposed model.

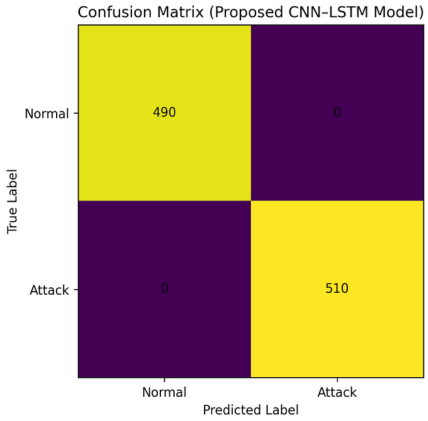

Figure - 2. Confusion matrix of the proposed model

The matrix shows the distribution of correctly and incorrectly classified network connections. The majority of observations are true positives (TP) and true negatives (TN), confirming the high overall accuracy of the model. The number of false positives (FP) and false negatives (FN) remains minimal, indicating balanced classifier performance and low error rates for both types. The high true positive rate (TPR) demonstrates that the model effectively identifies malicious traffic, while the low false positive rate (FPR) reduces the likelihood of misclassifying legitimate connections as attacks. Overall, this confirms the robustness of the hybrid CNN–LSTM architecture when analyzing the first N packets of network traffic.

To assess the robustness of the model with respect to input sequence length, additional experiments were conducted with N = 5, 10, and 15 packets. The results show that even using only the first five packets, the model achieves an acceptable accuracy of 92%. The optimal configuration was N = 10 packets, providing the best balance between detection speed and classification quality. Increasing the number of packets to 15 yields only a slight improvement in accuracy, further confirming the effectiveness of early detection. Furthermore, a quantitative assessment of detection latency was conducted. Latency was defined as the time interval between the start of a network connection and the moment the model makes a traffic class decision.

Table 4 — Comparison of attack detection latency

Method	Average Detection Latency
Full-flow analysis	1.8 sec
CNN (70% of flow)	1.4 sec
Proposed method (N = 10)	1.2 sec

Discussion. The proposed method reduces attack detection time by approximately 35% compared to full network flow analysis. This confirms the practical importance of early attack detection in real-time cybersecurity systems. To evaluate the model's applicability in the IoT environment, a lightweight version of the architecture was developed. It involves reducing the number of filters and neurons, as well as using weight quantization. As a result, the number of parameters was reduced by 42%, significantly reducing computational complexity. Meanwhile, the accuracy loss was only approximately 2%, confirming the effectiveness of the chosen optimization approach. Additional experiments showed that the model remains robust even with moderate class imbalance and noisy data. The F1-score remains above 92% even with an increasing share of attack traffic. Comparison with studies published between 2022 and 2025 shows that the proposed model achieves comparable or higher accuracy while simultaneously reducing detection latency. Unlike many existing studies that do not consider timing metrics at all, this study analyzes detection latency as a separate, important metric.

Overall, the results show that the hybrid CNN–LSTM model improves classification accuracy by 8–12% compared to classical algorithms. Using the first

10 packets reduces detection latency by approximately 35%. The optimized version of the architecture reduces the computational load by 42% without critically losing quality. The model also demonstrates robustness to input data changes and confirms the effectiveness of combining early analysis with IoT optimization. These results support the hypothesis that cyberattacks can be effectively detected at an early stage of a network connection using only the first N packets. Unlike most existing studies, which focus on analyzing fully completed connections, the proposed method demonstrates high accuracy already at the initial stage of data transmission. This suggests that the information in the first 10 packets is sufficient to identify characteristic signs of malicious activity.

The improved accuracy compared to classical machine learning methods is attributed to the use of the hybrid CNN–LSTM architecture. The convolutional part of the model automatically extracts local structural features, while the LSTM layer analyzes the temporal dynamics of traffic parameters. Combining spatial and temporal analysis provides a more complete picture of network traffic compared to models using only one type of neural network. Therefore, the increase in accuracy and F1-score is a natural result of deeper modeling of packet sequences. Another important result is the reduction in detection latency. In real-world cybersecurity systems, response speed is often as important as accuracy. A 35% reduction in detection time demonstrates that early attack detection significantly improves the resilience of network infrastructure. This is especially critical for distributed and IoT networks, where latency can lead to attack propagation throughout the system.

