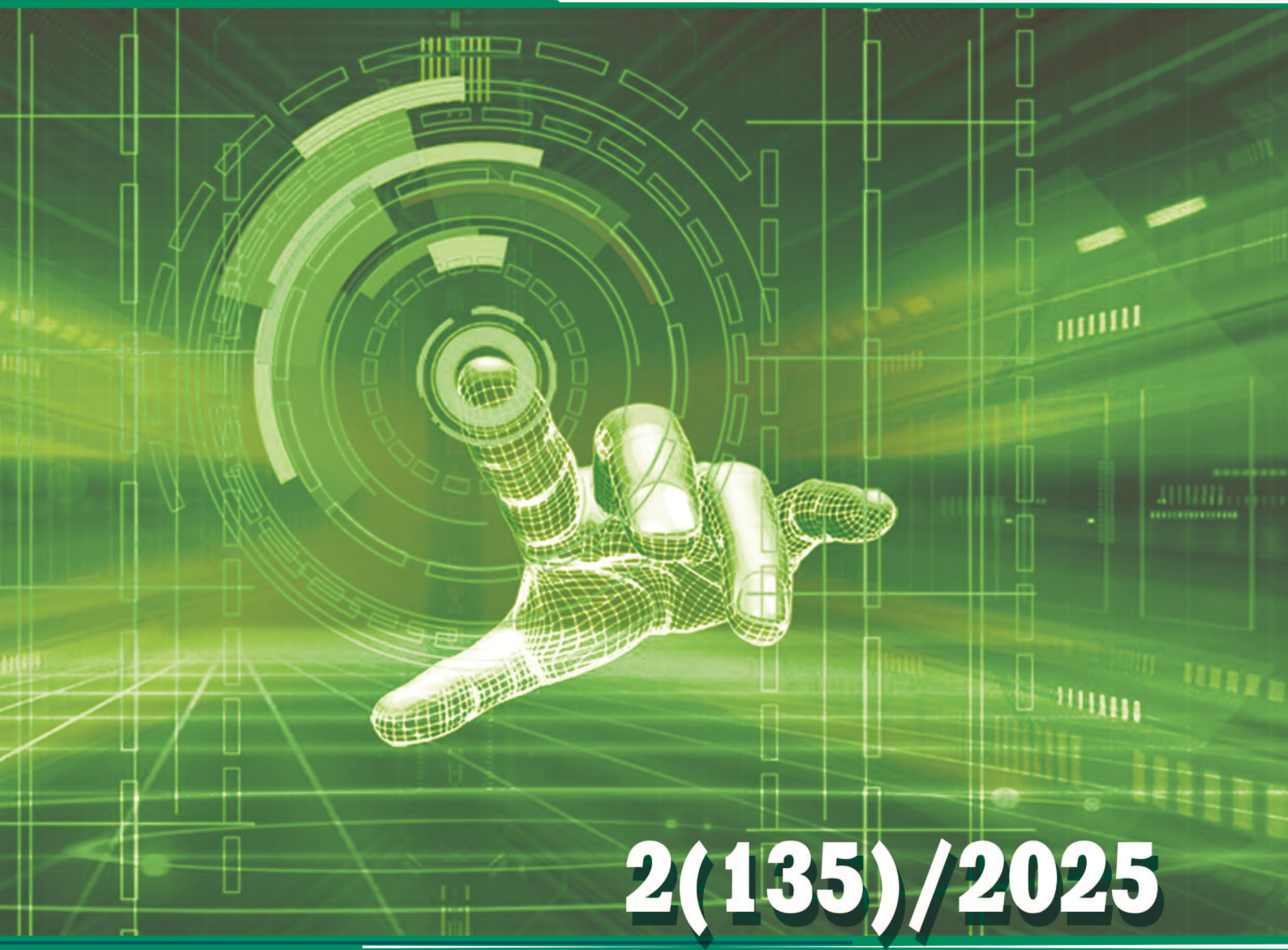
ISSN 1684-8853 (print); ISSN 2541-8610 (online)

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ



2(135)/2025

doi:10.31799/1684-8853-2025-2

ISSN 1684-8853 (print); ISSN 2541-8610 (online)

SCIENTIFIC JOURNAL

2(135)/2025

PEER REVIEWED JOURNAL

INFORMATSIONNO-UPRAVLIAIUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

Contact information
The Editorial and Publishing Center, SUAI
67A, Bol'shaya Morskaia, 190000, Saint Petersburg, Russia
Website: http://i-us.ru/en, e-mail: ius.spb@gmail.com
Tel.: +7 - 812 494 70 02

The Journal was registered in the Ministry of Press, Broadcasting and Mass Media of the Russian Federation.
Registration Certificate JD № 77-12412 from April, 19, 2002.
Re-registration in the Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications (ROSKOMNADZOR) due to change of the founder A. Vostrikov,
PI № FS 77-82226 from November, 23, 2021

Submitted for publication 05.03.25. Passed for printing 28.04.25. Signed for printing: 30.04.2025.
Format $60\times84_{1/8}$. Phototype CentSchbkCyrill BT. Digital printing.

Layout original is made at the Editorial and Publishing Center, SUAI.
67A, Bol'shaya Morskaia, 190000, Saint Petersburg, Russia
Printed from slides at the Editorial and Publishing Center, SUAI.
67A, Bol'shaya Morskaia, 190000, Saint Petersburg, Russia

The journal is indexed in Scopus.

Free distribution.

doi:10.31799/1684-8853-2025-2

ISSN 1684-8853 (print); ISSN 2541-8610 (online)

НАУЧНЫЙ ЖУРНАЛ

2(135)/2025

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

Учредитель
А. А. Востриков
Издатель
Санкт-Петербургский государственный университет аэрокосмического приборостроения

Редактор: А. Г. Ларионова
Корректор: Т. В. Звертановская
Дизайн: М. Л. Черненко, Ю. В. Умницына
Компьютерная верстка: Ю. В. Умницына

Адрес редакции: 190000, г. Санкт-Петербург,
ул. Большая Морская, д. 67, лит. А, ГУАП, РИЦ
Тел.: (812) 494-70-02, эл. адрес: ius.spb@gmail.com,
сайт: http://i-us.ru

Журнал входит в БД Scopus и в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук.

Сдано в набор 05.03.25. Подписано в печать 28.04.25. Дата выхода в свет: 30.04.2025.
Формат 60×841/8. Гарнитура CentSchbkCyrill BT. Печать цифровая.
Усл. печ. л. 9,7. Уч.-изд. л. 13,0. Тираж 1000 экз (1-й завод 50 экз.). Заказ № 128.
Оригинал-макет изготовлен в редакционно-издательском центре ГУАП.
190000, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. А.
Отпечатано в редакционно-издательском центре ГУАП.
190000, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. А.
Распространяется бесплатно.

Журнал зарегистрирован в Министерстве РФ по делам печати, телерадиовещания и средств массовых коммуникаций.
Свидетельство о регистрации ПИ № 77-12412 от 19 апреля 2002 г.
Перерегистрирован в Роскомнадзоре.
Свидетельство о регистрации ПИ № ФС77-82226 от 23 ноября 2021 г.

UDC 004.9
doi:10.31799/1684-8853-2025-2-2-15
EDN: MLUUSQ

Articles

Traffic sign recognition through the use of an Internet of Things system and deep learning

Vo Nhu Thanh[a], *PhD, Lecturer, orcid.org/0000-0001-5383-3119*
Pham Cong Thang[a], *PhD, Lecturer, orcid.org/0000-0002-6428-102X, pcthang@dut.udn.vn*
Vo Viet Truong[a], *Student, orcid.org/0009-0006-2601-0743*
Ho Van Thao[a], *Student, orcid.org/0009-0006-5016-9420*
Tran Minh Quan[a], *Student, orcid.org/0009-0001-8285-9171*
Hoang Nguyen Nhat Minh[a], *Student, orcid.org/0009-0000-1336-4426*
[a]*The University of Danang — University of Science and Technology, 54 Nguyen Luong Bang St., Danang 550000, Vietnam*

Introduction: *Automatic traffic sign recognition is an important component in modernizing traffic safety systems and minimizing related incidents.* ***Purpose:*** *To develop a Vietnamese traffic sign detection system that integrates IoT technology, cameras, sensors, and deep learning techniques.* ***Results:*** *In this study, we introduce an advanced computer vision solution that uses the neural architecture of YOLOv8 to recognize and classify traffic signs in Vietnam in real time. In addition, the system integrates the detected sign locations into Google Maps and provides a comprehensive database at all locations in urban areas. The data collection process is implemented by first capturing real-life images on the road and then combining them with the existing Vietnam traffic sign dataset from Kaggle data. To improve the reliability of the dataset, Mosaic data augmentation technique is applied. For real-time traffic sign recognition, a Raspberry Pi 4 board displays detected road signs on an HMI screen, while a custom-developed mobile application detects, recognizes, and notifies users of their locations. The trained system achieves a mAP50 value of 89% and an accuracy of 90%, which is pretty good for traffic signs detection. Moreover, the YOLOv8 model gives a precision and recall value of 89.7% and 86.8%, respectively, which is considerably higher as compared to the SVM and MSER/HOG methods.* ***Practical relevance:*** *We have succeeded in creating a real-time road sign recognition system integrated into Google Maps to meet the needs of traffic control, safety and accident reduction in the streets of Vietnam.*

Keywords — *traffic sign recognition, YOLOv8, Google Maps integration, real-time detection, road safety.*

For citation: Thanh V. N., Thang P. C., Truong V. V., Thao H. V., Quan T. M., Minh H. N. N. Traffic sign recognition through the use of an Internet of Things system and deep learning. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2025, no. 2, pp. 2–15. doi:10.31799/1684-8853-2025-2-2-15, EDN: MLUUSQ

Introduction

Traffic signs play an important role in conveying essential information, such as regulations, warnings, and instructions, which help guide drivers and reduce accidents. Recent studies have emphasized that traffic signs play a pivotal role in maintaining safety and regulating traffic, as they convey important information to drivers, pedestrians, and other road users [1]. Therefore, traffic sign recognition is essential for warning and reminding drivers to ensure road safety. However, accurate recognition of traffic signs remains a challenge due to multiple dynamic factors in real-world conditions. These include environmental variations such as extreme lighting conditions, adverse weather, and seasonal changes; physical constraints like sign deterioration and occlusions; motion-related challenges including high-speed image capture blur and vehicle vibrations; and infrastructure variations across different regions. To improve this situation, various automatic traffic sign recognition systems have been developed, each with distinct capabilities and limitations. Modern deep learning approaches like Convolutional Neural Networks (CNN) based systems [2] improve accuracy to 85–90% but demand significant computational resources, limiting their deployment on edge devices. Before the widespread adoption of neural networks, traffic sign recognition relied primarily on traditional computer vision techniques. Support Vector Machines (SVM) combined with Histogram of Oriented Gradients (HOG) [3] achieved 41.03% precision and 34.15% recall, offering computational efficiency suitable for embedded systems. Enhanced approaches using Maximally Stable Extremal Regions (MSER) with HOG features [4] improved performance to 88.75% precision and 81.35% recall. Recent commercial ADAS implementations have proposed two new systems: turn assist system and lateral clearance warning, which aim to further improve road safety in Russia [5]. Also, a multimodal mobile application called MIDriveSafely is designed to enhance road safety by detecting dangerous driving situations, providing driver feedback, offering entertainment, and enabling voice control through advanced audio-visual speech recognition

technologies [6]. IoT (Internet of Things) integrated systems have emerged as a promising solution, combining edge processing for real-time detection with cloud resources for complex analysis. However, existing IoT-based implementations face challenges in seamless integration, reliable connectivity, and scalable deployment across diverse traffic environments [7]. Integrating IoT with deep learning techniques offers a promising solution for real-time traffic sign recognition. Neural network systems, specifically CNNs, demonstrate outstanding performance in traffic sign detection tasks. Within the CNN architecture, the You Only Look Once (YOLO) method stands out as an effective method for instantaneous object recognition. Among them, YOLOv8 is the latest version that provides improved accuracy, speed, and performance in detecting small objects, making it suitable for traffic sign recognition applications. Additionally, IoT devices, such as Raspberry Pi, and mobile applications, facilitate real-time collection, processing, and transmission of traffic sign data, contributing to enhanced road safety. Recent advances in artificial intelligence and deep learning have offered innovative approaches for accurate and real-time traffic sign detection [8–10]. This study investigates the capabilities of the YOLOv8 model in recognizing traffic signs and its integration with Google Maps for location mapping. Through this holistic approach, our work advances the enhancement of automated transportation networks, supporting improved road safety and operational efficiency. By integrating detected traffic sign locations with Google Maps, the system enhances practical usability, offering real-time monitoring and management of traffic signs through a user-friendly interface. This supports informed decision-making for both drivers and traffic management consultants.

Our work focuses on enhancing road safety in Vietnam by implementing a traffic sign recognition system. Utilizing a Raspberry Pi 4 combined with a webcam and a dedicated mobile application, the system aims to raise awareness and understanding of traffic signs among road users [11]. The Raspberry Pi 4 is responsible for identifying and classifying traffic signs in real-time, while the mobile application serves as the primary tool for user interaction. Upon identifying a traffic sign, the application logs its coordinates and stores the data in a database, building upon previous research in this field [12–15]. The main point of this study focuses on improving Vietnam's roadway safety by developing an automated traffic sign detection system with the four specific objectives as follows.

1. The development of an automated traffic sign identification and classification system integrates community participation in database maintenance. The system architecture facilitates user contributions while simultaneously enhancing public understanding of traffic regulations.

2. The implementation of a dynamic traffic sign database provides real-time location information through an integrated mapping system. This database serves as a comprehensive repository of traffic sign distribution data, enabling systematic analysis and improvement of traffic information management.

3. A mobile platform is built to interact and provide detailed information about traffic signs to users. This mobile platform combines mapping functionality to illustrate the location and detailed content of traffic signs, supporting users to make rational decisions when participating in traffic.

4. The integration of advanced IoT technology with continuous real-time data collection from the community allows the system to be continuously improved and upgraded, thereby increasingly enhancing the safety of people participating in traffic.

This approach allows the development of applications to support and manage traffic management infrastructure in a sustainable manner with high scalability.

Literature survey

YOLO model

There have been many studies on different traffic sign detection and recognition schemes. Early research focused on developing recognition systems using color segmentation, moment invariance, and neural networks, which have demonstrated high recognition performance, good accuracy, and low computation time [16]. Subsequent work focused on developing systems that extract and alert drivers to the content of traffic signs using text or voice commands [17]. Other important contributions include methods for detecting and recognizing traffic signs using color information and symmetry features [18] and systems based on neural networks and swarm optimization [19]. CNN models have made some important advances in the field of traffic sign recognition. One study demonstrated the effectiveness of lightweight CNN models that achieved a reduction of more than 50% in the number of required parameters while maintaining test accuracy [20]. Another study successfully implemented CNN-UM simulations on live data stream processing platforms [21], while other works proposed automated CNN systems for image classification tasks [22]. Recent works have established mathematical models to evaluate the performance of CNN systems, providing accurate estimates of power consumption and response time [23].

The YOLO framework represents a significant advancement in real-time object detection, with YOLOv8 emerging as its latest iteration. Recent applications of YOLOv8 have demonstrated particular effectiveness in specialized domains, achieving a 7.7% improvement in drone-based detection accuracy [24, 25] and 86.09% accuracy at 28 frames per second in surveillance applications [26]. For small-object detection in remote sensing applications, enhanced variants have shown superior performance compared to the base architecture [27].

The YOLOv8 implementation selected for this study utilizes key architectural features essential for traffic sign recognition: advanced data preprocessing including Mosaic augmentation, efficient feature extraction through a specialized backbone network, and a detection system optimized for varying object scales. This architecture demonstrates particular suitability for traffic sign detection due to its ability to process real-time data streams while maintaining accuracy under diverse environmental conditions [28–30].

YOLOv8 uses a comprehensive loss function that integrates box loss to optimize bounding box coordinates, class loss to improve object classification accuracy, and object-specificity loss to improve the confidence score of detected objects. Building on this foundation, the model introduces several architectural improvements, including a split-head design that separates detection and classification tasks for better specialization, and dynamic anchor assignment that adapts to different object sizes and shapes. During detection and synthesis, feature scaling improves detection across different scales. These improvements combined make YOLOv8 particularly effective for complex, real-time applications such as traffic sign recognition, where accurate detection and classification must occur simultaneously under varying environmental conditions and time constraints.

Metrics and loss function

Intersection over Union (IoU) is a crucial metric in object detection that provides a numerical assessment of the overlap between an object detector's predicted (pd) bounding box and a ground truth bounding box (gt) [31]. It defines the following important words and is a basic component for assessing a model's accuracy:

True Positive (*TP*): A detection is labeled as TP when the IoU(gt, pd) is greater than or equal to a specified threshold α, indicating a meaningful overlap between the ground truth and the prediction.

False Positive (*FP*): If the IoU(gt, pd) falls below the threshold, the detection is marked as FP. FP denotes instances where the model incorrectly predicts an object that does not align with the ground truth, indicating a false positive result.

False Negative (*FN*): FN instances occur when a ground truth object is not detected by the model, typically due to IoU(gt, pd) being below the chosen threshold α. FN represents cases where the model misses an actual object, resulting in a false negative outcome.

The area of union between the gt and pd bounding boxes is divided by the area of intersection between them to get the IoU [32]:

$$IoU = \frac{\text{Area of Intersection between gt and pd}}{\text{Area of Union of gt and pd}}.$$

The IoU value is a number between 0 and 1, where 1 denotes a perfect match or total overlap and 0 denotes no overlap. The selection of the IoU threshold α is important since it affects how detections and ground truth objects are classified and makes it possible to adjust how well the model performs. Model evaluation can be made more flexible by adjusting the threshold, which can affect the trade-off between true positives and false positives.

Precision is a measure of a model's accuracy in accurately recognizing pertinent objects. It is computed by dividing the total number of positive predictions produced by the model (TP) by the number of all Positives (TP and FP) [33]:

$$Precision = \frac{\text{True Positives}}{\text{True Positives + False Positives}}.$$

Recall measures how well a model can identify all true positives. Recall is calculated as the number of TP divided by the total number of TP and FN. Recall estimates the proportion of data that are correctly identified compared to all true positives in the data set:

$$Recall = \frac{\text{True Positives}}{\text{True Positives + False Negatives}}.$$

Effective systems typically require high percentages in both precision and recall. However, optimizing these metrics involves balancing factors such as increasing precision often reduces recall and vice versa. This balance must be carefully adjusted during system optimization to suit operational needs.

Mean Average Precision (mAP) is a crucial parameter for assessing how well object identification models work. By combining precision and recall, it offers a thorough evaluation that gives important details about the model's correctness.

The object detection metric mAP50 uses a 0.5 overlap threshold to evaluate how well models can spot objects. When a model scores high on mAP50, it means it's doing a good job overall at finding things in images. For more thorough testing, we also look at mAP50-95, which checks detection qual-

ity across different overlap requirements from 0.5 to 0.95. This helps us understand how precisely the model can pinpoint object locations.

These measurements help us figure out if our object detection systems are actually working well in the real world. They give us useful insights about detection quality across different object types and scenarios. Systems that do well on both mAP50 and mAP50-95 tend to be more dependable, which is why they're often chosen for things like self-driving cars and security cameras.

Loss functions measure detection accuracy by comparing model predictions against ground truth data during training. Three key components form the evaluation framework:

– box loss quantifies coordinate prediction accuracy between predicted and ground truth bounding boxes;

– class loss evaluates classification accuracy within detected regions;

– defocus loss optimizes detection performance specifically for degraded image quality scenarios.

The combined loss metrics form an optimization index that guides model convergence toward optimal object localization and classification. This composite evaluation enables systematic performance assessment and model refinement across varying detection scenarios.

Mosaic augmentation

YOLOv8's training process integrates advanced data augmentation strategies to boost detection capabilities. The primary technique, Mosaic augmentation, merges four training images into one combined input. This method, which evolved through YOLOv4 and YOLOv5 iterations, exposes the model to varied object contexts and partial visibility scenarios. Such training conditions strengthen the model's ability to handle complex detection tasks in real environments.

Mosaic augmentation combines four training images into one image at a random ratio. The algorithm follows these steps [34]:

– acquire four images from data set;

– adjust dimensions for uniformity;

– combine into 4×4 matrix layout;

– crop a random image patch from the center, which becomes the final augmented image.

When applying Mosaic augmentation technique, the processed image matrices exhibit similar structural patterns as the examples shown in Fig. 1.

This algorithm enhances the sensitivity when detecting objects in heterogeneous environments, thereby reducing the dependence on the environment. This method optimizes the flexibility of the system to be able to operate on many different contexts, thereby improving the recognition accuracy. This data augmentation protocol is quite effective for datasets that exhibit spatial differences and contextual heterogeneity. However, its effectiveness decreases when applied to datasets that include text documents, objects with prominent features, or samples with invariant spatial locations.

Comparative analysis between YOLOv8 and YOLOv5 architectures in traffic sign detection applications shows a better improvement in performance (Table 1). The experimental evidence obtained demonstrates that the object detection accuracy of YOLOv8 is superior to YOLOv5 by about 2.82%. In contrast, YOLOv5 shows higher sensitivity (object detection speed), but the difference compared to YOLOv8 is quite small, about 0.54% [35].

■ ***Fig. 1.*** Examples of Mosaic augmentation applied to traffic sign images

■ ***Table 1.*** Performance comparison between YOLOv8 and YOLOv5 models

Model	Precision, %	Recall, %	Dataset size	Processing
YOLOv8	84.62	76.40	10,940	Real-time
YOLOv5	81.80	75.94	10,940	Real-time

Note: Results obtained from training on the Vietnamese traffic sign dataset.

However, based on the superior accuracy figures of YOLOv8 in other studies as mentioned above, the research team chose this network model to apply traffic sign recognition in this study.

Material and methods

The system implements two primary operational modes to effectively monitor and catalog traffic signs. In the mobile user scenario, individuals utilize a smartphone application to capture and identify traffic signs, with the system automatically processing images, determining sign types, recording GPS locations, and updating a central database. Simultaneously, fixed installations using Raspberry Pi units with cameras provide continuous monitoring at critical traffic locations, enabling real-time detection and automated database updates. The processing workflow integrates these inputs through a standardized pipeline: beginning with image acquisition from camera, followed by preprocessing for standardization, YOLOv8-based detection and classification, database updating with GPS coordinates and sign information, and finally providing user feedback through the mobile interface. This unified approach ensures comprehensive traffic sign monitoring while maintaining system accessibility and reliability. The subsequent sections provide detailed technical specifications and implementation methods for each system component.

Training model

The experimental setup utilized high-performance computing hardware and software configurations to ensure optimal training conditions. The training environment comprised a Windows 11 operating system with an Nvidia 3080Ti GPU (12 GB memory), PyTorch framework (version 1.10.0a0), and CUDA toolkit (version 12.0). The development environment was configured using Anaconda Navigator, with Ultralytics framework implementation for object detection and segmentation tasks.

Dataset preparation and processing

The dataset development followed a systematic three-phase approach:

Phase 1: Data acquisition

A comprehensive traffic sign image dataset was constructed through multiple acquisition channels, including: high-resolution digital photography; curated online repositories; manual field photography. This multi-source approach ensured diverse representation of traffic sign variations.

Phase 2: Data annotation

Image annotation was performed using specialized annotation tools (LabelImg and CVAT) to create accurate bounding box labels for each traffic sign instance. The annotation process followed strict guidelines to ensure consistency and quality of labeled data.

Phase 3: Dataset partitioning

The experimental dataset consisted of 10,940 preprocessed images obtained from Roboflow. The dataset partitioning followed standard deep learning practices with 9,540 images for training, 786 for validation, and 614 for testing. This distribution was determined based on empirical studies of deep learning model training requirements and the need to prevent overfitting while ensuring adequate validation capabilities.

YOLOv8 configuration

The post-labeling phase required systematic configuration of the YOLOv8 model parameters through a COCO-format YAML file. This configuration framework specified three critical components: the training data pathway for model learning, the validation data location for performance assessment, and the comprehensive class definitions for traffic sign detection. The structured YAML configuration ensures consistent model training conditions and facilitates experimental reproducibility. The implemented configuration structure is defined as:

path: dataset
train: train/images
val: val/images
test: test/images
nc: 58 # Number of classes
names: ['DP.135', 'P.102', 'P.103a', 'P.103b', 'P.103c', 'P.104', 'P.106a',...]

Setting up parameters and proceeding with the training

Model implementation utilizes YOLOv8-L (large variant) architecture with input resolution of 640x640 pixels, configured with specific hyperparameters optimized for traffic sign detection. The learning rate follows a cosine decay schedule starting at 0.001, using batch size of 16, weight decay of 0.0005, and momentum of 0.937, with 3 warmup epochs. The network is structured to recognize Vietnamese traffic signs according to national standards, classifying 25 warning signs (W series)

■ ***Table 2.*** YOLOv8-L (large) training configuration parameters and their values for traffic sign detection optimization

Model parameters	Description	Value
Model	Model that we want to use	yolov8l.pt
Data	Data file	coco.yaml
Imgsz	Image size	640
Workers	The number of processes that generate batches in parallel	0
Device	Device to run training	0
Batch	The number of images processed before updating the model	2
Epochs	The number of times the learning algorithm will work to process the entire dataset	300
Patience	Epoches to wait for no observable improvement for early stopping of traing	50
Name	Folder name	yolov8-traffic-sign

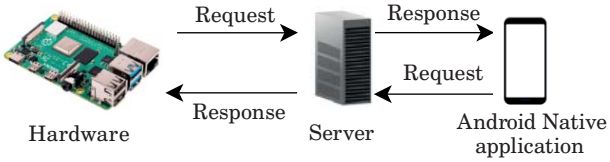
aligned with QCVN 41:2019, 20 regulatory signs (R series) following Circular 54/2019/TT-BGTVT, and 13 guide signs (G series) conforming to TCVN 7887:2018. The training follows a three-stage process: initial pre-training using ImageNet weights, domain adaptation using a COCO traffic sign subset, and final fine-tuning on our Vietnamese dataset using progressive learning rates. This configuration ensures optimal performance while maintaining compliance with local traffic standards and regulations. Through systematic parameter tuning and performance analysis, an optimal set of training parameters was established. Table 2 details these parameters, which were selected to maximize the model's traffic sign detection capabilities while maintaining computational efficiency.

The training time depends on the dataset size, the number of epochs, and the number of classes. After the training process completed, a trained YOLOv8 model is used to detect objects in real-time.

System architecture and design

The proposed system architecture integrates three fundamental components: server infrastructure, hardware platform, and software interface. Figure 2 shows this comprehensive architecture, demonstrating the interactions between components and data flow pathways.

The hardware implementation centers on a Raspberry Pi 4 computing platform, which manages image processing, display output, and camera input operations. This embedded system facilitates real-time traffic sign recognition through its integrated webcam module while maintaining continuous communication with the server infrastructure for data processing and result visualization.



■ ***Fig. 2.*** Overall system architecture

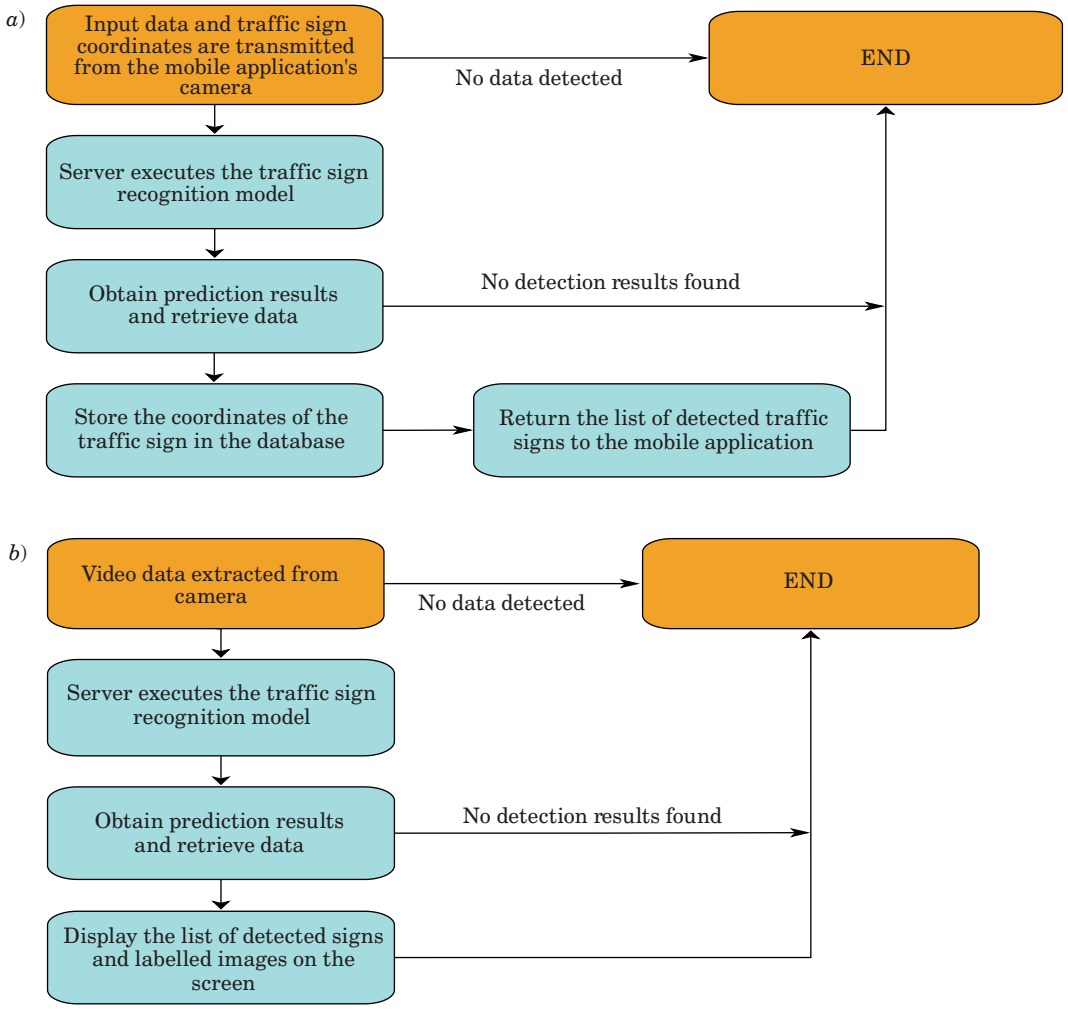
The software layer comprises an Android Native application developed in Java, implementing a client-server architecture for efficient data management and sign recognition processes. The application provides primary functionalities including user authentication, real-time sign detection, and location-based visualization through map integration. Secondary features encompass detailed sign information retrieval and supplementary data access capabilities.

The server infrastructure employs a dual-server architecture to optimize different processing requirements. A NodeJS-based server manages user authentication, session management, and geographical data storage, while a Flask-based Python server specializes in computational tasks related to traffic sign recognition. This separation of concerns ensures optimal performance for both data management and image processing operations.

Figure 3, *a*, *b* demonstrates how the web server acts as a bridge between software, hardware, and the system's AI server. Its role is to receive images collected by the mobile application for recognition processing and return the results to the mobile app while storing the location in the database. Additionally, on the hardware side, it receives streaming data to enable continuous recognition processing and displays the results on the web server interface. The diagram reveals the complete data flow and interaction pathways between all system components.

Hardware solution. The system hardware features these specifications: Dual-band networking via 2.4/5.0 GHz IEEE 802.11ac and gigabit ethernet connectivity. Four USB ports (two 3.0, two 2.0) enable camera connections. The processing unit incorporates a 40-pin GPIO interface, supports 1080p/30FPS video input, and drives a 320x480 LCD panel.

The processor handles concurrent display output and camera input operations. Visual data flows from camera sensor through the processing unit to the display module. Remote access operates through

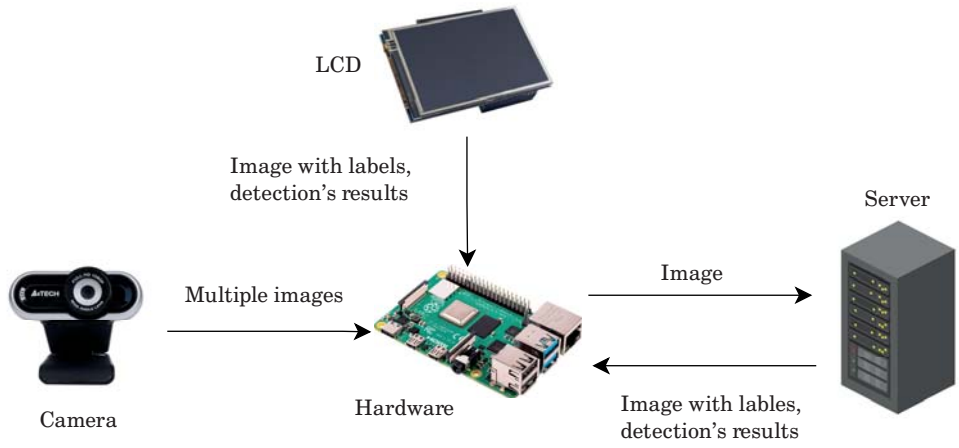


■ ***Fig. 3.*** System component interactions: *a* — web server and mobile application; *b* — web server and hardware

a NodeJS-based web portal. Captured images route to the central server for processing, with analysis results returned to the display. This configuration enables continuous visual analysis, supporting diverse research applications. Figure 4 illustrates the circuit architecture, detailing component interconnections between core processing unit, display system, and optical sensors.

Software solution. We have developed a mobile application that enables users to identify traffic signs by taking photos. This framework performs automatic sign classification, captures location data, and logs details to storage. When users submit images, the central processor analyzes them and returns comprehensive sign descriptions. The platform also enables alert parameter searches and profile customization options.

The application and interface of the traffic sign recognition system are built on Android Native (Java) software. The team chose this Android Native (Java) platform to take advantage of some of the main advantages of this platform that has been optimized for the Android operating system, providing superior performance and reliability when handling large data sets and diverse interactions between multiple devices. The platform uses the official Android language, which facilitates integration with many essential APIs, especially the Camera and Maps services, allowing full use of the system's features. Furthermore, thanks to the powerful and modern programming environment of Java and supported by continuous updates from the Android development community, access to cutting-edge technology is guaranteed to solve complex challenges. The combination of the Java platform and XML provides flexible user interface management that allows for a clear separation between logical and interface components. However, this implementation approach also poses certain challenges such as the time required for building and development is extended due to the need to build platform-specific solutions instead of using available cross-platform libraries. The platform requires quite comprehensive programming exper-



■ ***Fig. 4.*** Hardware implementation diagram

tise and extensive knowledge of the Android SDK, which can create certain difficulties for new developers. Additionally, the system scalability faces limitations such as expansion to other platforms such as iOS requires significant additional support library components compared to cross-platform alternatives. Despite these challenges, the advantages in performance, reliability, and system integration are adequate reasons to choose Android Native (Java) for this application.

Intermediary server. Figure 5 demonstrates our dual-server architecture implementation. The system employs a NodeJS server for database management and connectivity, alongside a Flask server dedicated to model storage and traffic sign recognition processing tasks.

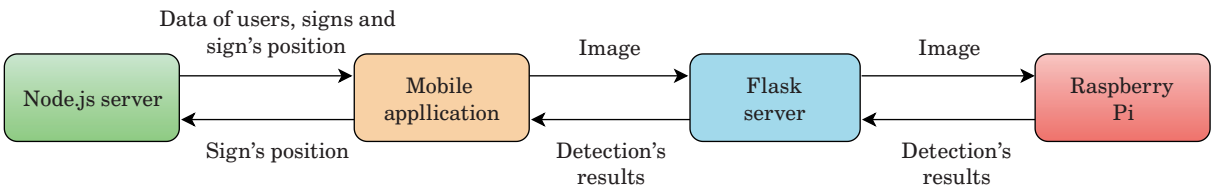
Building a server NodeJS. Used to manage all information about users, road signs, as well as all data about the location of road signs to be stored in the MySQL database.

Research has shown that Node.js demonstrates high memory efficiency compared to traditional multi-threading servers, allowing it to handle numerous concurrent requests without significant performance degradation [36]. These characteristics make it an ideal choice for managing data centrally, given its scalability, high performance, and cost-effectiveness. For the reasons mentioned above, we have built a Node.js server to handle the centralized management of data surrounding this mobile application.

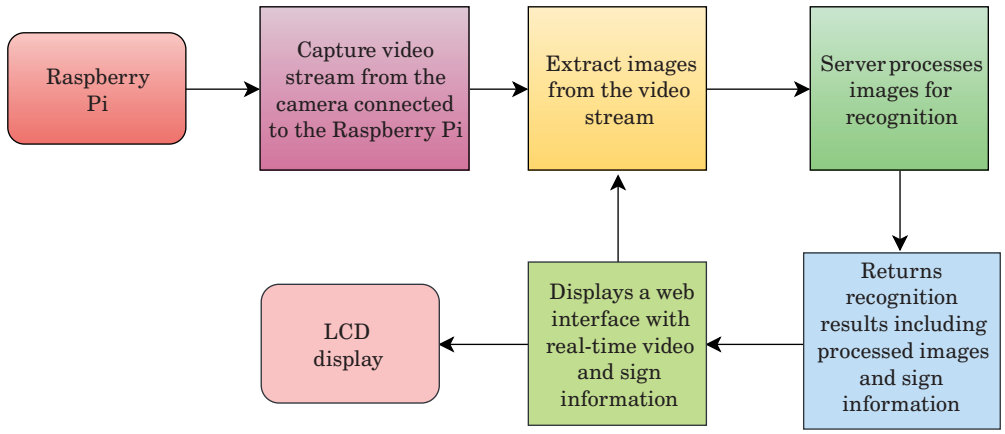
Building a server Flask. A web server was created to retrieve images, process recognition tasks, and provide a website interface capable of obtaining images from user devices for real-time recognition. The system implements Python as its core development language for neural network operations, utilizing its efficient code structure and established frameworks such as NumPy, Pandas, TensorFlow, and PyTorch. These libraries support effective implementation of complex models, making Python an ideal choice for research implementations. The system implements YOLOv8, the latest model in the YOLO family of object detection systems, which utilizes PyTorch to support traffic sign recognition with state-of-the-art performance.

Figure 6 shows the web server architecture, where Flask was selected as the framework for facilitating real-time traffic sign detection. While other robust frameworks like Django are available, research has shown that Flask's performance capabilities surpass alternatives due to its simple architecture [37]. The framework efficiently handles hundreds of requests per second without significant performance degradation, making it particularly suitable for rapid recognition tasks and real-time display of results on the web interface.

For further details, including access to the source code and comprehensive documentation, the complete repository, containing scripts, models, and additional resources, is available at (https://github.com/pacotha/DL_IOT2.git).



■ ***Fig. 5.*** Comprehensive system interaction flowchart



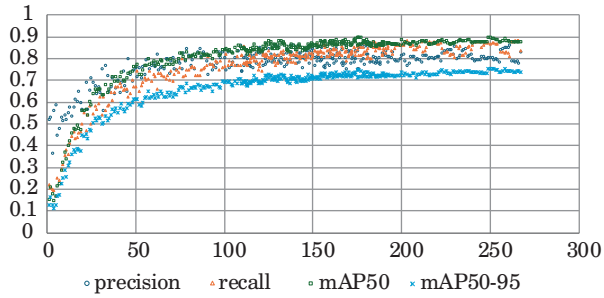
■ ***Fig. 6.*** Real-time image processing workflow on hardware components

The system employs a three-tier verification approach for sign status monitoring. Regular rescans by mobile users and fixed cameras provide updated sign condition data. A confidence scoring system flags potential sign removals when detection fails across multiple scans. Additionally, community reporting through the mobile interface allows users to flag missing or relocated signs, with automated verification triggered for reported changes.

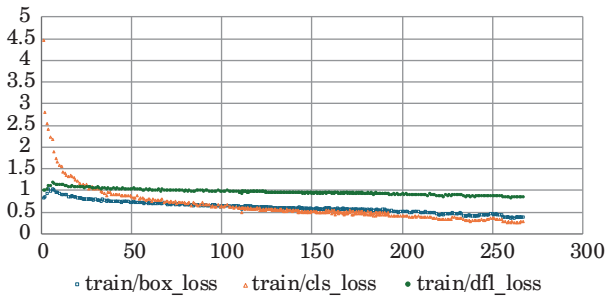
Results and discussion

The evaluation of our Traffic Sign Recognition system, implemented using the YOLOv8 architecture, high-lights its effective performance. Figure 7 illustrates the training losses for bounding box, class, and defocus over a 263-epoch training cycle. The box and class loss values, both below 0.5 after training process, indicate successful training and proficiency in object classification. Figure 8 presents the performance metrics of the YOLOv8 model, demonstrating that our system achieves a notable mAP50 score of approximately 89%. This score highlights the model's capability in recognizing objects with diverse levels of overlap across IoU thresholds ranging from 0.5 to 0.95. At the specific IoU threshold of 0.5, the system demonstrates significant accuracy, reaching nearly 90%, which validates its robust ability to accurately identify objects with moderate overlap.

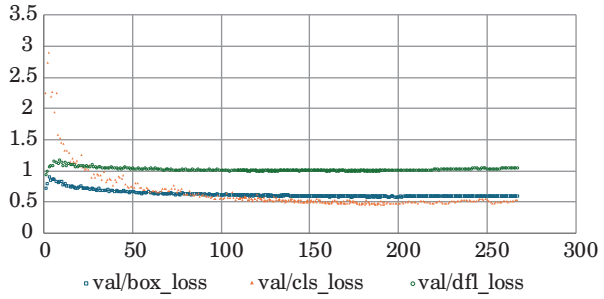
Figure 9 shows the validation phase results, where both box and class losses decreased and stabilized below 0.5, suggesting effective generalization of the model to unseen data. The defocus loss, indicated also known as dfl-loss, is a specialized component that improves object detection in challenging conditions such as blurry or unclear images. This loss remained stable around 1, indicating the model's resistance to overfitting while maintaining good



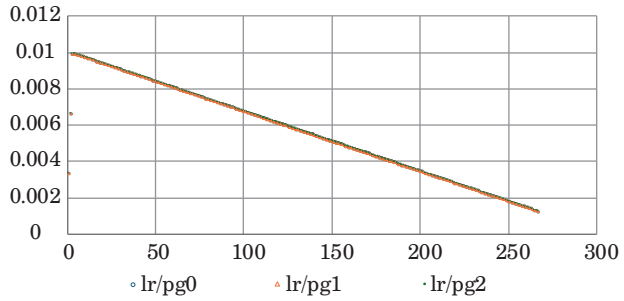
■ ***Fig. 8.*** Performance metrics evolution during training