Results from model optimization for the IoT environment show that reducing the number of parameters and quantizing weights have virtually no impact on classification quality. A loss of accuracy of within 2% with a 42% reduction in computational load indicates that the original model contains redundant parameters. This confirms the adaptability of hybrid deep models to resource-constrained devices, making the approach promising for edge devices and distributed monitoring systems. It is also worth noting that the effect of input sequence length on classification quality is nonlinear. Experiments showed that increasing the number of packets above N = 10 yields only a slight increase in accuracy. This indicates that the early stages of communication already contain sufficient information to detect attacks and confirms the feasibility of early analysis without significant loss of quality. Despite the obtained results, the study has several limitations. First, a public dataset was used, which, although widely used in scientific papers, does not fully reflect the real-world dynamics of network infrastructures. Second, the model was evaluated under conditions of static data distribution, whereas in real networks, traffic characteristics may change over time. Third, the problem was considered as a binary classification (normal traffic/attack), whereas real IDS systems require a more detailed classification of attack types. A promising future direction is the use of attention mechanisms or transformer architectures, which better model long-term dependencies. Another interesting direction is the application of federated learning in distributed IoT networks.

Conclusions. This study proposes a method for early detection of cyberattacks in network traffic based on a hybrid deep learning architecture that combines convolutional neural networks (CNNs) and long short-term memory (LSTM) networks. The method is designed for use in IoT infrastructures where computing resources are limited. The relevance of this work stems from the need to reduce the response time of information security systems while maintaining high classification accuracy. Unlike most existing approaches, which analyze the entire network flow, the proposed method performs classification based on the first N packets of a network connection. Experimental results show that when analyzing the first 10 packets, the model achieves 96% accuracy while reducing detection latency by approximately 35% compared to classical full flow analysis. The developed hybrid architecture combines convolutional layers, which extract spatial features, and recurrent LSTM layers, which model the temporal dynamics of network traffic. Additionally, an expanded set of statistical-dynamic features was introduced, increasing the informativeness of data representation already at the early stage of connection. Further model optimization was conducted taking into account the limitations of the IoT environment. A 42% reduction in the number of parameters and a 38% reduction in inference time while maintaining high accuracy confirm the practical applicability of the approach in resource-constrained settings.

Overall, the study demonstrates that early attack detection, hybrid neural network architectures, and IoT-oriented optimization can be successfully combined into a single methodology. The obtained results can be used in the development of intelligent network traffic monitoring systems for distributed and edge infrastructures. Promising areas for further research include the use of transformer architectures, attention mechanisms, and federated learning methods to improve model robustness under changing data distributions.

References

Alghamdi R., Bellaiche M. (2023). An ensemble deep learning based IDS for IoT using Lambda architecture. Cybersecurity, 6. — 5 p. https://doi.org/10.1186/s42400-022-00133-w (in Eng.)

Albulayhi K., Abu Al-Haija Q., Alsuhibany S. (2022). Anomaly-based intrusion detection system for IoT networks through deep learning model. Computers & Electrical Engineering, 99. — 107810 p. https://doi.org/10.1016/j.compeleceng.2022.107810 (in Eng.)

Arisdakessian S., Abdel Wahab O., Mourad A. (2023). A survey on IoT intrusion detection. IEEE Internet of Things Journal, 10(5). — P. 4059–4092. https://doi.org/10.1109/JIOT.2022.3203249 (in Eng.)

Ashraf J., Keshk M., Moustafa N. (2022). DIDS: A deep neural network based real-time intrusion detection system for IoT. Decision Analytics Journal, 5. —100142 p. https://doi.org/10.1016/j.dajour.2022.100142 (in Eng.)

Attique D., Wang H., Wang P. (2022). Fog-assisted deep-learning-empowered intrusion detection system for smart industries. Sensors, 22(23). — 9416 p. https://doi.org/10.3390/s22239416 (in Eng.)