■ ***Fig. 7.*** Evolution of training losses over 263 epochs



■ ***Fig. 9.*** Validation loss trends across training epochs

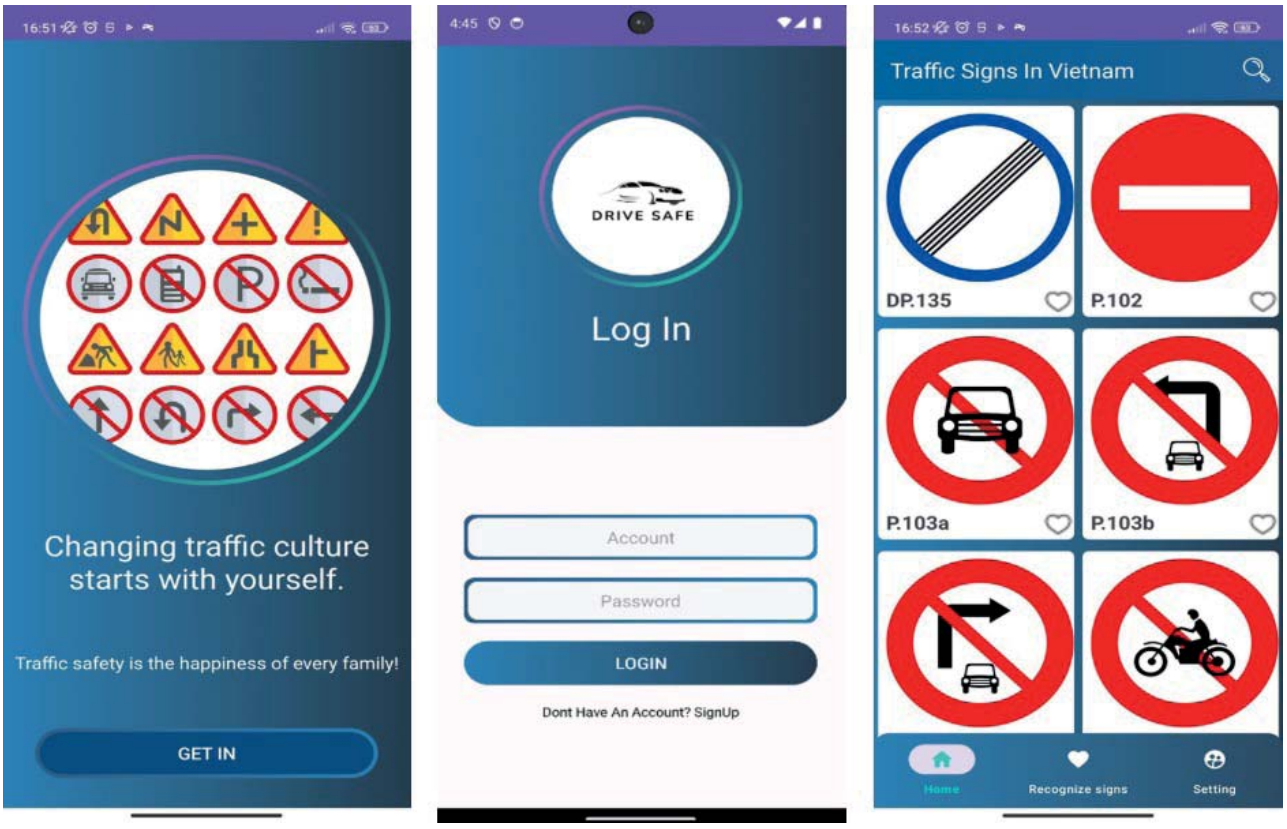


■ ***Fig. 10.*** Learning rate adaptation throughout training

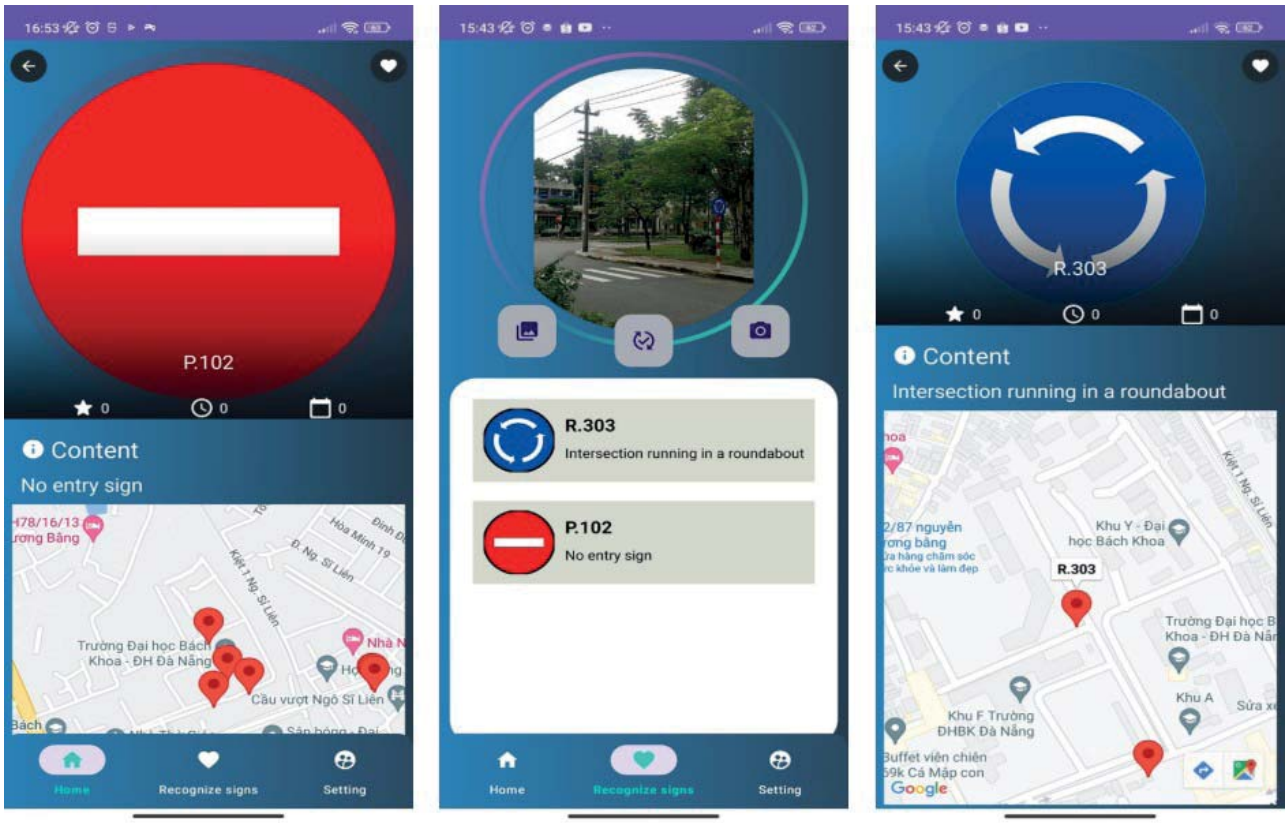
generalization capabilities. Figure 10 reveals the learning rate adaptation from 0.01 to 0.001, where the linear decrease and stabilization of training and validation losses suggest the model's convergence toward an optimal point.

The key performance metrics of the model include precision and recall, both of which are around 80%, indicating that this object detection model has a fairly good balance. This balance is important to minimize the rate of false positives while ensuring that relevant objects are correctly identified. Regarding the loss function, the box_loss decreases during training, indicating that the system effectively minimizes the error in training to predict bounding box coordinates. Similarly, the cls_loss decreases significantly during training, demonstrating the model's ability to accurately classify object types.

The mobile application interface is shown in Fig. 11. Although the positive metrics indicate that the YOLOv8 model has been properly trained, real-world evaluations show that the recognition ability is still not high. According to our survey on the street, the recognition accuracy of the model is around 70%. There are several reasons for this, including factors related to the training process and data processing. One important factor contributing to this phenomenon is overfitting, where the model focuses too much on the training data. The mobile application implements a hybrid storage approach. Frequently accessed map regions are cached locally, reducing server load and enabling offline functionality. Map data synchronizes automatically when connectivity resumes, with differential updates minimizing data transfer. User preferences determine cache size and update frequency. However, through practical observation, despite the limited recognition ability, the system still correctly places the recognized signs on Google Maps as shown in Fig. 12. This proves that our system works effectively in real-world applications.



■ ***Fig. 11.*** Mobile application user interface showcasing key features: sign recognition interface; location mapping display; user interaction components for data management



■ ***Fig. 12.*** Mobile application output displaying successful traffic sign recognition results with integrated location mapping and sign classification details

Table 3 presents a comprehensive performance comparison between the YOLOv8 model and two baseline methodologies: SVM [3] and MSER/HOG-enhanced SVM [4]. The evaluation metrics encompass recall, precision, and dataset utilization efficiency, providing quantitative insights into each model's detection capabilities in real-time traffic sign recognition.

Applying the YOLOv8 network to the system shows significant improvements in detection ability as shown by the recall data showing a 2.5x increase compared to traditional SVM methods. Specifically, the system's recall with YOLOv8 is 86.8% compared to the baseline SVM performance of 34.15%, showing a significant improvement in the ability to identify traffic signs. In addition, YOLOv8 shows better Precision than SVM, MSER/HOG, with an improvement of 39.67% (89.7% compared to 41.03% of the SWM method and 88.75% of the MSER/HOG method). This shows that YOLOv8 outperforms SVM in both object retention and classification. Compared to SVM-based MSER and HOG, YOLOv8 shows significantly higher recall value, with an improvement of 6.35% (86.8% vs. 81.45%). However, YOLOv8 has a slightly lower Precision, with a decrease of 8.05% (80.7% vs. 88.75%). This shows that SVM-based MSER and HOG have more accurate object retention, while YOLOv8 provides a higher recall.

Although the YOLOv8 network may show a slight decrease in precision compared to some methods, its significantly improved object retention and rapid object recognition, combined with its ability to learn efficiently from limited data, make it an effective choice for real-time traffic sign detection applications.

■ ***Table 3.*** Comparative analysis of traffic sign detection performance across different methodologies: YOLOv8, SVM, and MSER/HOG

Reference	Total sign	Precision, %	Recall, %
SVM [3]	104	41.03	34.15
MSER/ HOG [4]	104	88.75	81.35
YOLOv8-L	58	89.7	86.8

Conclusion

This study has successfully achieved its key objectives while advancing traffic sign recognition technology. First, our automated identification and classification system demonstrated strong techni-

cal performance, with the YOLOv8 architecture achieving an mAP50 score of 89% and 90% accuracy in controlled environments. The system enabled effective community participation through its mobile interface, though real-world accuracy averaged 70%, highlighting areas for future enhancement.

Second, we successfully implemented a dynamic traffic sign database with Google Maps integration, establishing Vietnam's first comprehensive traffic sign mapping system. The dual-server architecture, combining Node.js and Flask with Raspberry Pi edge computing, proved robust and scalable, enabling efficient data collection and real-time processing.

Third, our interactive mobile platform successfully delivered detailed sign information and educational content to users. The comprehensive mobile application facilitated user engagement and data collection, while maintaining system responsiveness and accessibility. Comparative analysis showed our approach outperformed traditional methodologies, with superior precision (89.7%) and recall (86.8%) compared to conventional SVM and MSER/HOG approaches.

Fourth, we established a scalable framework combining deep learning capabilities with community-driven data collection. This integration demonstrates the feasibility of our approach for traffic sign monitoring and maintenance. Future enhancements will focus on optimizing data augmentation techniques, improving model robustness against environmental variables, and refining real-time processing algorithms.

These achievements provide a foundation for intelligent transportation systems while identifying clear paths for future development. Initial deployment is planned for Da Nang's urban center, with phased expansion to other Vietnamese cities. The system's modular architecture enables integration with existing traffic management infrastructure and third-party ADAS systems. Future applications include real-time navigation assistance and automated traffic management. The system's architecture enables expansion to incorporate additional traffic management features and integration with broader transportation infrastructure, supporting the continued development of safer and more efficient road networks in Vietnam.

Acknowledgments

Pham Cong Thang (corresponding author) extends his gratitude to his colleagues at the Faculty of Information Technology, DUT, for their valuable feedback. The authors also sincerely appreciate the insightful comments and suggestions provided by the reviewers and editors.

References

1. Oviedo-Trespalacios O., Verity T., Barry W., Jane A. H. The impact of road advertising signs on driver behaviour and implications for road safety: A critical systematic review. *Transportation Research. Part A: Policy and Practice*, 2019, vol. 122, pp. 85–98. doi:10.1016/j.tra.2019.01.012
2. Maldonado-Bascón S., Lafuente-Arroyo S., Gil-Jiménez P., Gómez-Moreno H., López-Ferreras F. Road-sign detection and recognition based on support vector machines. *IEEE Transactions on Intelligent Transportation Systems*, 2007, vol. 8, no. 2, pp. 264–278. doi:10.1109/TITS.2007.895311
3. Greenhalgh J., Mirmehdi M. Real-time detection and recognition of road traffic signs. *IEEE Transactions on Intelligent Transportation Systems*, 2012, vol. 13, no. 4, pp. 1498–1506. doi:10.1109/TITS.2012.2209423
4. Wassouf Y., Korekov E. M., Serebrenny V. V. Decision making for advanced driver assistance systems for public transport. *2023 5th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)*, Moscow, Russian Federation, 2023, pp. 1–6. doi:10.1109/REEPE57272.2023.10086753
5. Ivanko D., Kashevnik A., Ryumin D., Kitenko A., Axyonov A., Lashkov I., Karpov A. MIDriveSafely: Multimodal interaction for drive safely. *Proceedings of the 2022 International Conference on Multimodal Interaction (ICMI '22)*, 2022, pp. 733–735. https://doi.org/10.1145/3536221.3557037
6. Lim K., Hong Y., Choi Y., Byun H. Real-time traffic sign recognition based on a general purpose GPU and deep-learning. *PLoS ONE*, 2017, vol. 12, no. 3, pp. e0173317. doi:10.1371/journal.pone.0173317
7. Shi Q., Abdel-Aty M. Big data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. *Transportation Research. Part C: Emerging Technologies*, 2015, vol. 58, pp. 380–394. doi:10.1016/j.trc.2015.07.011
8. Huangfu Z., Li S. Lightweight You Only Look Once v8: An upgraded You Only Look Once v8 algorithm for small object identification in unmanned aerial vehicle images. *Applied Sciences*, 2023, vol. 13, no. 22, pp. 12369. doi:10.3390/app132212369
9. Soylu E., Soylu T. A performance comparison of YOLOv8 models for traffic sign detection in the Robotaxi-full scale autonomous vehicle competition. *Multimedia Tools and Applications*, 2024, vol. 83, pp. 25005–25035. doi:10.1007/s11042-023-16451-1
10. Luo B., Sun Y., Tang Y., Gao X., Zhang F. A "Hardware-Friendly" foreign object identification method for belt conveyors based on improved YOLOv8. *Applied Sciences*, 2023, vol. 13, no. 20, Article 11464. doi:10.3390/app132011464

11. Kumar V. S., Prathibha D., Sreekanth Y., Bhanu M. Smart driver assistance system using raspberry pi and sensor networks. *Microprocessors and Microsystems*, 2020, vol. 79, iss. 8, Article 103275. doi:10.1016/j.micpro.2020.103275

12. Yang Y., Wang L., Liang J., Uddin M. Towards real-time traffic sign detection and classification. *IEEE Transactions on Intelligent Transportation Systems*, 2015, vol. 17, no. 7, pp. 2022–2031. doi:10.1109/TITS.2015.2462083

13. Akple M. S., Sogbe E., Atombo C. Evaluation of road traffic signs, markings and traffic rules compliance among drivers in Ghana. *Case Studies on Transport Policy*, 2020, vol. 8, no. 4, pp. 1295–1306. doi:10.1016/j.cstp.2020.09.004

14. Ravindran R., Basu P., Bandyopadhyay S., Roy S. Traffic sign identification using deep learning. *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, Las Vegas, NV, USA, 2019, pp. 318–323. doi:10.1109/CSCI49370.2019.00063

15. Siatwiko G. *Evaluating the effectiveness of the communication strategy used by the road transport and safety agency (RTSA) in promoting road safety in Lusaka urban*. Dis. the University of Zambia, 2023. Available at: https://dspace.unza.zm/handle/123456789/8090 (accessed 16 September 2024).

16. Hossain M., Hasan M., Ali M. A., Kabir M., Shawkat Ali A. S. Automatic detection and recognition of traffic signs. *2010 IEEE Conference on Robotics, Automation and Mechatronics*, Singapore, 2010, pp. 286–291. doi:10.1109/RAMECH.2010.5513177

17. Firame S., Jagtap K., Shendkar C., Dahiphale S., Sankapal S. Traffic sign detection. *Ijraset Journal for Research in Applied Science and Engineering Technology*, 2023, vol. 11, iss. IV, pp. 3438–3440. doi:10.22214/ijraset.2023.50966

18. Xu Q., Su J., Liu T. A detection and recognition method for prohibition traffic signs. *2010 International Conference on Image Analysis and Signal Processing*, Zhejiang, China, 2010, pp. 583–586. doi:10.1109/IASP.2010.5476048

19. Surinwarangkoon T., Nitsuwat S., Moore E. J. A traffic sign detection and recognition system. *International Journal of Circuits, Systems and Signal Processing*, 2013, vol. 7, no. 1, pp. 58–65.

20. He Y., Li T. A lightweight CNN model and its application in intelligent practical teaching evaluation. *2019 International Conference on Computer Science Communication and Network Security*, Sanya, China, 2019, pp. 22–23. doi:10.1051/matecconf/202030905016

21. Furedi L., Szolgay P. CNN model on stream processing platform. *2009 European Conference on Circuit Theory and Design*, Turkey, 2009, pp. 843–846. doi:10.1109/ECCTD.2009.5275115

22. Wang S., Sun L., Fan W., Sun J., Naoi S., Shirahata K., Fukagai T., Tomita Y., Ike A. An automated CNN recommendation system for image classification tasks. *2017 IEEE International Conference on Multimedia and Expo (ICME)*, Hong Kong, China, 2017, pp. 283–288. doi:10.1109/ICME.2017.8019347

23. Palit I., Lou Q., Acampora N., Nahas J., Niemier M., Hu X. Analytically modeling power and performance of a CNN system. *2015 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Austin, TX, USA, 2015, pp. 186–193. doi:10.1109/ICCAD.2015.7372569

24. Jocher G., Chaurasia A., Qiu J. *YOLO by Ultralytics*. Available at: https://github.com/ultralytics/ultralytics (accessed 07 October 2024).

25. Wang G., Chen Y., An P., Hong H., Hu J., Huang T. UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios. *Sensors*, 2023, vol. 23, no. 16, Article 7190. doi:10.3390/s23167190

26. Shaikh S., Chopade J., Kharate G. Object classification and tracking using scaled P8 YOLOv4 lite model. *Periodica Polytechnica Electrical Engineering and Computer Science*, vol. 67, no. 1, pp. 102–111. doi:10.3311/PPee.20685

27. Ma M., Pang H. SP-YOLOv8s: An improved YOLOv8s model for remote sensing image tiny object detection. *Applied Sciences*, 2023, vol. 13, no. 14, Article 8161. doi:10.3390/app13148161

28. Lim E. A., Hong T. W., Junoh A. K. A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond. *Mach. Learn. Knowl. Extr.*, 2023, vol. 5, pp. 1680–1716. doi:10.48550/arXiv.2304.00501

29. Wang C.-Y., Bochkovskiy A., Liao H.-Y. M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, 2023, pp. 7464–7475. doi:10.1109/CVPR52729.2023.00721

30. Akbarnezhad E. *YOLOv8 Projects #1 "Metrics, Loss Functions, Data Formats, and Beyond"*. Available at: https://www.linkedin.com/pulse/yolov8-projects-1-metrics-loss-functions-data-formats-akbarnezhad/ (accessed 16 October 2024).

31. Liu S., Qi X., Qin H., Jia J. Path aggregation network for instance segmentation. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, pp. 8759–8768. doi:10.1109/CVPR.2018.00913

32. Huang Z., Yang X., Zhang Q., Li Y. Research on traffic sign detection based on improved YOLOv8. *Journal of Computer and Communications*, 2023, vol. 11, no. 7, pp. 226–232. doi:10.4236/jcc.2023.117020

33. Neha Vishwakarma. *Understanding mosaic data augmentation*. Available at: https://www.analyticsvidhya.com/blog/2023/12/mosaic-data-augmentation/ (accessed 16 September 2024).

34. Sary I. P., Andromeda S., Armin E. U. Performance comparison of YOLOv5 and YOLOv8 architectures in human detection using aerial images. *Ultima Computing: Jurnal Sistem Komputer*, 2023, vol. 15, no. 1, pp. 8–13. doi:10.19184/ultima.2023.1510001

35. Chhetri N. *A comparative analysis of Node.js (server-side JavaScript)*. Culminating Projects in Computer Science and Information Technology, 2016. Available at: https://repository.stcloudstate.edu/csit_etds/5 (accessed 16 September 2024).

36. Idris N., Foozy C. F. M., Shamala P. A generic review of web technology: Django and Flask. *International Journal of Advanced Science Computing and Engineering*, 2020, vol. 2, no. 1, pp. 34–40. doi:10.1145/1124772.1124853

УДК 004.9
doi:10.31799/1684-8853-2025-2-2-15
EDN: MLUUSQ

Распознавание дорожных знаков с использованием системы Интернета вещей и глубокого обучения

Во Ньы Тхань[a], канд. техн. наук, преподаватель, orcid.org/0000-0001-5383-3119
Фам Конг Тханг[a], канд. техн. наук, преподаватель, orcid.org/0000-0002-6428-102X, pcthang@dut.udn.vn
Во Вьет Чыонг[a], студент, orcid.org/0009-0006-2601-0743
Хо Ван Тхао[a], студент, orcid.org/0009-0006-5016-9420
Чан Минь Куан[a], студент, orcid.org/0009-0001-8285-9171
Хоанг Нгуен Нят Минь[a], студент, orcid.org/0009-0000-1336-4426
[a]Университет Дананга — Университет науки и техники, Нгуен Лыонг Банг ул., 54, Дананг, 550000, Вьетнам

Введение: автоматическое распознавание дорожных знаков является важной частью модернизации систем безопасности дорожного движения и минимизации связанных с ним инцидентов. **Цель:** разработать систему распознавания дорожных знаков Вьетнама с интеграцией IoT-технологий, камер, датчиков и методов глубокого обучения. **Результаты:** представлено передовое решение компьютерного зрения, использующее нейронную архитектуру YOLOv8 для распознавания и классификации дорожных знаков Вьетнама в реальном времени. Система интегрирует местоположение обнаруженных знаков в Google Maps и предоставляет обширную базу данных местоположения всех знаков в городских районах. Сбор данных осуществляется путем первоначальной съемки изображений с реальных дорог и их объединения с существующим набором данных дорожных знаков Вьетнама из Kaggle. Для повышения надежности набора данных применяется техника аугментации Mosaic. Для распознавания дорожных знаков в режиме реального времени плата Raspberry Pi 4 отображает обнаруженные дорожные знаки на экране HMI, в то время как специально разработанное мобильное приложение обнаруживает, распознает и уведомляет пользователей об их местоположении. Обученная система достигает 89 % значения mAP50 и 90 % точности, что довольно хорошо для распознавания дорожных знаков. Кроме того, модель YOLOv8 демонстрирует 89,7 % точности (Precision) и 86,8 % полноты (Recall), что значительно выше по сравнению с методами SVM и MSER/HOG. **Практическая значимость:** созданная интегрированная в Google Maps система распознавания в реальном времени дорожных знаков удовлетворяет потребности регулирования дорожного движения, обеспечения безопасности и снижения аварийности на улицах Вьетнама.

Ключевые слова — распознавание дорожных знаков, YOLOv8, интеграция в Google Maps, обнаружение в реальном времени, безопасность на дорогах.

Для цитирования: Thanh V. N., Thang P. C., Truong V. V., Thao H. V., Quan T. M., Minh H. N. N. Traffic sign recognition through the use of an Internet of Things system and deep learning. *Информационно-управляющие системы*, 2025, № 2, с. 2–15. doi:10.31799/1684-8853-2025-2-2-15, EDN: MLUUSQ

For citation: Thanh V. N., Thang P. C., Truong V. V., Thao H. V., Quan T. M., Minh H. N. N. Traffic sign recognition through the use of an Internet of Things system and deep learning. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2025, no. 2, pp. 2–15. doi:10.31799/1684-8853-2025-2-2-15, EDN: MLUUSQ

УДК 004.627
doi:10.31799/1684-8853-2025-2-16-26
EDN: NMLHJC

Научные статьи
Articles

Применение пространственного подхода для задачи архивации телеметрических данных

И. В. Богачев[a], *канд. техн. наук, доцент, orcid.org/0000-0001-8203-7779, ilya.bogachev@yahoo.com*
А. В. Левенец[a], *доктор техн. наук, доцент, orcid.org/0000-0002-6973-6185*
Д. П. Зайцев[a], *лаборант, orcid.org/0009-0009-7241-0627*
[a]*Тихоокеанский государственный университет, ул. Тихоокеанская, 136, Хабаровск, 680035, РФ*

Введение: *для обработки данных телеметрии, которым в последние годы не уделялось достаточного внимания, существует ряд алгоритмов как специализированных, так и общего назначения, ни один из которых не показывает хотя бы приемлемую эффективность, и все имеют ряд фундаментальных ограничений.* ***Цель:*** *разработать эффективный алгоритм сжатия/восстановления наборов кадров телеметрических данных на основе пространственного подхода к их представлению и формата хранения их сжатого представления для задачи архивации.* ***Результаты:*** *предложен алгоритм сжатия наборов кадров телеметрических данных, представляющих собой трехмерные разностно-битовые матрицы, рассмотрены различные способы их отображения в тело фрактального куба. Для хранения сжатого представления данных разработан специализированный формат описания, который эффективно может быть применен в задаче архивации данных. Приведены оценки скорости и качества работы разработанного алгоритма, который по эффективности сравнивался с алгоритмом, основанным на построении префиксных кодов Хаффмана, и с алгоритмом DEFLATE, входящим в состав утилиты WinRAR, который в свою очередь является фактическим отраслевым стандартом для решения задач сжатия и архивации. В результате доказано существенное превосходство разработанного алгоритма над алгоритмами Хаффмана и DEFLATE в скорости и коэффициенте сжатия.* ***Практическая значимость:*** *использование предложенного алгоритма может привести к значительным технико-экономическим выгодам для промышленных предприятий и улучшению эффективности работы систем передачи телеметрических данных.* ***Обсуждение:*** *основным недостатком разработанного алгоритма является нелинейная зависимость времени сжатия от размерности сжимаемых данных, являющаяся следствием как невозможности эффективно сегментировать данные для их раздельной обработки, так и необходимости в многопроходной обработке. Решение данных проблем открывает ряд перспективных направлений для дальнейшего развития алгоритма.*

Ключевые слова *— наборы кадров, телеметрические данные, обратимое сжатие, архивация, телеметрические системы.*

Для цитирования: Богачев И. В., Левенец А. В., Зайцев Д. П. Применение пространственного подхода для задачи архивации телеметрических данных. *Информационно-управляющие системы*, 2025, № 2, с. 16–26. doi:10.31799/1684-8853-2025-2-16-26, EDN: NMLHJC

For citation: Bogachev I. V., Levenets A. V., Zaitsev D. P. Application of spatial approach to the problem of archiving measurement data. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2025, no. 2, pp. 16–26 (In Russian). doi:10.31799/1684-8853-2025-2-16-26, EDN: NMLHJC

Введение

В современном мире информация играет ключевую роль во многих сферах жизни общества, при этом темпы ее роста постоянно увеличиваются, что обусловлено развитием технологий и глобализацией. Ежедневно вырабатываются огромные объемы данных: видео, аудио, текстовые данные и многие другие их типы. Не являются исключением и данные, получаемые от систем телеметрии. Телеметрические данные играют критически важную роль в современном мире, предоставляя ценную информацию о том, как функционируют процессы и системы, позволяя принимать обоснованные решения и улучшать эффективность их работы.

Очевидно, что с ростом объемов генерируемых данных возникает необходимость их эффективного сжатия в целях как дальнейшей передачи по каналам связи, так и длительного хранения. Происходит неуклонный рост значимости данной процедуры, поскольку, сокращая объемы данных, можно снизить технико-экономические затраты на их передачу, хранение и обработку и повысить производительность работы систем.

Однако стоит учесть, что разные типы данных требуют разных подходов к сжатию. Для изображений и видео обычно используются специализированные алгоритмы сжатия с потерями, такие как JPEG или H.264/AVC [1, 2]. Для текстовых данных зачастую привлекают универсальные алгоритмы сжатия без потерь, такие как LZW или DEFLATE, последний основан на комбинации алгоритмов LZ77 и Хаффмана [3–6]. В случае данных телеметрии лишь немногие универсальные алгоритмы показывают хотя бы приемлемую эффективность, при этом обращение к специализированным алгоритмам сильно ограничено в связи с их спецификой [6–19]. Предложенные в этих работах алгоритмы сжатия данных, полученные от беспроводных сенсорных сетей [6, 7, 19] и систем телеизмерения

[9–12], взятые для обработки данных ультразвукового исследования [15] и биомедицинских сигналов [13, 14] хорошо работают с телеметрическими данными только определенного типа или обладающими только определенными свойствами, что значительно сужает область их эксплуатации.

Таким образом, в настоящее время существует практическая необходимость в разработке эффективных алгоритмов сжатия телеметрических данных широкого спектра.

Представление кадров телеметрических данных

В данной работе предложен алгоритм обратимого сжатия телеметрических данных для задачи их архивации, основанный на пространственном подходе [20], идея которого заключается в представлении отсчетов данных в разностно-битовом виде с их последующей группировкой и расположением на поверхности или внутри тела некоторого пространственного объекта. Применение такого подхода позволяет производить более глубокий анализ данных с сохранением корреляционных связей как между отсчетами, полученными в один момент времени и объединенными в рамках единого кадра данных, так и между кадрами, полученными за все время работы системы и объединенными в наборы.

В качестве пространственного объекта для отображения данных был выбран фрактальный параллелепипед, при этом для переноса бит предлагается использовать способы последовательного и группового отображения.

Последовательный способ подразумевает отображение кадров друг за другом в глубину, в результате чего формируется трехмерная битовая матрица $\mathbf{G_C}$ (фрактальный параллелепипед) размерностью $n \times m \times l$, где l — количество кадров в наборе, n — разрядность отсчетов и m — количество отсчетов в кадре.

Групповой способ представления является модификацией последовательного, но различия заключаются в том, что каждые четыре кадра формируют группу, а каждая группа располагается последовательно друг за другом. При этом если рассматривать получившийся объект со стороны геометрии, то он останется все тем же параллелепипедом, а при рассмотрении со стороны битового отображения — трехмерной матрицей $\mathbf{G_G}$ размерностью $(n \times 2) \times (m \times 2) \times (l / 4)$.

Важно отметить, что в случае группового представления классический способ расположения бит в телеметрическом кадре (слева направо от старшего разряда к младшему и сверху вниз от первого отсчета к последнему) следует изменить, так как при его использовании на текущей структуре будут нарушены корреляционные зависимости между кадрами. Чтобы не допустить этого, нужно отображать биты следующим образом:

– для первого кадра (верхнего левого) биты записываются, как и при классическом способе, по строкам сверху вниз и слева направо в каждой строке;

– для второго кадра (верхнего правого) биты записываются по строкам сверху вниз и справа налево в каждой строке;

– для третьего кадра (нижнего правого) биты записываются по строкам снизу вверх и слева направо в каждой строке;

– для четвертого кадра (нижнего левого) биты записываются по строкам снизу вверх и справа налево в каждой строке.

Также необходимо изменить и алгоритм преобразования данных в бинарный вид на более подходящий для полученной структуры, которое нужно производить следующим образом:

1) число представляется в бинарном виде в зависимости от знака по следующему принципу:

а) если число является положительным, оно представляется в бинарном виде разности самого числа и единицы;

б) если число является отрицательным, оно представляется в своем бинарном виде без учета знака;

в) ноль представляется в бинарном виде без изменений, т. е. нулем;

2) к полученной бинарной последовательности младшим разрядом дописывается один бит, отвечающий за знак числа и равный:

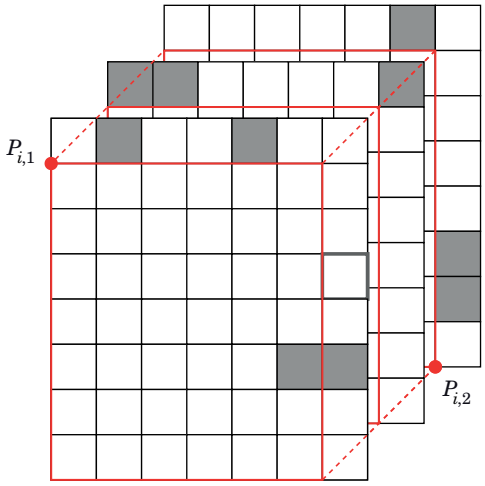
а) нулю, если число отрицательное или равно нулю;

б) единице, если число положительное.

Алгоритм сжатия, основанный на отображении данных в фрактальный параллелепипед

В основе данного алгоритма лежит поиск областей, сумма бит внутри которых больше некоторого порогового значения в теле фрактального параллелепипеда, образованного в ходе последовательного и группового способов отображения телеметрических кадров. При этом фиксировать обнаруженные области предлагается с использованием двух точек в системе координат фрактального параллелепипеда: $P_{i,1}[x_{i,1}, y_{i,1}, z_{i,1}]$ и $P_{i,2}[x_{i,2}, y_{i,2}, z_{i,2}]$ (где i — номер области), — как это представлено на рис. 1.

В качестве алгоритма поиска областей предлагается использовать подход, в основу которого положен алгоритм, разработанный Д. Каданом [21],



■ ***Рис. 1.*** Пример поиска области в теле фрактального параллелепипеда

■ ***Fig. 1.*** Example of searching for an area in the body of a fractal parallelepiped

изначально предназначенный для поиска максимальной суммы подмассива для заданного одномерного массива целых чисел. Тогда для поиска области с максимальной суммой в теле фрактального параллелепипеда нужно получить его взвешенное представление [20], после чего выявить вложенный параллелепипед с максимальной суммой значений всех его элементов в теле основного параллелепипеда. Для этого следует пройти по всем возможным вложенным параллелепипедам с подсчетом сумм значений входящих в них элементов.

Параллелепипед предлагается называть фрактальным, так как по окончании алгоритма поиска он разобьется на множество непересекающихся областей, обладающих свойством самоподобия, т. е. каждая область будет представлять собой все тот же фрактальный параллелепипед.

Следует отметить, что при групповом способе представления данных возможна ситуация, при которой количество кадров в наборе не равно числу, кратному четырем, вследствие чего объект, образованный после процедуры предварительной обработки, не будет являться параллелепипедом. Предлагается два варианта решения данной проблемы:

1) при формировании набора телеметрических кадров строго ограничивать количество кадров числом, кратным четырем;

2) ввести предшествующую первому шагу алгоритма поиска процедуру проверки количества кадров в наборе на соблюдение условия кратности четырем. В ситуации, когда условие соблюдается и количество кадров в наборе кратно четырем, выполняется описанный ранее алгоритм без каких-либо дополнительных операций. В случае, когда условие не соблюдается и количество кадров не кратно четырем, предлагается дополнить набор посредством дублирования последнего кадра до тех пор, пока условие кратности не выполнится.

Формат описания сжатых и архивированных данных

В результате работы алгоритма поиска областей весь набор кадров телеметрических данных разделится на три составляющие:

— список координат начала $P_{i,1}$ и конца $P_{i,2}$ найденных в ходе поиска областей;

— список координат бит, находящихся в найденных областях, значение которых равно единице ($P_{\text{един},i}$);

— список бит, не вошедших ни в одну из областей.

Формирование структуры сжатых данных происходит путем представления каждой из составляющих в бинарном виде и их последующей записи в единую последовательность. Однако данный процесс отличается в зависимости от выбранного способа представления набора телеметрических кадров.

Для алгоритма сжатия телеметрических данных, основанного на их отображении в теле фрактального параллелепипеда последовательным способом, процедуру формирования структуры описания сжатых данных можно описать с помощью следующей последовательности шагов:

1) определение количества бит, которое нужно выделить для хранения числа найденных областей;

2) определение количества бит, которое нужно выделить для хранения значений пары координат одной области;

3) определение количества бит, которое нужно выделить под хранение числа бит, находящихся в найденных областях, значение которых равно единице;

4) определение бит, которые не вошли ни в одну из найденных областей;

5) запись бит в определенном порядке.

Количество бит, которое нужно выделить для хранения числа найденных областей, определяется по формуле

$$Q_q = 1 + \lfloor \log_2 (q-1) \rfloor, \quad (1)$$

где q — количество найденных областей.

Количество бит, выделяемых для хранения числа бит, имеющих значение логической единицы в найденных областях, определяется по формуле

$$Q_g = 1 + \lfloor \log_2 (g-1) \rfloor, \quad (2)$$

где g — количество единичных бит в найденных областях.

Количество бит, которое нужно выделить для хранения значений одной координаты по осям X, Y и Z, определяется по следующим формулам:

$$Q_X = 1 + \lfloor \log_2 (n-1) \rfloor; \quad (3)$$

$$Q_Y = 1 + \lfloor \log_2 (m-1) \rfloor; \quad (4)$$

$$Q_Z = 1 + \lfloor \log_2 (l-1) \rfloor. \quad (5)$$

Для алгоритма сжатия телеметрических данных, основанного на их отображении в теле фрактального параллелепипеда последовательным способом, порядок записи бит представлен на рис. 2, *а*.

Декодирование данных из сжатого формата происходит с учетом того, что количество кадров в наборе и отсчетов в кадре заранее определено, а также известна разрядность данных. При этом выполняется следующая последовательность шагов:

1) строится пустая трехмерная матрица $\mathbf{G}_{\text{вос}}$ размерностью $n \times m \times l$;

2) считываются биты, хранящие количество найденных областей, после чего считываются биты, хранящие значения координат найденных областей. На основе полученной информации происходит восстановление найденных областей в матрице $\mathbf{G}_{\text{вос}}$;

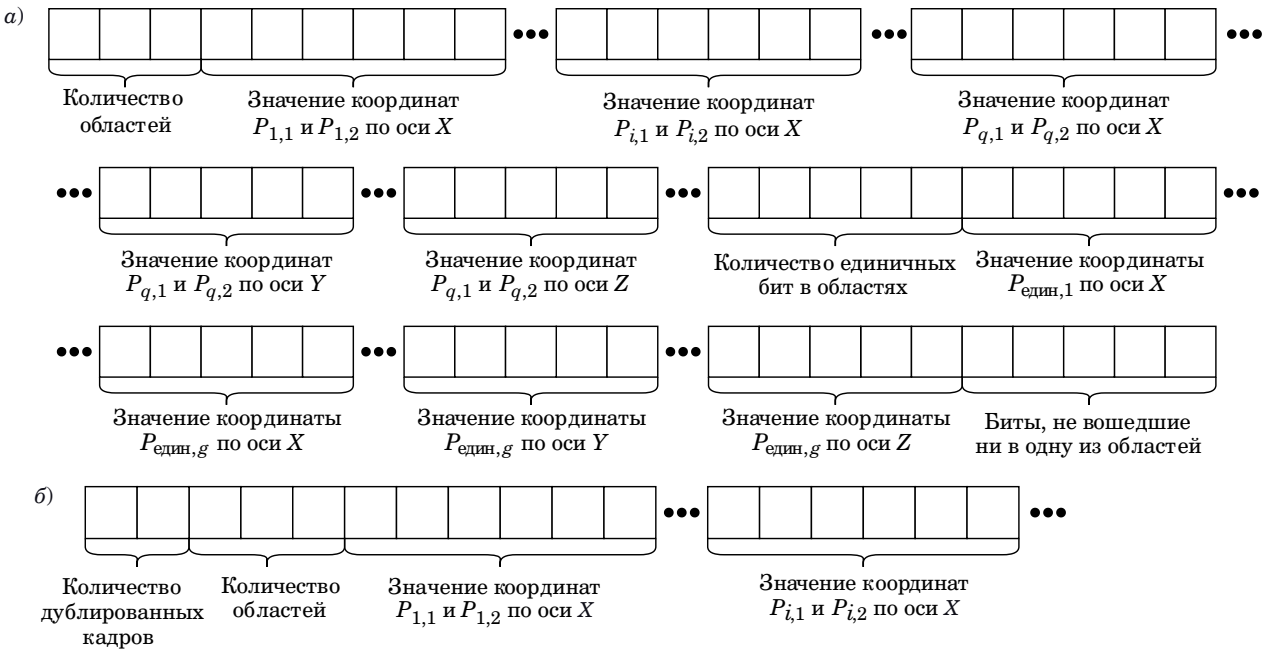
3) считываются биты, хранящие количество единичных бит, находящихся в найденных областях, после чего считываются биты, хранящие значения их координат. На основе полученной информации производится восстановление единичных бит, находящихся в найденных областях;

4) последовательно считываются биты несжатых данных и записываются в оставшиеся свободными элементы матрицы $\mathbf{G}_{\text{вос}}$;

5) выполняются процедуры обратного преобразования данных из бинарного вида в десятичный и обратного дельта-кодирования.

Для алгоритма сжатия телеметрических данных, основанного на их отображении в теле фрактального параллелепипеда групповым способом, процедура формирования структуры описания сжатых данных дополняется шестым шагом, который необходим для учета ситуации, при которой количество кадров в наборе не равно числу, кратному четырем. На данном шаге к полученной результирующей бинарной последовательности в начало добавляются несколько бит, отведенных для хранения числа дублирований последнего кадра набора (рис. 2, *б*).

Исходя из того, что при групповом способе отображения телеметрические кадры группируются по четыре штуки, можно сделать вывод, что возможное количество дублированных кадров варьируется от нуля в лучшем случае до трех в худшем. Вследствие чего для минимизации объема результирующей последовательности при кодировании числа дублированных кадров предлагается выделять именно два бита.



■ ***Рис. 2.*** Порядок записи сжатого набора кадров телеметрических данных при последовательном (*а*) и групповом (*б*) способе отображения

■ ***Fig. 2.*** The order of recording a compressed set of telemetry data frames with a sequential (*а*) and group (*б*) display method

Количество бит, необходимых для хранения числа найденных областей, а также количество бит, выделяемых для хранения числа единичных бит в этих областях, определяется по формулам (1) и (2) соответственно.

Однако стоит заметить, что формулы (3)–(5), по которым определяется количество бит, необходимое для хранения значений одной координаты по осям X, Y и Z, в данном случае нужно изменить в соответствии с размерностью матрицы $\mathbf{G}_{\mathbf{G}}$ следующим образом:

$$Q_X = 1 + \left\lfloor \log_2 \left((n \cdot 2) - 1 \right) \right\rfloor;$$

$$Q_Y = 1 + \left\lfloor \log_2 \left((m \cdot 2) - 1 \right) \right\rfloor;$$

$$Q_Z = 1 + \left\lfloor \log_2 \left(\left(\frac{l+d}{4} \right) - 1 \right) \right\rfloor,$$

где d — количество дублированных кадров.

Процедура декодирования данных из сжатого формата при групповом способе отображения кадров имеет всего два отличия:

1) на первом шаге строится трехмерная матрица $\mathbf{C}_{\mathbf{G\ вос}}$ размерностью $(n \times 2) \times (m \times 2) \times ((l + d) / 4)$;

2) после процедур обратного преобразования данных из бинарного вида в десятичный и обратного дельта-кодирования выполняется удаление дублированных кадров.

Применение предложенных выше форматов описания сжатых данных возможно только при условии, что при декодировании заранее известны такие показатели, как количество кадров в закодированном наборе, количество отсчетов в кадре (т. е. количество источников, с которых отправлялись данные), а также разрядность этих данных. Вследствие чего они не подходят для сжатия телеметрических данных в целях их архивации.

Для задачи сжатия телеметрических данных в целях их архивации нужно дополнить бинарные последовательности, полученные в ходе работы предложенных выше процессов формирования структуры описания сжатых данных, информацией о параметрах системы: количестве кадров в наборе, количестве отсчетов в кадре и их разрядности, — а также информацией о использованном методе сжатия, необходимой для полного восстановления данных. Следует учитывать, что при декодировании архивированных данных эта информация заранее не известна. Для кодирования каждого из параметров предлагается выделить по десять бит и вставить их в начало последовательности. Поскольку в работе показаны два варианта способа представления данных для одного алгоритма сжатия, для кодирования выбранного алгоритма хватит двух бит, которые нужно поставить в начало последовательности. Сформированную данным способом группу бит предлагается называть «служебной частью». Таким образом, служебная часть будет сформирована из последовательности бит, показанной на рис. 3.

Полученную бинарную последовательность, состоящую из служебной и описательной частей, надо закодировать в соответствии с расширенной восьмибитной таблицей ASCII. Вследствие чего полностью заархивированный файл будет состоять из однобайтовых символов, что в свою очередь говорит о том, что в худшем случае он получит семь избыточных бит по сравнению с еще не архивированной бинарной последовательностью. Стоит заметить, что такое преобразование необходимо в целях последующего сравнения результатов кодирования предложенного в работе алгоритма с уже известными алгоритмами сжатия, и данный этап можно пропустить в случае передачи данных по каналам связи систем телеметрии.

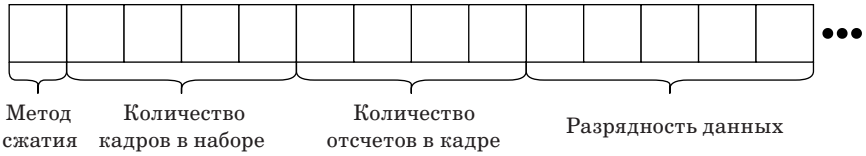
Предложенный алгоритм является асимметричным, так как процесс разархивации затрачивает в разы меньше времени и необходимых системных ресурсов в сравнении с процессом сжатия, при этом включая в себя две стадии: обработку архивированного формата данных и восстановление исходных данных.

Стадия обработки подразумевает декодирование с выявлением основных параметров, необходимых для восстановления исходных телеметрических данных, таких как применяемый алгоритм сжатия и размерность исходной фигуры.

Обработка архивированного формата данных состоит из следующих шагов:

1) выполняется восстановление бинарной последовательности путем представления каждого символа архивированного файла в двоичный вид в соответствии с таблицей ASCII;

2) определяется алгоритм сжатия путем извлечения из полученной бинарной последовательности первых двух бит;



■ ***Рис. 3.*** Порядок записи бит служебной части данных при архивации

■ ***Fig. 3.*** The order of recording the bits of the service part of the data during archiving

3) определяется количество кадров в наборе, количество отсчетов в кадре, а также разрядность данных, т. е. определяется размерность исходной фигуры на основе следующих 30 бит последовательности по 10 на каждый параметр путем преобразования данных из двоичной системы счисления в десятичную.

По окончании обработки архивированного формата становятся известны параметры системы, необходимые для корректного декодирования данных из сжатого формата и восстановления исходных телеметрических данных, описанных ранее для каждого из алгоритмов сжатия.

Исследование эффективности работы алгоритма

Исследование предлагаемого в работе алгоритма сжатия было проведено с использованием шести наборов кадров телеметрических данных, из которых:

— четыре набора содержали реальные данные, полученные от различных предприятий энергетики Дальнего Востока;

— два набора содержали данные, искусственно сгенерированные генератором случайных чисел по нормальному закону распределения.

Наборы реальных данных (НРД) включали в себя от 10 до 20 тысяч кадров, состоящих из восьмиразрядных отсчетов, полученных от 32 до 56 источников, при этом наборы искусственно сгенерированных данных (НИСД) состояли строго из 64 кадров, каждый из которых включал в себя строго 32 восьмибитных отсчета.