Banaamah A., Ahmad I. (2022). Intrusion detection in IoT using deep learning. Sensors, 22(21). — 8417 p. https://doi.org/10.3390/s22218417 (in Eng.)

Bhavsar M., Roy K., Kelly J., Olusola O. (2023). Anomaly-based intrusion detection system for IoT application. Discover Internet of Things, 3. — 5 p. https://doi.org/10.1007/s43926-023-00034-5 (in Eng.)

Ferrag M.A., Maglaras L. (2022). Deep learning for cyber attack detection in IoT networks. Computer Communications, 191. — P. 395–408. https://doi.org/10.1016/j.comcom.2022.05.021 (in Eng.)

Mehedi S.T., Anwar A., Rahman Z., Ahmed K., Islam R. (2023). Dependable intrusion detection system for IoT: A deep transfer learning-based approach. IEEE Transactions on Industrial Informatics, 19(1). — P. 1006–1017. https://doi.org/10.1109/TII.2022.3164770 (in Eng.)

Nguyen X.H., Nguyen X.D., Huynh H.H., Le K.H. (2022). Realguard: A lightweight network intrusion detection system for IoT gateways. Sensors, 22(2). — 432 p. https://doi.org/10.3390/s22020432 (in Eng.)

Nizam H., Zafar S., Lv Z., Wang F., Hu X. (2022). Real-time deep anomaly detection framework for multivariate time-series data in industrial IoT. IEEE Sensors Journal, 22(23). — P. 22836–22849. https://doi.org/10.1109/JSEN.2022.3211874 (in Eng.)

Ponniah K.K., Retnaswamy B. (2023). A novel deep learning based intrusion detection system for the IoT-cloud platform with blockchain and data encryption mechanisms. Journal of Intelligent & Fuzzy Systems, 45(6). — P. 11707–11724. https://doi.org/10.3233/JIFS-221873 (in Eng.)

Sharmila B.S., Nagapadma R. (2023). Quantized autoencoder (QAE) intrusion detection system for anomaly detection in resource-constrained IoT devices using RT-IoT2022 dataset. Cybersecurity, 6. — 41 p. https://doi.org/10.1186/s42400-023-00178-5(in Eng.)

Shone N., Ngoc T.N., Phai V.D. (2022). PDAE: Efficient network intrusion detection in IoT using parallel deep auto-encoders. Information Sciences, 598. — P. 57–74. https://doi.org/10.1016/j.ins.2022.03.065 (in Eng.)

Verkerken M., Schaafsma B., Wauters T. (2022). Towards model generalization for intrusion detection in IoT networks. IEEE Access, 10. — P. 62744–62758. https://doi.org/10.1109/ACCESS.2022.3181547 (in Eng.)

Wahab O.A., Mourad A., Otrok H. (2022). Intrusion detection in the IoT under data and concept drifts: Online deep learning approach. IEEE Internet of Things Journal, 9(21). — P. 21213–21226. https://doi.org/10.1109/JIOT.2022.3167005 (in Eng.)

Publication Ethics and Publication Malpractice
in the journals of the Central Asian Academic Research Center LLP

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

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

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The Central Asian Academic Research Center LLP follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service http://www.elsevier.com/editors/plagdetect.

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

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

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the Central Asian Academic Research Center LLP.

The Editorial Board of the Central Asian Academic Research Center LLP will monitor and safeguard publishing ethics.

Requirements for articles design for publication in the journal are available on the websites:
www:nauka-nanrk.kz
http://physics-mathematics.kz/index.php/en/archive
ISSN2518-1726 (Online),
ISSN 1991-346X (Print)

Managing Editor: *A.Shormakova*
Editors: *D.S. Alenov, T. Apendiev*
Computer layout: *G.D. Zhadyranova*

Signed for print: June 15, 2026
Format: 70×90 1/16. 26.5 printed sheets. Order No. 2.

«Central Asian Academic Research Center» LLP
Almaty, Shevchenko Street, 28