Стоит отметить, что, несмотря на то, что все исследования проводились лишь на восьмибитных данных, предложенный в работе алгоритм сжатия не имеет подобного ограничения и способен обрабатывать данные любой разрядности. Также важно заметить, что применение процедуры дельта-кодирования проводилось над всеми наборами данных, что обусловлено необходимостью проведения их исследования в равных

■ ***Таблица 1.*** Исследование пространственного алгоритма сжатия
■ ***Table 1.*** Research of spatial compression algorithm

Алгоритм сжатия	Параметр	Количество кадров в выборке	Набор данных					
			НРД1	НРД2	НРД3	НРД4	НИСД1	НИСД2
ФП-П	Коэффициент сжатия, ед.	8	13,31	4,56	10,85	8,59	6,75	5,97
		16	14,58	4,64	14,08	10,17	7,90	6,55
		32	19,68	4,85	16,38	10,95	8,44	6,85
		64	23,10	4,88	18,70	11,66	8,70	5,62
	Время сжатия, с	8	0,44	0,86	0,62	0,11	0,42	0,46
		16	0,12	0,43	0,23	0,74	0,13	0,12
		32	0,82	2,60	1,12	2,91	0,41	0,67
		64	4,19	24,13	6,21	9,29	2,11	3,91
ФП-Г	Коэффициент сжатия, ед.	8	12,02	4,36	10,85	8,77	6,65	5,79
		16	14,58	4,56	13,39	10,17	7,67	6,39
		32	19,12	4,69	15,64	10,97	8,65	6,71
		64	22,64	4,75	18,66	12,11	8,92	5,57
	Время сжатия, с	8	0,42	0,68	0,64	0,15	0,45	0,43
		16	0,16	0,27	0,35	0,51	0,17	0,16
		32	0,60	0,97	0,79	1,47	0,58	0,59
		64	1,92	5,63	2,78	5,17	1,61	1,93

условиях, при этом очевидно, что данная процедура не может повысить эффективность сжатия НИСД.

Представим результаты исследования эффективности работы предложенного алгоритма сжатия, основанного на отображении телеметрических данных в тело фрактального параллелепипеда последовательным (ФП-П) и групповым (ФП-Г) способом (табл. 1), при этом оценивались коэффициент сжатия и время сжатия.

Анализируя результаты, представленные в табл. 1, можно сделать вывод, что вне зависимости от количества кадров в выборке и степени стационарности исследуемых наборов наибольшие коэффициенты сжатия показывает алгоритм, основанный на представлении кадров в теле ФП-П. Однако стоит заметить, что по данному критерию ФП-Г проигрывает лишь в незначительной степени. То же самое касается и времени сжатия.

Рассматривая результаты, полученные на наборах с реальными и искусственно сгенерированными данными, можно увидеть, что во втором случае коэффициенты и время сжатия в разы меньше, что можно объяснить бóльшим динамическим диапазоном и отсутствием аномальных значений, вызванных случайными погрешностями в наборах данных при разностном представлении кадров, по сравнению с первым.

Важно заметить, что предложенный алгоритм показывает повышение коэффициента сжатия с увеличением количества кадров в выборке, при этом затрачиваемое на сжатие время значительно возрастает. На рис. 4, *а* и *б* приведены графики зависимости времени сжатия и коэффициента сжатия от количества кадров в сжимаемом наборе для алгоритма ФП-П.

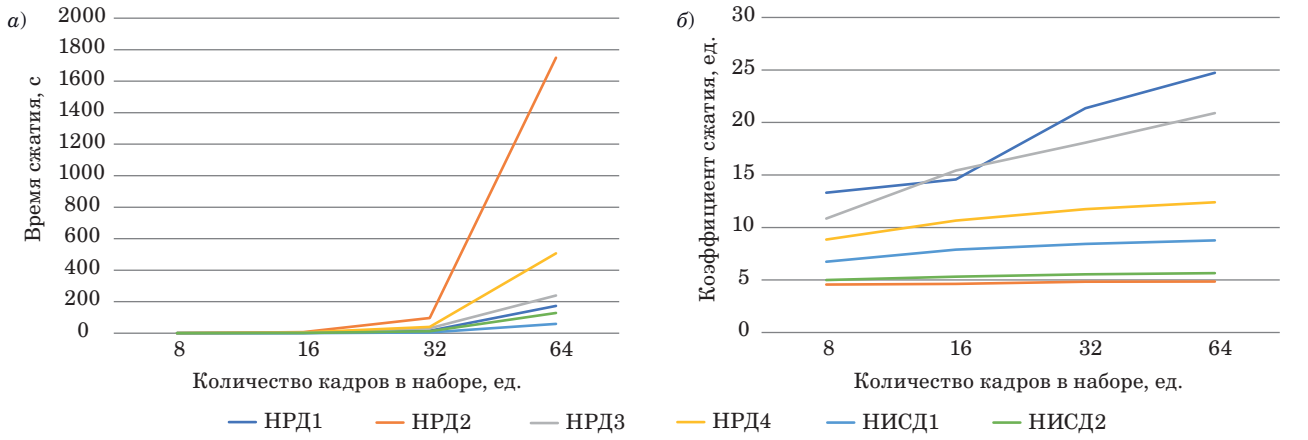
Анализ графиков позволяет утверждать, что с увеличением количества кадров в сжимаемом наборе возрастает время сжатия, причем стоит заметить, что если при увеличении объема кадров в наборе с восьми до 16 наблюдается незначительное увеличение времени сжатия, то при переходе с 16 до 32 и с 32 до 64 время сжатия возрастает на порядок. Вследствие чего можно сделать вывод, что зависимость времени сжатия от количества кадров в сжимаемом наборе носит экспоненциальный характер. При этом зависимость коэффициента сжатия от количества кадров в сжимаемом наборе носит окололинейный характер.

Основываясь на данных результатах, можно прийти к выводу, что использование пространственных алгоритмов сжатия на наборах с количеством кадров, превышающим 16 единиц, является нерациональным.

Для качественной оценки предложенного алгоритма сжатия были проведены исследования применения на тестовых наборах алгоритма, основанного на кодировании Хаффмана, а также алгоритма DEFLATE, реализованного в утилите WinRAR, являющейся фактическим лидером в области архивации данных [22].

Важно отметить, что при использовании алгоритмов Хаффмана и DEFLATE над тестовыми наборами проводилась предварительная обработка, а именно процедура дельта-кодирования. Результаты исследований представлены в табл. 2.

По результатам, представленным в табл. 2, можно сделать вывод, что использование алгоритма DEFLATE позволяет получить бóльшие коэффициенты сжатия, в свою очередь алгоритм Хаффмана имеет выигрыш во времени сжатия. Стоит заметить, что использование как алгорит-



■ ***Рис. 4.*** Зависимость времени сжатия (*а*) и коэффициента сжатия (*б*) от количества кадров в сжимаемом наборе для ФП-П

■ ***Fig. 4.*** Dependence of compression time (*а*) and compression ratio (*б*) on the number of frames in the compressed set for a compression algorithm based on mapping telemetry data into the body of a fractal parallelepiped in a sequential way

■ ***Таблица 2.*** Сравнение эффективности сжатия алгоритмом Хаффмана и DEFLATE
■ ***Table 2.*** Comparison of compression efficiency of the Huffman algorithm and DEFLATE

Алгоритм сжатия	Параметр	Набор данных					
		НРД1	НРД2	НРД3	НРД4	НИСД1	НИСД2
Хаффмана	Коэффициент сжатия, ед.	3,90	2,55	3,85	3,68	3,02	2,58
	Время сжатия, с	0,11	0,51	0,12	0,14	0,16	0,37
DEFLATE	Коэффициент сжатия, ед.	7,54	2,96	7,06	6,82	3,23	2,19
	Время сжатия, с	0,56	0,58	0,70	0,59	0,55	0,68

ма Хаффмана, так и DEFLATE является наиболее эффективным для стационарных наборов кадров.

Сравнение результатов обеих таблиц приводит к выводу, что алгоритм сжатия на основе пространственного подхода обеспечивает бóльший коэффициент сжатия: в среднем в 2,1 раза по сравнению с DEFLATE и в 2,9 раза по сравнению с алгоритмом Хаффмана. При этом по скорости работы предложенный алгоритм выигрывает у DEFLATE в среднем в 3,5 раза и в среднем в 1,2 раза у алгоритма Хаффмана. Однако стоит учитывать, что сравнение происходило на относительно малой выборке данных, состоящей из 16 кадров.

Важно заметить, что предлагаемый метод сжимает данные из наборов НИСД лучше, чем алгоритм Хаффмана. Причина заключается в том, что согласно принципам пространственного подхода набор исходных данных преобразуется в битовую последовательность, которая сворачивается в пространственный объект, при определенных условиях имеющий более упорядоченный вид, чем исходный набор данных. В свою очередь это позволяет получать более высокие значения коэффициента сжатия по сравнению с алгоритмами, которые учитывают энтропию данных, представленных классическим способом.

Таким образом, результаты, полученные в ходе проведенных исследований, позволяют сделать вывод, что предложенный в работе специализированный алгоритм пространственного сжатия превосходит популярные в настоящее время алгоритмы универсального назначения при задаче архивации телеметрических данных.

Заключение

В результате проведенного исследования разработан основанный на пространственном подходе алгоритм архивации телеметрических данных, использование которого может привести к значительному снижению требований к объемам хранилищ данных или к увеличению количества хранимых данных. Рассмотрены различные подходы к сжатию данных и показано, что пространственный подход позволяет достичь значительного результата в задаче сжатия и архивации телеметрических данных, выигрывая по коэффициенту сжатия у алгоритма DEFLATE в среднем в 2,1 и по скорости сжатия — 3,5 раза.

Полученные результаты демонстрируют, что использование пространственного подхода к представлению данных может быть эффективным способом решения проблемы избыточности телеметрических данных. Однако следует отметить, что для более широкого применения данного метода необходимы дополнительные исследования и оптимизация алгоритма. Так, необходим более эффективный способ поиска областей, чем подход, основанный на использовании алгоритма Кадана.

Таким образом, данная работа может послужить основой для дальнейших исследований в области сжатия и обработки телеметрических данных с использованием пространственного подхода, а также может быть полезной для разработчиков телеметрических систем и инженеров, работающих в промышленности.

Финансовая поддержка

Исследование выполнено за счет гранта Российского научного фонда № 24-21-20063, https://rscf.ru/project/24-21-20063/ и Министерства образования и науки Хабаровского края (соглашение № 113С/2024).

Литература

1. **Быковский П. С., Ляшенко В. А., Липницкий А. А., Пак М. А.** Анализ алгоритмов сжатия изображений JPEG И JPEG 2000. *Актуальные научные исследования в современном мире*, 2020, № 6-2 (62), с. 23–32. EDN: SDACKV
2. **Deepa Mathew, Binish M. C., Bijoy A. Jose.** Computationally efficient intra and inter mode decision in

H.264/AVC. *Procedia Computer Science*, 2020, vol. 171, pp. 360–368. doi:10.1016/j.procs.2020.04.037

3. **Hilal H. Nuha, Hassan Rizky Putra Sailellah, Ahmed Abo Absa.** Huffman tree compression and Lempel – Ziv coding using Java. *2024 International Conference on Decision Aid Sciences and Applications (DASA)*, Kingdom of Bahrain (Hybrid), 11–12 December 2024, 2024, 5 p. doi:10.1109/DASA63652.2024.10836416
4. **Бакулина М. П.** Эффективное сжатие без потерь больших массивов информации данных. *Проблемы информатики*, 2022, № 4 (57), с. 63–69. doi:10.24412/2073-0667-2022-4-63-69, EDN: HIQWPK
5. **Сидякин И. М.** Произвольный доступ к данным архивов телеметрической информации. *Наука и образование: научное издание МГТУ им. Н. Э. Баумана*, 2013, № 10, с. 343–354. doi:10.7463/1013.0616065, EDN: RSLHQP. http://technomag.bmstu.ru/doc/616065.html (дата обращения: 05 августа 2023).
6. **Chen S., Liu J., Wang K., Wu M.** A hierarchical adaptive spatio-temporal data compression scheme for wireless sensor networks. *Wireless Networks*, 2019, vol. 25, pp. 429–438. https://doi.org/10.1007/s11276-017-1570-6
7. **Chen C., Zhang L., Tiong R. L. K.** A new lossy compression algorithm for wireless sensor networks using Bayesian predictive coding. *Wireless Networks*, 2020, vol. 26, no. 8, pp. 5981–5995. https://doi.org/10.1007/s11276-020-02425-w
8. **Бакулина М. П.** Эффективное сжатие изображений на основе кодирования низкоэнтропийных источников. *Автометрия*, 2011, т. 47, № 1, с. 59–66. EDN: NEJTOZ
9. **Ломтев Е. А., Мясникова М. Г., Мясникова Н. В., Цыпин Б. В.** Совершенствование алгоритмов сжатия-восстановления сигналов для систем телеизмерений. *Измерительная техника*, 2015, № 3, с. 11–15. EDN: TORCTX
10. **Орешко В. В.** Алгоритмы устранения избыточности информации, передаваемой от бортовых телеметрических систем на Землю. *Ракетно-космическое приборостроение и информационные системы*, 2017, т. 4, № 2, с. 85–94. doi:10.17238/issn2409-0239.2017.2.85, EDN: ZIBDFR
11. **Тулекбаев Е. Т.** Эффективные методы сжатия телеметрической информации для наземных комплексов управления. *Вестник науки и образования*, 2017, № 10 (34), с. 14–20. EDN: ZMIRMZ
12. **Байбекова Ф. Н., Подольцев В. В., Беспалова Н. М., Сологубова Л. А.** Обзор способов снижения избыточности телеметрической информации. *Радиопромышленность*, 2019, № 2, с. 8–16. EDN: HUSNKK
13. **Capurro I., Lecumberry F., Martin A., Ramirez I., Rovira E., Seroussi G.** Efficient sequential compression of multichannel biomedical signals. *IEEE Journal of Biomedical and Health Informatics*, 2017, vol. 21, no. 4, pp. 904–916. doi:10.1109/JBHI.2016.2582683
14. **Huang C.-W., Ding J.-J.** Efficient EEG signal compression algorithm with long length improved adaptive arithmetic coding and advanced division and encoding techniques. *23rd International Conference on Digital Signal Processing (DSP)*, Proceedings, Shanghai, China, 2018, 5 p. doi:10.1109/ICDSP.2018.8631886
15. **Zhang X., Saniie J.** Unsupervised learning for 3D ultrasonic data compression. *2021 IEEE International Ultrasonics Symposium (IUS)*, Xi'an, China, 2021, 3 p. doi:10.1109/IUS52206.2021.9593654
16. **Shi X., Zhang Y., Dong X.** Evaluation of BAQ on Tiangong-2 interferometric imaging radar altimeter data compression. *22nd International Microwave and Radar Conference (MIKON)*, Poznan, Poland, 2018, 2 p. doi:10.23919/MIKON.2018.8405306
17. **Bose R., Pathak S.** Combined data encryption and compression using chaos functions. *Proceedings of SPIE – The International Society for Optical Engineering*, 2004, vol. 5561, pp. 164–175. doi:10.1117/12.561800
18. **Candes E. J., Wakin M. B.** An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 2008, vol. 25, no. 2, pp. 21–30. doi:10.1109/MSP.2007.914731
19. **Marco F. D., Godwin Sh., Antonio O.** Signal compression in wireless sensor networks. *Philosophical Transactions of the Royal Society A*, 2012, vol. 370, pp. 118–135. doi:10.1098/rsta.2011.0247
20. **Богачев И. В., Левенец А. В., Чье Е. У.** *Алгоритмы предварительной обработки и обратимого сжатия телеметрических кадров на основе геометрического подхода к представлению данных*. М., Русайнс, 2021. 166 с.
21. **Le N. P., Alouini M.-S.** Performance analysis of RIS-Aided THz wireless systems over α–μ fading: An approximate closed-form approach. *IEEE Internet of Things Journal*, 2024, vol. 11, no. 1, pp. 1328–1343. doi:10.1109/JIOT.2023.3288588
22. **Salomon D.** *Data Compression: The Complete Reference*. Fourth ed. Springer, 2007. 241 p.

UDC 004.627
doi:10.31799/1684-8853-2025-2-16-26
EDN: NMLHJC

Application of spatial approach to the problem of archiving measurement data

I. V. Bogachev[a], PhD, Tech., Associate Professor, orcid.org/0000-0001-8203-7779, ilya.bogachev@yahoo.com
A. V. Levenets[a], Dr. Sc., Tech., Associate Professor, orcid.org/0000-0002-6973-6185
D. P. Zaitsev[a], Laboratory Assistant, orcid.org/0009-0009-7241-0627
[a]Pacific National University, 136, Tihookeanskaya St., 680035, Khabarovsk, Russian Federation

Introduction: In the case of telemetry data, which has received little attention in recent years, there are a number of both specialized and general-purpose algorithms available for processing it, yet none of these have demonstrated even acceptable performance and have a number of fundamental limitations. **Purpose:** To develop an efficient algorithm for compressing/restoring measurement data frame sets based on a spatial approach to their representation, and a format for storing their compressed representation for the archiving task. **Results:** We propose an algorithm for compressing measurement data frame sets which are three-dimensional difference-bit matrices, and consider various methods for displaying them in the body of a fractal cube. To store a compressed data representation, we develop a specialized description format that can be implemented effectively in the task of archiving them. We give the estimates of the speed and quality of the developed algorithm, the efficiency of which is compared with both the algorithm based on the construction of Huffman codes and the DEFLATE algorithm, which is part of the WinRAR utility, the latter, in its turn, being the de facto industry standard for solving compression and archiving problems. We prove that the developed algorithm is significantly superior to the Huffman and DEFLATE algorithms in both speed and compression ratio. **Practical relevance:** The use of the proposed algorithm can lead to significant technical and economic benefits for industrial enterprises and improve the efficiency of telemetry data transmission systems. Discussion: The main disadvantage of the developed algorithm is the nonlinear dependence of the compression time on the dimension of the compressed data, which is a consequence of both the impossibility of effectively segmenting the data for their separate processing and the need for multi-pass processing. Solving these problems opens up a number of promising directions for further development of the algorithm.
Keywords — frame sets, measurement data, reversible compression, archiving, telemetry systems.

For citation: Bogachev I. V., Levenets A. V., Zaitsev D. P. Application of spatial approach to the problem of archiving measurement data. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2025, no. 2, pp. 16–26 (In Russian). doi:10.31799/1684-8853-2025-2-16-26, EDN: NMLHJC

Financial support

The research was supported by a grant from the Russian Science Foundation No. 24-21-20063 (https://rscf.ru/project/24-21-20063/) and the Ministry of Education and Science of the Khabarovsk Territory (agreement No. 113C/2024).

References

1. Bykovsky P. S., Lyashenko V. A., Lipnitsky A. A., Pak M. A. Analysis of JPEG and JPEG 2000 image compression algorithms. *Aktual'nye nauchnye issledovaniya v sovremennom mire* [Current Scientific Research in the Modern World], 2020, no. 6-2 (62), pp. 23–32 (In Russian). EDN: SDACKV
2. Deepa Mathew, Binish M. C., Bijoy A. Jose. Computationally efficient intra and inter mode decision in H.264/AVC. *Procedia Computer Science*, 2020, vol. 171, pp. 360–368. doi:10.1016/j.procs.2020.04.037
3. Hilal H. Nuha, Hassan Rizky Putra Sailellah, Ahmed Abo Absa. Huffman tree compression and Lempel – Ziv coding using Java. *2024 International Conference on Decision Aid Sciences and Applications (DASA)*, 2024, 5 p. doi:10.1109/DASA63652.2024.10836416
4. Bakulina M. P. Efficient lossless compression of large information arrays. *Problemy informatiki* [Problems of Informatics], 2022, no. 4 (57), pp. 63–69 (In Russian). doi:10.24412/2073-0667-2022-4-63-69, EDN: HIQWPK
5. Sidyakin I. M. Random access to telemetry data files compressed with deflate algorithm. *Science and Education: Scientific Publication of Bauman Moscow State Technical University*, 2013, no. 10, pp. 343–354 (In Russian). doi:10.7463/1013.0616065, EDN: RSLHQP. Available at: http://technomag.bmstu.ru/doc/616065.html (accessed 5 August 2023).
6. Chen S., Liu J., Wang K., Wu M. A hierarchical adaptive spatio-temporal data compression scheme for wireless sensor networks. *Wireless Networks*, 2019, vol. 25, pp. 429–438. https://doi.org/10.1007/s11276-017-1570-6
7. Chen C., Zhang L., Tiong R. L. K. A new lossy compression algorithm for wireless sensor networks using Bayesian predictive coding. *Wireless Networks*, 2020, vol. 26, no. 8, pp. 5981–5995. https://doi.org/10.1007/s11276-020-02425-w
8. Bakulina M. P. Efficient image compression by coding of low-entropy sources. *Optoelectronics, Instrumentation and Data Processing*, 2011, vol. 47, no. 1, pp. 59–66 (In Russian). EDN: NEJTOZ
9. Lomtev E. A., Myasnikova M. G., Myasnikova N. V., Tsypin B. V. Measurements in information technologies: improvement of signal compression-restoration algorithms for remote measurement systems. *Measurement Techniques*, 2015, vol. 58, no. 3, pp. 250–255. EDN: TORCTX
10. Oreshko V. V. Algorithms for elimination of the redundancy of the information transmitted from onboard telemetry system to the Earth. *Rocket-Space Device Engineering and Information Systems*, 2017, vol. 4, no. 2, pp. 85–94 (In Russian). doi:10.17238/issn2409-0239.2017.2.85, EDN: ZIBDFR
11. Tulekbayev Ye. T. Effective methods of compression of telemetric information for land complexes of management. *Vestnik nauki i obrazovaniya* [Bulletin of Science and Education], 2017, no. 10, pp. 14–20 (In Russian). EDN: ZMIRMZ
12. Baybekova F. N., Podoltsev V. V., Bespalova N. M., Sologubova L. A. Overview of the ways to reduce telemetric information redundancy. *Radio Industry*, 2019, no. 2, pp. 8–16 (In Russian). EDN: HUSNKK
13. Capurro I., Lecumberry F., Martin A., Ramirez I., Rovira E., Seroussi G. Efficient sequential compression of multichannel biomedical signals. *IEEE Journal of Biomedical and Health Informatics*, 2017, vol. 21, no. 4, pp. 904–916. doi:10.1109/JBHI.2016.2582683
14. Huang C.-W., Ding J.-J. Efficient EEG signal compression algorithm with long length improved adaptive arithmetic coding and advanced division and encoding techniques. *23rd International Conference on Digital Signal Processing (DSP)*, Proceedings, Shanghai, China, 2018, 5 p. doi:10.1109/ICDSP.2018.8631886
15. Zhang X., Saniie J. Unsupervised learning for 3D ultrasonic data compression. *2021 IEEE International Ultrasonics Symposium (IUS)*, Xi'an, China, 2021, 3 p. doi:10.1109/IUS52206.2021.9593654
16. Shi X., Zhang Y., Dong X. Evaluation of BAQ on Tiangong-2 interferometric imaging radar altimeter data compression. *22nd International Microwave and Radar Conference*

(MIKON), Poznan, Poland, 2018, 2 p. doi:10.23919/MIKON.2018.8405306
17. Bose R., Pathak S. Combined data encryption and compression using chaos functions. *Proceedings of SPIE — The International Society for Optical Engineering*, 2004, vol. 5561, pp. 164–175. doi:10.1117/12.561800
18. Candes E. J., Wakin M. B. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 2008, vol. 25, no. 2, pp. 21–30. doi:10.1109/MSP.2007.914731
19. Marco F. D., Godwin Sh., Antonio O. Signal compression in wireless sensor networks. *Philosophical Transactions of the Royal Society A*, 2012, vol. 370, pp. 118–135. doi:10.1098/rsta.2011.0247
20. Bogachev I. V., Levenets A. V., Chie E. U. *Algoritmy predvaritel'noj obrabotki i obratimogo szhatiya telemetricheskih kadrov na osnove geometricheskogo podhoda k predstavleniyu dannyh* [Algorithms for Pre-Processing and Reversible Compression of Telemetry Frames Based on a Geometric Approach to Data Representation]. Moscow, Rusajns Publ., 2021. 166 p. (In Russian).
21. Le N. P., Alouini M.-S. Performance analysis of RIS-Aided THz wireless systems over α–μ fading: An approximate closed-form approach. *IEEE Internet of Things Journal*, 2024, vol. 11, no. 1, pp. 1328–1343. doi:10.1109/JIOT.2023.3288588
22. Salomon D. *Data Compression: The Complete Reference.* Fourth ed. Springer, 2007. 241 p.

УВАЖАЕМЫЕ АВТОРЫ!

Научные базы данных, включая Scopus и Web of Science, обрабатывают данные автоматически. С одной стороны, это ускоряет процесс обработки данных, с другой — различия в транслитерации ФИО, неточные данные о месте работы, области научного знания и т. д. приводят к тому, что в базах оказывается несколько авторских страниц для одного и того же человека. В результате для всех по отдельности считаются индексы цитирования, что снижает рейтинг ученого.

Для идентификации авторов в сетях Thomson Reuters проводит регистрацию с присвоением уникального индекса (ID) для каждого из авторов научных публикаций.

Процедура получения ID бесплатна и очень проста, есть возможность провести регистрацию на 12 языках, включая русский (чтобы выбрать язык, кликните на зеленое поле вверху справа на стартовой странице): https://orcid.org

УДК 004:623-9
doi:10.31799/1684-8853-2025-2-27-36
EDN: QEFDEK

Научные статьи
Articles

Классификация прецедентов группового управления

В. К. Абросимов*[а], *доктор техн. наук, старший научный сотрудник, orcid.org/0000-0001-6701-2389, avk787@yandex.ru
Е. С. Михайлова*[а,б], *аспирант, orcid.org/ 0009-0008-0917-5524
[а]*Главный научно-исследовательский испытательный межвидовой центр перспективного вооружения Министерства обороны РФ, Октябрьская ул., 26, кор. 1, Москва, 127018, РФ*
[б]*ПАО «НПО «Алмаз» им. академика А. А. Расплетина, Авиамоторная ул., 57, кор. 14, Москва, 111024, РФ*

Введение: *в условиях развития автономности беспилотных средств важно совершенствовать методы повышения автономности. Одним из таких методов является использование прецедентов, которые позволяют системе принимать решения в текущей ситуации, опираясь на прошлый опыт.* ***Цель:*** *для эффективного использования прецедентов разработать классификацию прецедентов группового управления с помощью иерархического и фасетного метода, оценить емкость и коэффициент заполненности классификации таких прецедентов.* ***Результаты:*** *введены новые определения прецедентного состояния, прецедентной коллизии, прецедентной ситуации, прецедентных фактов. Подтверждена применяемость распространенных методов классификации — иерархического и фасетного — и к прецедентам группового управления. Приведена глубина иерархической классификации для различных классов, достаточная для решения практических задач. Введены два новых определения: «ядро» и «бахрома» прецедента, — которые требуются для построения фасетной классификации. Они позволяют выбрать несколько вариантов типовых прецедентов, состояния группы в которых можно описать только «ядрами» признаков, не обращая особого внимания на «бахрому». На основе логико-лингвистической модели создана фасетная формула для прецедентов, учитывающая «ядро» и «бахрому» прецедента. Также проведена оценка емкости и коэффициента заполненности классификации прецедентов, в результате для каждой классификации определена емкость. Анализ признаков прецедентов позволяет сделать вывод о том, что возникновение прецедента в групповом управлении является достаточно сложным явлением, учитывающим различные аспекты прецедента. В результате на основе классификации прецедентов сформулировано определение группового прецедента.* ***Практическая значимость:*** *классификация прецедентов позволяет более эффективно управлять беспилотными средствами, так как облегчает поиск подходящих прецедентов при необходимости принятия решения в ходе выполнения миссии.*

Ключевые слова *— прецедент, актор, группа, рой, среда, решение, ситуация, состояние.*

Для цитирования: Абросимов В. К., Михайлова Е. С. Классификация прецедентов группового управления. *Информационно-управляющие системы*, 2025, № 2, с. 27–36. doi:10.31799/1684-8853-2025-2-27-36, EDN: QEFDEK
For citation: Abrosimov V. K., Mikhailova E. S. Classification of group control precedents. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2025, no. 2, pp. 27–36 (In Russian). doi:10.31799/1684-8853-2025-2-27-36, EDN: QEFDEK

Введение

Классификация является необходимым элементом любого исследования. Как хорошо известно, сущность классификации состоит в «…распределении *предметов* какого-либо *рода* на взаимосвязанные *классы* согласно существенным *признакам*, присущим предметам данного *рода* и отличающим их от предметов других родов…» [1]. В процессе классификации понятие делится по определенному основанию (признаку, критерию) на множество классов, классы — на множество подклассов и т. д. При разделении на классы основой является сходство либо различие классов между собой с учетом вводимых правил, позволяющих с высокой степенью достоверности отнести объект к заданному классу. Главным аспектом для классификации считается система качественных или количественных классификационных признаков (свойств или характеристик объекта), присущих для данного класса. Количество признаков классификации зависит от сложности и количества классифицируемых объектов и целей классификации.

Понятие «прецедент» имеет известное и несложное объяснение [2], но несколько принципиально различающихся смыслов. Можно говорить о «гуманитарном прецеденте», ведущем свою историю из Древнего Рима и относящемся к предметной области «право». Далее будем рассматривать так называемый «технический прецедент», акторами которого являются, как правило, технические объекты, обученные соответствующим парадигмам поведения в сложной, часто противодействующей среде. Теория прецедентов в области технических систем стала активно развиваться в связи с созданием систем автоматизации обработки инцидентов в автоматизированных системах, управления технологическими процессами и др.

В последнее время тренд автономности активно развивается и внедряется в беспилотные средства. Следовательно, в условиях роста автономности необходимо совершенствовать методы

повышения автономности, и одним из таких методов являются прецеденты. Прецеденты позволяют системе принимать решения, опираясь на прошлый опыт. Для эффективного использования прецедентов необходима их классификация, которая до настоящего времени не разработана, что делает поставленную цель актуальной.

Формирование прецедентной ситуации при выполнении групповой миссии

Нередко при исследовании групповых миссий рассматриваются весьма важные, но все же частные вопросы, например формирование группы, распределение ролей, сроки выполнения задач, выбор стратегии выполнения задания, принятие коллективных решений [3, 4] и др. Как правило, исследование производится с точки зрения как бы внешнего наблюдателя, сама же группа является «не живой», а материальной, технической системой. Вместе с тем все чаще сущность и поведение технической системы, в которой реализуются технологии искусственного интеллекта, связываются и сравниваются с человеческими свойствами. Так, роботов уже наделяют эмоциями [5], обучают писать вполне осознанные тексты и т. д. [6, 7]. Однако групповое управление на данный момент все еще находится на начальных стадиях развития и в реальной практике почти не применяется из-за сложности реализации методов управления группой, и чаще всего используются компьютерные модели [3, 8]. В связи с этим сложно получать прецеденты, необходимые для составления базы данных, таким образом, для набора нужного количества прецедентов предлагается использовать не только данные, полученные из реальной практики, но и синтетические данные.

Интересно рассмотреть выполнение группой объектов коллективной миссии «изнутри», с точки зрения самой группы как своего рода «живого организма». Это тем более возможно, если, во-первых, наделить объекты управления интеллектом, например обучить их определенным стратегиям поведения в сложных ситуациях [9], и, во-вторых, представить объекты управления как «сеть слабо связанных решателей частных проблем (агентов), которые существуют в общей среде и взаимодействуют друг с другом для достижения тех или иных общих целей системы и/или частных целей отдельных агентов. Взаимодействие может выполняться агентами либо напрямую — путем обмена сообщениями, либо косвенно, когда одни агенты воспринимают присутствие других агентов через изменения во внешней среде, с которой они взаимодействуют» [8]. В этом случае полагают, что агенты обладают рядом так называемых «ментальных свойств», в том числе знаниями, убеждениями, желаниями, намерениями, правдивостью, рациональностью и, что, возможно, самое главное, общественным поведением, предполагающим способность функционировать в сообществе с другими агентами, обмениваясь с ними сообщениями с помощью некоторого общепонятного языка коммуникаций [10].

Итак, рассмотрим выполнение коллективной миссии гетерогенной группой интеллектуальных агентов. В наших допущениях эти агенты считаются обученными различным стратегиям поведения, способными к адаптации и выработке коллективных решений в ходе «переговоров» на конкурентной или коллективной основе.

В гетерогенной группе выделим несколько акторов — действующих объектов, моделируемых как агенты и совершающих совместные действия, направленные на выполнение коллективной миссии в некоторой пространственно-временной среде. В такой среде при выполнении миссии в течение определенного промежутка времени или «в данном моменте» может складываться совокупность различных условий, в которых наблюдаются некоторые факты, регистрируются события, фиксируются ситуации, т. е. существует некоторая динамика, требующая действий. Назовем такую совокупность условий *прецедентным обстоятельством*.

В возникших прецедентных обстоятельствах могут появляться противоречия между целями выполнения миссии и интересами среды, приводящие к столкновению противоположных стратегий миссии группы и противодействию среды. Такие противоречия целесообразно называть *прецедентной коллизией*.

В теории управления активно используется термин «*состояние*». Состояние определяется как набор значений (или интервалов допустимых значений) переменных параметров объекта [1]. Состояние устойчиво до тех пор, пока над ним не будет произведено действие.

Группа в процессе выполнения коллективной миссии все время находится в некотором состоянии, характеризующемся текущим состоянием группы и окружающей среды. Состояние описывает переменные свойства конкретного объекта. Такое состояние подразумевает появление во времени ситуаций, которые требуют принятия решений.

Итак, *прецедентное состояние* — это состояние, которое установлено или определено предшествующим действием, решением или событием и которое может служить основой для последующих действий или решений. Прецедентное состояние может быть использовано для определения последующих действий группы.

Прецедентные состояния могут описываться совокупностью прецедентных фактов — т. е. зафиксированных средствами измерений объекта управления явлений, отражающих реальное проявление действующего на группу объективного или субъективного фактора. Наступление темноты, обнаружение средства противодействия, огневого воздействия на объекты группы, неисправность видеокамеры системы технического зрения есть примеры прецедентных фактов.

Явление, наступившее одновременно или последовательно в течение непродолжительного промежутка времени в результате действия разнообразных факторов, свершившихся определенных обстоятельств и совокупности зафиксированных фактов, связанных единой логикой возникновения и последствий, целесообразно назвать *прецедентным событием*. Пропадание связи между объектами группы в ходе воздействия средств радиоэлектронной борьбы, неисправность системы управления объектом, уничтожение объекта при огневом воздействии являются примерами прецедентных событий.

Совокупность связанных единой логикой, свершившихся одновременно или последовательно в течение определенного промежутка времени прецедентных событий, приведших к определенному стечению обстоятельств, возникновению новых сцен в окружающей среде, изменению параметров обстановки среды выполнения миссии, вызвавшая необходимость принятия решений по дальнейшим действиям, формирует *прецедентную ситуацию*. Невозможность взаимодействия объектов по информации и управлению, возникновение необходимости реконфигурации группы в связи с требованием перераспределения задач и ролей в группе являются примерами прецедентных ситуаций.

Таким образом, при выполнении группой, находящейся в определенном прецедентном состоянии, коллективной миссии вследствие прецедентной коллизии сформировалась прецедентная ситуация, требующая принятия решения.

Разработка классификационных признаков

Как следует из изложенного, в процессе выполнения нескольких категорий коллективных миссий объекты разных классов объединяются в различные виды групп, выполняют различные типы задач, фиксируют текущее состояние, факты, события, сталкиваются с коллизиями, приводящими к сложным ситуациям, принимают прогнозируемо эффективные решения и реализуют их, изменяя состояние группы.

Выделим классы и классификационные признаки классов всех затрагиваемых выше сущностей прецедентов применительно к задачам оборонного значения, решаемым автономными объектами управления в составе групп (табл. 1).

Выделено 11 типов классов прецедентов с независимыми классификационными признаками, в каждом типе несколько классов (от двух до 10). В последнем столбце табл. 1 приведена минимальная глубина классификации для различных классов, достаточная для решения практических задач.

Принципиально возможны два типа классификации — иерархическая и фасетная.

Иерархическая классификация прецедентов

При использовании иерархических методов классификации строится некоторое дерево, где каждый уровень представляет собой более детализированную категорию. Основным видом отношений при этом является подчинение.

Анализ показывает, что примерно 80 % классов сущностей, образующих прецедент, имеют четко выраженную иерархическую структуру. Действительно, такие классы, как объекты, мощность группы, целеполагание легко разбиваются по дополнительным классификационным признакам на более мелкие классы.

Приведем в качестве примера второй уровень иерархии для некоторых классов (табл. 2).

К числу недостатков иерархической классификации относятся жесткость самой структуры, а также проблемы распределения новых сущностей по классам, если объект обладает признаками, которые заранее не были предусмотрены. Авторская практика построения алгоритмов управления группами по прецедентам не сталкивалась, по крайней мере на настоящем этапе, с необходимостью учета указанных недостатков [11].

Фасетная классификация прецедентов

Основой фасетной классификации является деление объектов или понятий на независимые группировки — фасеты, каждый из которых характеризует объекты или понятия в одном аспекте, по одному признаку. Данный метод значительно облегчает многоаспектное отражение объектов, поскольку можно строить классы из различных сочетаний признаков [12].

Фасетная структура понятий представляется в виде пересечения ряда признаков. Для прецедентов фасетная структура представляет собой совокупность характеристик, описывающих определенную связку «ситуация — решение». Фасетная структура прецедентов позволяет более точно определить и классифицировать объекты или явления, учитывая их различные

■ ***Таблица 1.*** Классы и классификационные признаки прецедентов
■ ***Table 1.*** Classes and classification features of precedents

№	Тип класса сущностей прецедентов	Классификационный признак	Наименование классов	Количество классов	Глубина классификации
1	Объекты (агенты) прецедента	По видам объектов, образующих группу	РТК, БПЛА, БЭК, АНПА	4	2–3
2	Мощность группы	Количество объектов в группе	Группа, стая, рой	3	2
3	Спектр группы	Процент входящих в состав объектов одного вида	Гомогенная, гетерогенная	2	2
4	Стратегия управления	Тип стратегии управления	Группа с лидером, группа без лидера, самоорганизующаяся группа	3	3
5	Стратегия принятия решений в группе	Принцип принятия решений	Конкурентная, коллективная	2	2
6	Целеполагание группы	Задача – для одиночного объекта Миссия – для группы	Мониторинг (наблюдение, разведка, контроль), патрулирование (инспектирование, поиск), доставка груза, ретрансляция сигнала, атака охраняемого объекта, защита охраняемого объекта	6	3–4
7	Пассивная среда прецедента	Тип пассивного противодействия	День/сумерки/ночь/ туман/дождь/ грязь/ветер/волнение поверхности/ муть/препятствия	10	2
8	Активная среда прецедента	Направленность противодействия	Противодействие миссии посредством оказания влияния: на функции/на перемещение/на связь/ на управление/на ресурсы	5	4
9	Решения для одиночного объекта	Выбранное действие	Отложить выполнение миссии, прекратить выполнение миссии, запросить помощь функцией, запросить помощь ресурсами, оказать информационное содействие, оказать воздействие на источник опасности	6	2
10	Решения при выполнении коллективной миссии	Выбранное коллективное действие	Отложить выполнение миссии, прекратить выполнение миссии, осуществить реконфигурацию структуры группы, осуществить реконфигурацию задач в группе, оказать воздействие на источник опасности	5	2
11	Достигнутый эффект	Значение параметра эффективности (числовое или зависит от выполняемой задачи)	Оценивается лингвистическим значением нечеткой переменной «Эффект»: высокий, приемлемый, средний, удовлетворительный, неудовлетворительный	5	3

■ ***Таблица 2.*** Примеры второго уровня иерархии для выбранных классов прецедентов
■ ***Table 2.*** Examples of the second level of hierarchy for selected classes of precedents

Класс сущностей прецедентов	Классификационный признак	Наименование классов второго уровня иерархии
Объекты класса БПЛА	Полезная нагрузка	Разведчик, камикадзе, логист, связист
Рой (класс мощности группы)	Процент боевых и ложных элементов в составе роя	Ложный рой, провоцирующий рой, ударный рой
Коллективная стратегия принятия решений в группе	Направление обучения стратегии поведения	Альтруизм, прагматизм, эгоизм
Мониторинг (целеполагание группы)	Приоритетная задача	Наблюдение, разведка, контроль

аспекты и свойства. Фасетная структура благодаря своей гибкости позволяет сформировать наиболее полную классификацию прецедентов с учетом их неоднозначности [11].

Целесообразно создать так называемую фасетной формулу для прецедентов. В основу формирования фасетных формул закладывается логико-лингвистическая модель, которая может стать основой фасетной формулы.

Исходя из изложенного, типовая фасетная формула для прецедентов группового управления имеет следующий вид:

$$GP = (RL \vee TY \wedge A \vee DS \vee V \wedge EF \vee TR) \wedge$$
$$\wedge (F \wedge M \wedge ST \vee N \vee K \vee FT) \wedge (GE \wedge C \wedge PF \vee$$
$$\vee AF \vee TC) \wedge (T1 \vee T2 \vee T3) \wedge (RO \vee RG),$$

где RL — роль объекта в группе; TY — тип объекта; A — уровень автономности группы; DS — максимальная дальность действия группы; V — уязвимость группы; EF — эффективность действия группы; TR — обученность группы; F — тип группы; M — миссия группы; ST — стратегия группы; N — мощность группы; K — вид группы; FT — геометрия действий группы в пространстве; GE — географический район; C – сложность местности/пространства мониторинга; PF — пассивные факторы среды; AF — активные факторы среды; TC — тип противодействия; $T1$ — факты (фиксация происходящего); $T2$ — события как логическая связка фактов; $T3$ — ситуации как логическая связка событий (задача — решение — эффект); RO — решения для действий объектов, включенных в группу; RG — решения для действий группы.

В качестве фасетов можно использовать сущности, составляющие прецедент, например, перечисленные во втором столбце табл. 1. Количество возникающих прецедентов при выполнении различными группами различных коллективных задач может быть довольно весомым, но значения основных параметров, как правило, весьма близки, поэтому количество групп фасетов можно сокращать.

Введем два определения.

Определение 1. «Ядром» прецедента называется совокупность его признаков, относящихся к различным сущностям, незначительное изменение которых в заданных пределах существенно меняет как решение, так и эффективность решения по прецеденту. Для прецедентов наиболее значимы тип объекта (TY), роль объекта (RL), тип группы (F), мощность группы (N), миссия группы (M), геометрия действий группы в пространстве (FT), среда (GE, C, PF, AF), задача ($T3$), поставленная перед группой, и решение (RO, RG), принятое при решении задачи.

Определение 2. «Бахромой» прецедента называется совокупность его признаков, относящихся к различным сущностям, значительное изменение которых в заданных пределах несущественно меняет как решение, так и эффективность решения с использованием прецедента (понятия «ядро» и «бахрома» впервые использованы при описании идеологии построения нейронных сетей с ансамблевой организацией и ассоциативно-проективной структурой в книге «Ассоциативные нейроподобные структуры» [13]).

Приведенные определения не требуются для построения иерархической классификации. Однако для фасетной классификации они позволяют выбрать несколько вариантов типовых прецедентов, состояния группы в которых можно описать только «ядрами» признаков, не обращая особого внимания на «бахрому». К ним относятся типовые практические ситуации, связанные, прежде всего, с выполняемыми типовыми группировками миссиями. Тогда описание прецедентов методом фасетной классификации [14–18] будет иметь вид, представленный в табл. 3.

Для каждого конкретного прецедента состояние группы представляется как пересечение признаков классов, которые независимы, но не исключают друг друга.

■ ***Таблица 3.*** Описание прецедентов методом фасетной классификации
■ ***Table 3.*** Description of precedents using faceted classification

Наименование прецедента	Фасет 1 — объект	Фасет 2 — группа	Фасет 3 — задача	Фасет 4 — среда	Фасет 5 — решение	Фасет 6 — эффект	Фасет 7 — фасетная формула
«Мониторинг» [14, 15]	Оснащен техническим зрением	Одиночный БЛА	Обнаружение объектов интереса	Ночь	Переход с оптического на ИК-диапазон	Повышение вероятности обнаружения	$(TY \vee RL) \wedge (N \vee M) \wedge PF \wedge T3 \wedge RO$
«Транспорт» [16]	Грузовой дрон, обладающий высокой грузоподъемностью	Группа БПЛА	Доставка полезного груза до места назначения	Туман	Взаимный обмен информацией о целях	Высокая точность доставки грузов	$(TY \vee RL) \wedge (K \vee M) \wedge PF \wedge T3 \wedge RG$
«Атака» [17]	МикроБПЛА	Рой БЛА	Диффузная бомба	Наличие системы ПВО	Создание целенаправленных формаций	Снижение потерь боевых элементов роя	$(TY \vee RL) \wedge (K \vee M) \wedge AF \wedge T3 \wedge RG$
«Антенна» [18]	Обладают станциями радиолокации	Стая БЛА	Создание воздушной системы предупреждения	Наличие системы РЭП	Запрос на подавление системы РЭП	Повышение эффективности обнаружения угроз	$(TY \vee RL) \wedge (K \vee M) \wedge AF \wedge T3 \wedge RG$

Оценка емкости и коэффициента заполненности классификации прецедентов

Под емкостью классификации понимается максимальное число классификационных группировок позиций, которое может вместить в себя классификатор. Этот показатель позволяет оценить объем информации, который может быть закодирован в классификаторе исходя из всех возможных кодовых комбинаций с учетом принятых методов кодирования и структуры кода.

Емкость иерархической классификации прецедентов определяется наибольшим количеством уровней и ветвей классификационных группировок, допускаемых в данной системе классификаций.

Характерными особенностями иерархической системы являются:

— возможность использовать неограниченное количество признаков классификации;

— соподчиненность признаков классификации, что выражается разбиением каждой классификационной группировки, образованной по одному признаку, на множество классификационных группировок по нижестоящему (подчиненному) признаку.

В иерархической классификации емкость можно определить как количество уровней и ветвей в иерархии классов. Чем больше уровней и ветвей, тем более сложная иерархическая структура и, следовательно, более высокая емкость классификации. Таким образом, емкость иерархической классификации можно определить как количество классификационных группировок на последнем уровне иерархии, если известно число уровней иерархии и максимальное количество значений признаков классификации на каждом уровне.

Таким образом, емкость иерархической классификации прецедентов составляет не менее нескольких десятков (до 100) единиц.

Фасетная классификация представляет из себя параллельное разделение множества объектов на независимые классификационные группировки. Поэтому емкость фасетной классификации прецедентов определяется количеством классов, на которые можно разделить данные. Следовательно, емкость можно определить как величину, соответствующую количеству фасетов и значений признаков для каждого фасета.

Объект одновременно имеет классификационные признаки из различных фасетов, а классификационные группировки создаются путем задания фасетной формулы — последовательности фасетов и значений классификационных признаков выбранных фасетов, следовательно, комбинация фасетов позволяет создать полное представление об объекте.

В табл. 3 выделено шесть основных фасетов. Значения признаков фасетов различны для различных фасетов, они варьируются от нескольких единиц (группа, задача, эффект) до нескольких десятков при более глубокой декомпозиции сущностей, например объектов. Таким образом, емкость фасетной классификации прецедентов составляет до 40–50 единиц.

В зависимости от наличия информации классификатор может быть ограниченно заполнен. Коэффициент (т. е. уровень) заполненности — это фактическое число группировок в системе, деленное на ее показатель емкости [19].

В настоящий момент групповое применение интеллектуальных объектов крайне ограничено. Существуют лишь отдельные описания (см., например, [7, 20, 21]). Анализ показывает, что как для иерархического, так и для фасетного классификатора с точки зрения практики, т. е. наличия информации о прецедентах, коэффициент (уровень) заполненности применительно к решению оборонных задач крайне мал и не превышает 2–3 %. Это означает, что классификаторы прецедентов обладают более чем 95 %-й резервной емкостью. Указанное только подтверждает необходимость сбора и систематизации прецедентов группового управления в соответствующих базах знаний.

Определение группового прецедента на основе классификации прецедентов

Наш анализ позволяет сделать вывод о том, что возникновение прецедента в групповом управлении является достаточно сложным явлением. В нем оказываются задействованными различные акторы, он возникает как спонтанно, так и предсказуемо в некоторой быстроменяющейся среде, требует фиксации состояния как объектов, так и среды, анализа и поиска эффективного решения в складывающейся ситуации.

На все типовые вопросы, задаваемые при реализации сложных многофакторных и динамичных процессов [22], а именно: кто? (акторы), что? (действие), где? (область действия), как? (способ действия), когда? (временной фактор) и зачем? (целеполагание), — классификация прецедентов группового управления позволяет сформулировать определенные ответы.

Определение 3. Прецедентом группового управления называется воспринимаемое разносредовыми и разнотипными техническими средствами явление, возникающее при выполнении (*кто?*) группой объектов управления заданной миссии (*где?*) в некоторой среде, (*когда?*) в определенный момент времени, (*как?*) заключающееся в формировании совокупности условий и обстоятельств, которые (*почему?*) в соответствии с логикой действий каждой из сторон провоцируют (*что?*) наступление последовательности событий, фиксируемых в виде совокупности фактов и, как следствие, формирующих ситуацию, описываемую текущим состоянием всей группы с возникновением коллизии между стратегиями поведения акторов и среды, потенциально влияющей на эффективность выполнения миссии и приводящей к необходимости принятия решений по дальнейшим действиям, направленным на обеспечение достижения целей миссии, с учетом того, что аналогичная/схожая/подобная/идентичная/близкая ситуация уже имела место в прошлом и по ней были приняты решения, эффективность которых априори известна.

Заключение

В настоящей работе описано формирование прецедентной ситуации, разработаны классификационные признаки, предложены иерархическая и фасетная классификации прецедентов, дана оценка емкости и заполненности классификации прецедентов и в результате сформулировано определение группового прецедента.

При выполнении группой, находящейся в определенном прецедентном состоянии, коллективной миссии вследствие прецедентной коллизии сформировалась прецедентная ситуация, требующая принятия решения. В процессе выполнения нескольких категорий коллективных миссий объекты разных классов объединяются в различные виды групп, выполняют различные типы задач, фиксируют текущее состояние, факты, события, сталкиваются с коллизиями, приводящими к сложным ситуациям, принимают прогнозируемо эффективные решения и реализуют их, изменяя состояние группы. Для разработки классификационных признаков были выделены классы всех характеристик группы. В результате выделено 11 типов классов с независимыми классификационными признаками, в каждом типе несколько классов (от двух до 10).

Принципиально возможны два типа классификации — иерархическая и фасетная. В работе проанализированы оба типа классификации; исследование выявило, что фасетная классификация является более гибкой в плане добавления новых классов, но иерархическая классификация является более емкой.

Анализ показывает, что как для иерархического, так и для фасетного классификатора с точки зрения практики (т. е. наличия информации о прецедентах) коэффициент (т. е. уровень) заполненности применительно к решению оборонных задач крайне мал и не превышает 2–3 %. Это означает, что классификаторы прецедентов обладают более чем 95 %-й резервной емкостью. Указанное только подтверждает необходимость сбора и систематизации прецедентов группового управления в соответствующих базах знаний.

Литература

1. *Теория управления*: словарь системы основных понятий/под общ. ред. Д. А. Новикова. М., ЛЕНАНД, 2024. 128 с.
2. **Татарников Д. Г.** Прецедент в римском праве: основные термины и их содержание. *Известия Саратовского университета. Новая серия. Серия: Экономика. Управление. Право*, 2023, т. 23, № 1, с. 106–111. doi:10.18500/1994-2540-202323-1-106-111, EDN: CUQJKD
3. **Иванов Д. Я.** Распределение ролей в коалициях роботов при ограниченных коммуникациях на основе роевого взаимодействия. *Управление большими системами*, 2019, № 78, с. 23–45. doi:10.25728/ubs.2019.78.2, EDN: DUXNRB
4. **Бычков И. В., Давыдов А. В., Кензин М. Ю., Нагул Н. В., Толстихин А. А.** Интеллектуальное планирование стратегий и управление группой мобильных роботов в условиях неполноты информации. *Известия ЮФУ. Технические науки*, 2023, № 1, с. 170–184. doi:10.18522/2311-3103-2023-1-170-184
5. **Кулинич А. А., Карпов В. Э., Карпова И. П.** *Социальные сообщества роботов*. М., ЛЕНАНД, 2019. 352 с.
6. **Краснояров А. Ю., Аргузова М. А., Хужамурадов Ж. А., Рахимов С. Р.** «Речевое творчество» искусственного интеллекта: какие тексты пишет машина и чем они отличаются от людских. *Социальные и гуманитарные науки. Отечественная и зарубежная литература. Сер. 6, Языкознание*, 2022, № 2, с. 41–49. doi:10.31249/ling/2022.02.02
7. **Плаксин С. И.** Подход к автономному возвращению интеллектуального мобильного робота в случае потери связи с оператором. *Известия Тульского государственного университета. Технические науки*, 2023, № 3, с. 413–415. doi:10.24412/2071-6168-2023-3-413-416, EDN: AQXMQX
8. **Городецкий В. И., Пантелеев М. Г.** Сети автономных агентов реального времени в среде с противодействием: особенности и компоненты модели. *Информационные технологии в управлении: тр. 13-й мультиконференции по проблемам управления*, Санкт-Петербург, 06–08 октября 2020 г. СПб., 2020, с. 22–31.
9. **Абросимов В. К.** Права, обязанности, обязательства и ограничения автономного робота в киберфизической группе. *Правовая информатика*, 2022, № 4, с. 67–75. doi:10.21681/1994-1404-2022-4-67-75
10. **Исаев И. Д.** Многоагентные системы, алгоритм распознавания образов интеллектуальными агентами. *Вестник науки*, 2023, т. 4, № 11 (68), с. 651–659. EDN: VOQPRA. https://www.вестник-науки.рф/article/11143 (дата обращения: 02.02.2025).
11. **Басюк В. С., Илалтдинова Е. Ю., Мандрова Н. А.** Методология, подходы и принципы проектирования сообществ: постановка проблемы. *Ценности и смыслы*, 2023, № 4 (86), с. 6–27. doi:10.24412/2071-6427-2023-4-6-27, EDN: AJJSLY
12. **Максимова И. В., Шпальченко Э. П.** Лексикографическая типология терминологических словарей: когнитивный вектор развития (на материале терминологических словарей русского, английского и французского языков предметной области «Военная авиация»). *Филологические науки. Вопросы теории и практики*, 2023, т. 16, № 2, с. 615–623. doi:10.30853/phil20230081
13. **Куссуль Э. М.** Ассоциативные нейроподобные структуры. Киев, Наукова думка, 1991. 144 с.
14. **Абросимов В. К., Райков А. Н.** Агропромышленные роботы и искусственный интеллект. *Управление развитием крупномасштабных систем (MLSD'2022): тр. Пятнадцатой Международной конференции*, Москва, 26–28 сентября 2022 г. М., 2022, с. 329–335. doi:10.25728/mlsd.2022.0329
15. **Парников В. Е., Афонин В. В., Далбараев А. С.** Использование беспилотных летательных аппаратов для проведения мониторинга земель. *Международный журнал гуманитарных и естественных наук*, 2023, № 11-3 (86), с. 135–138. doi:10.24412/2500-1000-2023-11-3-135-138, EDN: QZHGVZ
16. **Матюха С. В.** Беспилотные авиационные системы в грузоперевозках. *Транспортное дело России*, 2022, № 1, с. 141–143, doi:10.52375/20728689_2022_1_141, EDN: SKTLOJ
17. **Малышев В. А., Митрофанов Д. В.** Анализ боевых возможностей беспилотных летательных аппаратов по поражению наземных целей и порядок их применения. *Воздушно-космические силы. Теория и практика*, 2024, № 29, с. 21–33. EDN: EURKWA
18. **Макарецкий Е. А., Корнеев К. Г.** Исследование возможностей построения РЛС на беспилотных летательных аппаратах. *Известия Тульского государственного университета. Технические науки*, 2022, № 4, с. 172–177. doi:10.24412/2071-6168-2022-4-172-177
19. **Абросимов В. К.** Искусственный интеллект и проблемы развития вооружения и военной техники. *Вооружение и экономика*, 2021, № 2(56), с. 5–2. EDN: BYCJCX. https://viek.ru/ (дата обращения: 05.08.2023).
20. **Бычков И. В., Давыдов А. В., Кензин М. Ю., Нагул Н. В.** Иерархическое планирование действий разнородной группы автономных мобильных роботов. *Проблемы искусственного интеллекта*, 2024, № 2, с. 4–20. doi:10.24412/2413-7383-2024-2-4-20, EDN: EUEYQM
21. **Ле В. Н., Ронжин А. Л.** Способы и технические средства позиционирования и навигации роботов в водной среде. *Известия Кабардино-Балкарского научного центра РАН*, 2023, № 6(116), с. 167–178. doi:10.35330/1991-6639-2023-6-116-167-178, EDN: NCBFQL
22. **Данилин А. В.** *Электронные государственные услуги и административные регламенты: от политической задачи к архитектуре электронного правительства*. М., Инфра-М, 2024. 336 с.

UDC 004:623-9
doi:10.31799/1684-8853-2025-2-27-36
EDN: QEFDEK

Classification of group control precedents

V. K. Abrosimov[a], Dr. Sc., Tech., Senior Researcher, orcid.org/0000-0001-6701-2389, avk787@yandex.ru
E. S. Mikhailova[a,b], Post-Graduate Student, orcid.org/ 0009-0008-0917-5524
[a]Main Research and Testing Inter-Species Center for Advanced Weapons of the Ministry of Defense of the Russian Federation, 26, bldg. 1, Oktyabrskaya St., 127018, Moscow, Russian Federation
[b]PJSC «NPO «Almaz» named after A. A. Raspletin, 57, bldg. 14, Aviamotornaya St., 111024, Moscow, Russian Federation

Introduction: In the context of the development of autonomy for unmanned vehicles, it is important to improve the methods for increasing the level of autonomy. One of such methods is the use of precedents that allow the system to make decisions in the current situation, based on past experience. **Purpose:** To develop a classification of group management precedents using the hierarchical and faceted method for the effective use of precedents, to estimate the capacity and fill factor of the classification of such precedents. **Results:** We introduce new definitions of a precedent state, precedent collision, precedent situation and precedent facts. We confirm the applicability of common classification methods – hierarchical and faceted ones – to group management precedents. We present the depth of the hierarchical classification for various classes that is sufficient for solving practical problems. We introduce two new definitions: "core" and "fringe" of a precedent, which are required to build a faceted classification. They make it possible to select several variants of typical precedents, in which the states of the group can described only by the "cores" of features, without paying special attention to the "fringe". Using the logical-linguistic model as a base we create a facet formula for precedents, taking into account the "core" and "fringe" of the precedent. Finally, we carry out an assessment of the capacity and fill factor of the precedent classification. As a result, the capacity for each classification is determined. The analysis of the features of precedents allows us to conclude that the occurrence of a precedent in group management is a fairly complex phenomenon where various aspects of the precedent are taken into account. On the basis of the classification of precedents, a definition of a group precedent is formulated. **Practical relevance:** The classification of precedents makes it possible to implement more efficient control of unmanned vehicles, as it facilitates the search for suitable precedents when it is necessary to make a decision during the mission.

Keywords — precedent, actor, group, swarm, environment, decision, situation, condition.

For citation: Abrosimov V. K., Mikhailova E. S. Classification of group control precedents. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2025, no. 2, pp. 27–36 (In Russian). doi:10.31799/1684-8853-2025-2-27-36, EDN: QEFDEK

References

1. *Teoriya upravleniya:* slovar' sistemy osnovnykh ponyatiy [Control theory: Dictionary of the system of basic concepts]. D. A. Novikov ed. Moscow, LENAND Publ., 2024. 128 p. (In Russian).
2. Tatarnikov D. G. Precedent in Roman law: Basic terms and their semantic content. *Izvestiya of Saratov University. Economics. Management. Law*, 2023, vol. 23, no. 1, pp. 106–111 (In Russian). doi:10.18500/1994-2540-2023-23-1-106-111, EDN: CUQJKD
3. Ivanov D. Distribution of roles in coalitions of robots with limited communications based on the swarm interaction. *Large-Scale Systems Control*, 2019, no. 78, pp. 23–45. doi:10.25728/ubs.2019.78.2, EDN: DUXNRB
4. Bychkov I. V., Davydov A. V., Kenzin M. Yu., Nagul N. V., Tolstikhin A. A. Intelligent strategy planning and control of a group of mobile robots under conditions of incomplete information. *Izvestiya SFedU. Engineering Sciences*, 2023, no. 1, pp. 170–184. doi:10.18522/2311-3103-2023-1-170-184
5. Kulinich A. A., Karpov V. E., Karpova I. P. *Sotsial'nyye soobshchestva robotov* [Social communities of robots]. Moscow, LENAND Publ., 2019. 352 p. (In Russian).
6. Krasnoyarov A. Yu., Arguzova M. A., Khuzhamuradov Zh. A., Rakhimov S. R. "Speech creativity" of artificial intelligence: what texts does a machine write and how do they differ from human ones. *Social and Humanitarian Sciences. Domestic and Foreign Literature. Series 6, Linguistics*, 2022, no. 1, pp. 41–49 (In Russian). doi:10.31249/ling/2022.02.02
7. Plaksin S. I. An approach to autonomic return of intelligent mobile robot in the event of communication loss with operator. *Izvestiya Tula State University. Technical Sciences*, 2023, no. 3, pp. 413–415. doi:10.24412/2071-6168-2023-3-413-416, EDN: AQXMQX
8. Gorodetsky V. I., Panteleev M. G. Networks of autonomous agents operating in an adversary environment: Basic features of the conceptual mode. *Informatsionnyye tekhnologii v upravlenii: 13 mul'tikonferentsiya po problemam upravleniya* [Proc. 13th Multi-Conference on Management Problems "Information Technologies in Management"]. Saint-Petersburg, 2020, pp. 22–31 (In Russian).
9. Abrosimov V. K. Rights, duties, obligations and restrictions of an autonomous robot in a cyber-physical group. *Legal Informatics*, 2022, no. 4, pp. 67–75. doi:10.21681/1994-1404-2022-4-67-75
10. Isaev I. D. Multi-agent systems, algorithm for paern recognition by intelligent agents. *Science Bulletin*, 2023, vol. 4, no. 11 (68), pp. 651–659. Available at: https://www.вестник-науки.рф/article/11143 (accessed 02 February 2025) (In Russian).
11. Basyuk V. S., Ilaltdinova E. Y., Mandrova N. A. Methodology, approaches and principles of community design: problem statement. *Cennosti i smysly* [Values and Meanings], 2023, no. 4 (86), pp. 6–27 (In Russian). doi:10.24412/2071-6427-2023-4-6-27, EDN: AJJSLY
12. Maksimova I. V., Shpal'chenko E. P. Lexicographic typology of terminological dictionaries: A cognitive vector of development (on the basis of terminological dictionaries of the Russian, English and French languages in the subject area "Military aviation"). *Philology. Theory & Practice*, 2023, vol. 16, no. 2, pp. 615–623 (In Russian). doi:10.30853/phil20230081
13. Kussul E. M. *Assotsiativnyye neyropodobkyye struktury* [Associative neuron-like structures]. Kiev, Naukova dumka Publ., 1991. 144 p. (In Russian).
14. Abrosimov V. K., Raikov A. N. Artificial intelligence and robots in agriculture. *Trudy 15 Mezhdunarodnoj konferencii "Upravlenie razvitiem krupnomasshtabnyh system" (MLSD'2022)* [Proc. of the Fifteenth International Conference "Management of Large-Scale Systems Development" (MLSD'2022)]. Moscow, 2022, pp. 329–335 (In Russian). doi:10.25728/mlsd.2022.0329
15. Parnikov V. E., Afonin V. V. Dalbaraev A. S. Use of unmanned aircraft for lands monitoring. *International Journal of Humanities and Natural Sciences*, 2023, vol. 11-3 (86), pp. 135–138 (In Russian). doi:10.24412/2500-1000-2023-11-3-135-138, EDN: QZHGVZ
16. Matyukha S. Unmanned aerial systems in cargo transportation. *Transport Business in Russia*, 2022, no. 1, pp. 141–143 (In Russian). doi:10.52375/20728689_2022_1_141, EDN: SKTLOJ
17. Malyshev V. A., Mitrofanov D. V. Analysis of the combat capabilities of unmanned aerial vehicles to defeat ground targets and the procedure their applications. *Vozdushno-kosmicheskie sily. Teoriya i praktika*, 2024, no. 29, pp. 21–33 (In Russian). EDN: EURKWA
18. Makaretsky E. A., Korneev K. G. Investigation of the possibilities of building radar on unmanned aerial vehicles. *Izves-*

tiya Tula State University. Technical Sciences, 2022, vol. 4, pp. 172–177 (In Russian). doi:10.24412/2071-6168-2022-4-172-177
19. Abrosimov V. K. Artificial intelligence and problems of development of weapons and military equipment. *Vooruzheniye i ekonomika* [Armament and Economics], 2021, no. 2(56), pp. 5–21. EDN: BYCJCX. Available at: https://viek.ru/ (accessed 5 August 2023) (In Russian).
20. Bychkov I. V., Davydov A. V., Kenzin M. Yu., Nagul N. V. Hierarchical planning of actions of a heterogeneous group of autonomous mobile robots. *Problems of Artificial Intelligence*, 2024, no. 2 (33), pp. 4–20 (In Russian). doi:10.24412/2413-7383-2024-2-4-20, EDN: EUEYQM
21. Le V. N., Ronzhin A. L. Methods and technical means of positioning and navigation of robots in the aquatic environment. *News of the Kabardino-Balkarian Scientific Center of RAS*, 2023, no. 6(116), pp. 167–178 (In Russian). doi:10.35330/1991-6639-2023-6-116-167-178, EDN: NCBFQL
22. Danilin A. V. *Elektronnye gosudarstvennye uslugi i administrativnye reglamenty: ot politicheskoj zadachi k arhitekture elektronnogo pravitel'stva* [Electronic government services and administrative regulations: From a political task to the architecture of electronic government]. Moscow, Infra-M Publ., 2024. 336 p. (In Russian).

УДК 004.056.53
doi:10.31799/1684-8853-2025-2-37-49
EDN: PVFXDP

Научные статьи
Articles

Метод активной защиты объектов критической информационной инфраструктуры от кибератак на основе прерывания процесса воздействия нарушителя

В. А. Липатников[a], *доктор техн. наук, профессор, orcid.org/0000-0002-3736-4743, lipatnikovanl@mail.ru*
А. А. Шевченко[a], *канд. техн. наук, старший научный сотрудник, orcid.org/0000-0001-9113-1089*
К. В. Мелехов[a], *адъюнкт, orcid.org/0009-0007-3474-412X*
В. А. Задбоев[a], *младший научный сотрудник, orcid.org/0009-0003-9362-1307*
[a]*Военная академия связи им. Маршала Советского Союза С. М. Буденного, Тихорецкий пр., 3, Санкт-Петербург, 194064, РФ*

Введение: *развитие IT-технологий и особенности динамического взаимодействия сторон конфликта объектов критической информационной инфраструктуры и нарушителей приводит к появлению новых кибератак.* ***Цель:*** *разработать новый подход к мониторингу, анализу и прерыванию цепочки атаки еще до достижения цели вторжения на ранних этапах атак с учетом результатов моделирования процессов противоборства.* ***Результаты:*** *структурирован процесс реализации кибератаки, включающий в себя основные фазы: анализ и внедрение, активное воздействие и завершение с выводом данных. Учтены особенности современных методов многоэтапных атак и используемых программ нарушителем. Разработаны временная модель многоэтапной атаки и граф состояний действий нарушителя, что позволило рассчитать вероятностно-временные характеристики успешного воздействия нарушителя на сеть. Предложен алгоритм прерывания кибератаки на этапе сканирования сети, предполагающий выявление попыток сканирования сети, сбор и обработку информации о нарушителе, а также формирование ответных мер для прерывания атаки и автоматизированный контроль эффективности системы с возможностью корректировки мер защиты. В результате смоделирован процесс реализации активной защиты информационно-вычислительной сети от кибератаки, продемонстрировавший эффективность предложенного метода защиты.* ***Практическая значимость:*** *применение метода прерывания цикла кибератаки на объекты критической информационной инфраструктуры позволит повысить оперативность пресечения воздействий на ранних стадиях проникновения.*

Ключевые слова *— информационно-вычислительная сеть, кибератака, моделирование угрозы, прерывание атаки.*

Для цитирования: Липатников В. А., Шевченко А. А., Мелехов К. В., Задбоев В. А. Метод активной защиты объектов критической информационной инфраструктуры от кибератак на основе прерывания процесса воздействия нарушителя. *Информационно-управляющие системы*, 2025, № 2, с. 37–49. doi:10.31799/1684-8853-2025-2-37-49, EDN: PVFXDP
For citation: Lipatnikov V. A., Shevchenko A. A., Melekhov K. V., Zadboev V. A. Active protection method against cyberattacks for the objects of critical information infrastructure based on the interruption of the process of an intruder's impact. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2025, no. 2, pp. 37–49 (In Russian). doi:10.31799/1684-8853-2025-2-37-49, EDN: PVFXDP

Введение

В современных условиях геополитической нестабильности обеспечение безопасности объектов критической информационной инфраструктуры (КИИ), а именно информационных систем и в частности информационно-вычислительных сетей (ИВС), является актуальным направлением [1], которое включает в себя развитие методик повышения защищенности ИВС от кибератак (КА). Развитие IT-технологий и постоянный рост количества пользователей и аппаратных мощностей информационных систем приводят к неконтролируемому появлению новых уязвимостей в них, которые позволяют нарушителю получать доступ к ИВС КИИ, сканировать ресурсы сети, повышать права доступа в атакуемом сегменте сети, внедрять скрипты, удаленно подключаться к оборудованию ИВС и выключать его, искать конфиденциальную информацию, блокировать учетные записи пользователей [2, 3].

Воздействие КА приводит к повышению в ИВС нелегитимной активности [4], которая влечет за собой изменение пропускной способности телекоммуникационного оборудования, искажению сведений в хранилищах данных или перехвату конфиденциальной информации.

В настоящее время для обеспечения защищенности объектов КИИ, например ИВС государственного учреждения, используется государственная система обнаружения, предупреждения и ликвидации последствий компьютерных атак (ГосСОПКА). Анализ применения ГосСОПКА [5] показал, что в ней развернуты и действуют подсистемы обнаружения, предупреждения и ликвидации последствий КА. Основное внимание по функционированию уделено подсистемам обнаружения и ликвидации

последствий КА, в то время как к подсистеме предупреждения КА, согласно нормативно-правовым документам [5], предъявляются только требования к ее созданию и реализации, но не рассматривается вопрос функционирования. Ввиду этого необходима разработка новых методов предупреждения КА.

В работах [6, 7] представлены способы, в которых делается акцент на управление информационной безопасностью (ИБ) инфраструктуры на основе выявления уязвимостей в процессе функционирования. В данных источниках управление ИБ является реактивным, так как не подразумевается прогнозирование развития атаки на сеть. Вновь разрабатываемые способы и методики управления ИБ должны быть направлены на анализ не только динамики действий нарушителя, но и содержания блоков данных протоколов по этапам атаки.

Наряду с этим одним из требований к управлению ИБ ИВС является реализация способов контроля защищенности в режиме времени, близком к реальному. Отсюда вытекает противоречие между интенсивно развивающимися способами и их реализующими возможностями воздействия на ИВС с одной стороны и применяемыми методами управления ИБ с другой стороны.

В отличие от традиционных методов, которые чаще всего основаны на обнаружении и реагировании на угрозы, активная защита предполагает меры по прерыванию цепочки атаки еще до достижения цели вторжения. Необходим анализ содержания цикла КА в ИВС, т. е. последовательности действий нарушителя.

В статье [8] исследуется цикл КА при наблюдении за параметрами трафика внутри ИВС. Представлено четкое позиционирование и места хранения результатов сканирования, с которыми придется работать администратору ИБ. Однако в предложенной системе отсутствует прогнозирование возможных нарушений ИБ или какие-нибудь противодействия уже проведенной атаке на ИВС.

Целью статьи является представление нового метода мониторинга, анализа и прерывания цепочки атаки еще до достижения цели вторжения с учетом результатов моделирования процесса реализации КА.

Моделирование этапов реализации кибератаки на информационно-вычислительную сеть

Прежде чем переходить к разработке способов противодействия нарушителю, необходимо подробно изучить процесс воздействия на ИВС. Такой подход позволяет не только фиксировать потенциальные угрозы, но и структурировать данные, которые будут необходимы администратору ИБ для оперативного реагирования и принятия мер.

Перед разработкой временной диаграммы (рис. 1) были проанализированы документ ФСТЭК [9] и зарубежная матрица MITRE ATT & CK, описывающие современные тактики и техники, которые применяют нарушители в ИВС. Так как целью исследования является пресечение внедрения нарушителя в ИВС на начальных этапах атаки, то при создании временной диаграммы был сделан акцент на начальные этапы реализации КА с дальнейшим обобщением последующих этапов.

Представлен обобщенный цикл КА при реализации угрозы, который состоит из трех фаз.

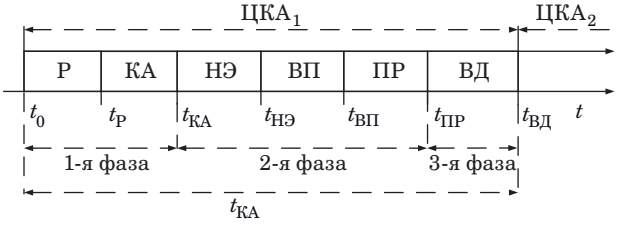
В первой фазе «Анализ и внедрение» нарушитель первоначально проводит разведку, собирая данные об объектах КИИ и изучая систему для выявления слабых мест в ее структуре. Наличие этого этапа зависит от типа атаки и не всегда является обязательным, как, например, для применения SQL-инъекции на публичный веб-ресурс.

После сбора данных осуществляется этап «КА», в котором нарушитель выбирает вид атаки и принимает решение по воздействию на ИВС на основе доступных ему методов в виде вирусов и эксплойтов с целью получить доступ к ИВС для дальнейших действий.

Во второй фазе «Активное воздействие» нарушитель загружает выбранные на предыдущем этапе программы или код в зараженную систему.

После внедрения программы или кода нарушитель начинает их выполнение на зараженной системе, включая сбор информации, установку дополнительного вредоносного ПО, запуск процессов и т. д.

С помощью реализованных ранее программ повышается доступ к более значимым данным



■ ***Рис. 1.*** Временная диаграмма цикла реализации КА: *ЦКА* – цикл кибератаки; *Р* – разведка; *НЭ* – начальная эксплуатация; *ВП* – выполнение программ; *ПР* – повышение ранга; *ВД* – вывод данных

■ ***Fig. 1.*** Time chart of the cyberattack implementation cycle: *ЦКА* – the cyberattack cycle; *Р* – intelligence; *НЭ* – initial exploitation; *ВП* – program execution; *ПР* – rank increase; *ВД* – data output

или выполняются более разрушительные действия в системе.

В третьей фазе «Вывод данных/вредоносные действия» нарушитель завершает свои действия путем кражи или уничтожения данных [10].

Рассмотрим подробнее этапы реализации представленной КА.

При необходимости разведки нарушитель проводит сбор данных с помощью сканирования ИВС (рис. 2).

Слева на рис. 2 показана стандартная передача пакетов при общении сервера с легитимным пользователем, перед началом работы происходит трехстороннее рукопожатие (TCP handshake), в котором пользователь и сервер синхронизируются путем подтверждения (SYN + ACK) пакетов, после чего начинается свободная передача файлов.

Нарушитель при данном методе синхронизации с сервером использует уязвимость, заключающуюся в том, что он отправляет большое количество рукопожатий, предварительно добавив в них точечные запросы о системе с принудительным прерыванием синхронизации, тем самым получая возможность продолжать посылать такие запросы.

Нарушитель при сканировании ИВС способен получить сведения о сервере, обнаружить активные устройства в сети и их IP- и MAC-адреса, определить статусы портов (открыт, закрыт, фильтруется и т. д.) и протоколов (какие используются и уровни их защиты), выявить неактивные или слабозащищенные ресурсы, сформировать карту сети, а при усложненных запросах или недостаточной защите ИВС — получить сведения об уязвимостях сервера, на котором обрабатываются данные, или обнаружить закрытые разделы сайта [11, 12].

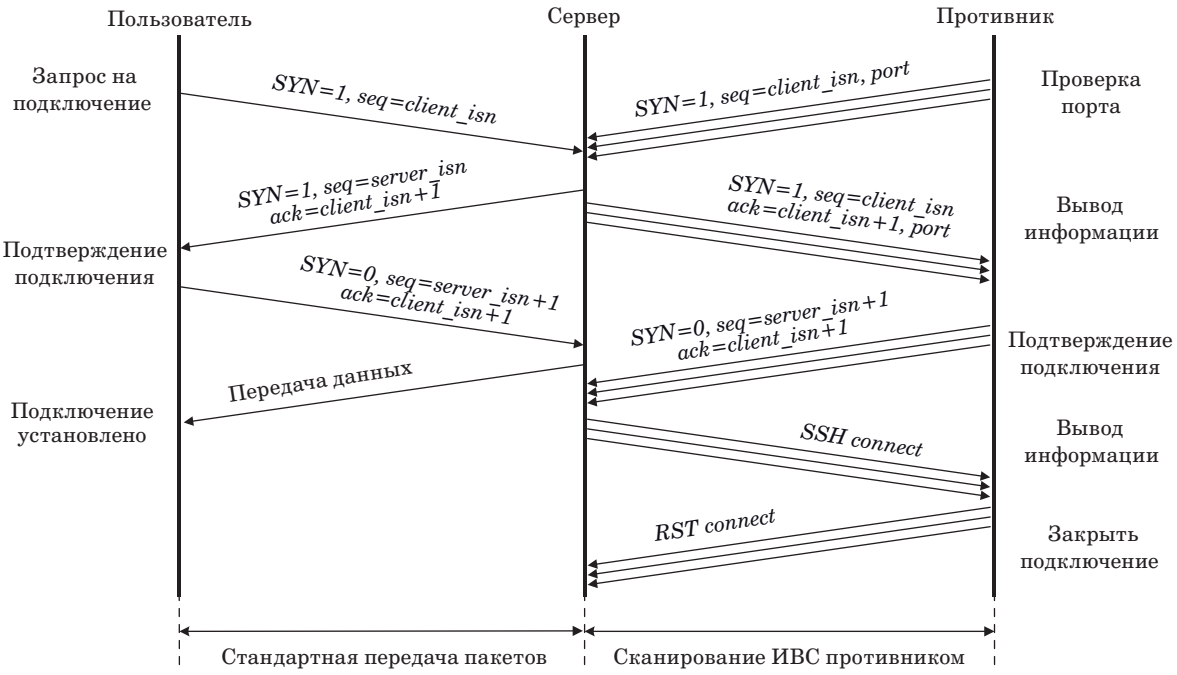
Таким образом нарушитель может получить ценный набор информации о цели атаки, после изучения которой проводит планирование атаки. Например, полезными сведениями могут послужить открытые уязвимости (vulners), сведения о структуре сервера (http-enum) и статусы портов (открыт, закрыт, фильтруется и т. д.) [13].

Ввиду вышесказанного готовое решение по воздействию на ИВС будет выглядеть следующим образом: «Загрузка вредоносного ПО на /admin/index.php через порт 22/tcp с использованием уязвимости CVE-2023-51385».

Загрузка вредоносного ПО представляет собой атаки типа Troyan, Ransomware, Keylogger, Spyware и т. д., которые заражают компьютеры или целые ИВС, в целях кражи данных или получения прав доступа к управлению системой. В случае успешной атаки нарушитель приступает к активным действиям для выполнения своей задачи.

Атаки на ИВС могут быть разнообразными. Возникает необходимость учитывать все сигнатуры, по которым возможно определить нарушителя. Перечень сигнатур представлен на сайте ФСТЭК [9], наиболее известными являются:

1. Угроза утечки информации, удаления информационных ресурсов, ненадлежащего (нецелевого) использования.



■ ***Рис. 2.*** Способ сканирования ИВС нарушителем

■ ***Fig. 2.*** Method of scanning the information and computer network by an intruder

При эксплуатации уязвимостей в ПО нарушитель внедряет в ИВС эксплойты, используя уязвимости в операционных системах, веб-приложениях или другом ПО для получения несанкционированного доступа к системе, выполнения кода или кражи данных. Для реализации своих эксплойтов нарушитель может воспользоваться различными инструментами:

— Metasploit Framework — мощный набор инструментов для разработки, тестирования и использования эксплойтов;

— ExploitDB — онлайн-база данных, содержащая большое количество готовых эксплойтов для различных уязвимостей;

— Immunity Canvas — коммерческое ПО для разработки и тестирования собственных эксплойтов;

— Burp Suite — набор инструментов, используемый для тестирования безопасности веб-приложений, включающий модуль для эксплуатации уязвимостей [14].

2. Угрозы несанкционированного доступа, модификации (искажения), подмены.

Атака типа MiTM (человек посередине) работает следующим образом: нарушитель устанавливает маршрутизатор между двумя устройствами, которые обмениваются данными. Для анализа трафика нарушитель использует программу, например WireShark. Нарушитель создает точку доступа с идентичным названием и паролем, что и у исходной сети, перенаправляет через нее трафик клиентов и запускает WireShark для его прослушивания. Такая атака позволяет ему перехватывать, изменять или внедрять данные в передаваемый трафик.

После достижения своих целей нарушитель забирает всю необходимую информацию, параллельно скрывая или удаляя следы своих действий. Это происходит путем:

— удаления и модификации системных логов в серверах и SIEM;

— удаления временных файлов в хосте, находящихся в папках /temp и оперативной памяти;

— удаления вредоносного ПО или его скрытия для дальнейшего использования;

— создания и маскировки бэкдоров для упрощенного доступа к сети;

— использования VPN и прокси для скрытия точек подключения;

— маскировки вредоносного сетевого трафика под обычные протоколы;

— шифрования и удаления файлов системы.

3. Угроза отказа в обслуживании, нарушения функционирования (работоспособности).

DDoS-атаки (Distributed Denial of Service — атаки, направленные на распределенный отказ в обслуживании) представляют собой один из видов, целью которых является нарушение работоспособности или затруднение доступа к системе для обычных пользователей. При реализации данной атаки нарушитель использует ботнеты (сети зараженных компьютеров) для одновременного направления огромного объема запросов на серверы или сети, что приводит к их перегрузке и временному отказу в обслуживании для пользователей данной сети.

APDoS-атака (Advanced Persistent Denial of Service — усовершенствованная постоянная атака отказа в обслуживании) представляет собой разновидность КА, связанной с высокой степенью угрозы, требующей особых методов защиты от DDoS-атак.

В подобных сценариях нарушитель может намеренно переключаться между разными целями, чтобы отвлечь внимание и обойти меры противодействия, сосредотачивая основную мощь атаки на одной выбранной жертве. Такой нарушитель, обладая постоянным доступом к мощным сетевым ресурсам, способен организовать длительную атаку, сопровождающуюся созданием огромного объема DDoS-трафика.

Особенности APDoS-атак:

— использование OSINT (Open Source Intelligence — разведка по открытым источникам) для тщательного анализа перед атакой, а также сканирование инфраструктуры, позволяющее скрываться в течение длительного времени;

— тактическое исполнение атак по нескольким целям для отвлечения внимания от основной цели атаки;

— большая вычислительная мощность (использование значительных сетевых и аппаратных ресурсов для организации атак);

— одновременные многопоточные атаки на нескольких уровнях OSI (от 3-го до 7-го);

— долговременная активность вышеперечисленных действий [15].

Для реализации атаки, как говорилось выше, первоначально нарушителю необходимо заразить большое количество компьютеров для создания своего ботнета, самым известным ПО является Mirai, способное, помимо компьютеров, заражать устройства интернета вещей (IoT), такие как IP-камеры, маршрутизаторы и умные устройства.

Далее после выбора цели атаки нарушитель удаленно активирует свой ботнет с использованием таких программ, как DDoS-as-a-Service (использование сторонних услуг для DDoS), NTP Reflection Attack Tools и Memcached DDoS Tools для реализации TCP/UDP Flood, HTTP Flood и DNS-amplification [16].

4. Угроза получения информационных ресурсов из недоверенного или скомпрометированного источника.

Использование недостатков, связанных с некорректной настройкой сетевого доступа, — это тип атаки, при котором нарушитель эксплуатирует ошибки в конфигурации сетевых служб, маршрутизации или прав доступа для получения доступа к чувствительным данным или системам. Такие уязвимости могут возникать из-за неправильной настройки брандмауэров (Firewall Misconfiguration Attack), маршрутизаторов (Default Credential Exploitation), серверов (Network Sprawl Attack) и других сетевых компонентов.

5. Угроза распространения противоправной информации.

Целенаправленная рассылка через рассылку рекламных писем. Атака реализуется после получения доступа к ресурсам организации, например для рассылки электронных писем от лица доверенной организации, с целью распространить вредоносное ПО.

6. Угроза несанкционированного массового сбора информации.

SQL-инъекции — это атаки, направленные на взлом веб-приложений путем внедрения злонамеренного SQL-кода в поле ввода на вебстранице, что может привести к несанкционированному доступу к базе данных и краже или модификации данных. Ниже описан пример реализации SQL-инъекции с использованием программы Havij.

Нарушитель использует Havij для сканирования веб-сайтов на наличие уязвимых точек ввода, в которые вводится злонамеренный SQL-код, например ' *OR* 1=1 --, который заставляет вывести все записи таблицы, из которых с помощью «парсинга» нарушитель добывает ценные сведения [17].

Математическая модель процесса реализации кибератаки на информационно-вычислительную сеть

Проведем моделирование воздействия нарушителя на ИВС, так как это позволит детально изучить процесс реализации угроз, оценить вероятность их успешного выполнения и на основе полученных результатов выявить наиболее уязвимые этапы цепочки КА и разработать эффективные способы противодействия.

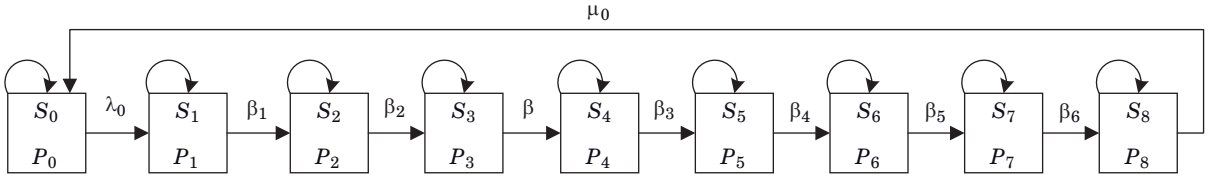
В качестве примера воздействия нарушителя разработана модель реализации угрозы типа DDoS-атаки с реализацией проникновения в ИВС с использованием эксплуатации уязвимостей.

Для исследования процесса реализации КА, а также оценки вероятностно-временных характеристик КА в рамках настоящего исследования на основе теории марковских процессов разработана математическая модель в виде графа состояний (рис. 3, где P_i — вероятность нахождения нарушителя в i-м состоянии), параметры процесса КА представлены в таблице. Данная модель позволит выполнить расчет вероятностных характеристик успешного проведения атаки.

Взяв за основу полученный выше граф состояний, составим систему дифференциальных уравнений

$$\begin{cases} \frac{dP_0(t)}{dt} = \mu_0 P_8(t) - \lambda_0 P_0(t); \\ \frac{dP_1(t)}{dt} = \lambda_0 P_0(t) - \beta_1 P_1(t); \\ \frac{dP_2(t)}{dt} = \beta_1 P_1(t) - \beta_2 P_2(t); \\ \frac{dP_3(t)}{dt} = \beta_2 P_2(t) - \beta P_3(t); \\ \frac{dP_4(t)}{dt} = \beta P_3(t) - \beta_3 P_4(t); \\ \frac{dP_5(t)}{dt} = \beta_3 P_4(t) - \beta_4 P_5(t); \\ \frac{dP_6(t)}{dt} = \beta_4 P_5(t) - \beta_5 P_6(t); \\ \frac{dP_7(t)}{dt} = \beta_5 P_6(t) - \beta_6 P_7(t); \\ \frac{dP_8(t)}{dt} = \beta_6 P_7(t) - \mu_0 P_8(t), \end{cases} \quad (1)$$

решение которой представлено выражением



■ ***Рис. 3.*** Граф состояний действий нарушителя при реализации угрозы

■ ***Fig. 3.*** Graph of the states of the attacker's actions during the implementation of the threat

$$
\begin{cases}
P_0 = \dfrac{1}{\left(1+\lambda_0\left(\dfrac{1}{\beta_1}+\dfrac{1}{\beta_2}+\dfrac{1}{\beta}+\dfrac{1}{\beta_3}+\dfrac{1}{\beta_4}+\dfrac{1}{\beta_5}+\dfrac{1}{\beta_6}+\dfrac{1}{\mu_0}\right)\right)}; \\
P_1 = \dfrac{\lambda_0}{\beta_1}\dfrac{1}{\left(1+\lambda_0\left(\dfrac{1}{\beta_1}+\dfrac{1}{\beta_2}+\dfrac{1}{\beta}+\dfrac{1}{\beta_3}+\dfrac{1}{\beta_4}+\dfrac{1}{\beta_5}+\dfrac{1}{\beta_6}+\dfrac{1}{\mu_0}\right)\right)}; \\
P_2 = \dfrac{\lambda_0}{\beta_2}\dfrac{1}{\left(1+\lambda_0\left(\dfrac{1}{\beta_1}+\dfrac{1}{\beta_2}+\dfrac{1}{\beta}+\dfrac{1}{\beta_3}+\dfrac{1}{\beta_4}+\dfrac{1}{\beta_5}+\dfrac{1}{\beta_6}+\dfrac{1}{\mu_0}\right)\right)}; \\
P_3 = \dfrac{\lambda_0}{\beta}\dfrac{1}{\left(1+\lambda_0\left(\dfrac{1}{\beta_1}+\dfrac{1}{\beta_2}+\dfrac{1}{\beta}+\dfrac{1}{\beta_3}+\dfrac{1}{\beta_4}+\dfrac{1}{\beta_5}+\dfrac{1}{\beta_6}+\dfrac{1}{\mu_0}\right)\right)}; \\
P_4 = \dfrac{\lambda_0}{\beta_3}\dfrac{1}{\left(1+\lambda_0\left(\dfrac{1}{\beta_1}+\dfrac{1}{\beta_2}+\dfrac{1}{\beta}+\dfrac{1}{\beta_3}+\dfrac{1}{\beta_4}+\dfrac{1}{\beta_5}+\dfrac{1}{\beta_6}+\dfrac{1}{\mu_0}\right)\right)}; \\
P_5 = \dfrac{\lambda_0}{\beta_4}\dfrac{1}{\left(1+\lambda_0\left(\dfrac{1}{\beta_1}+\dfrac{1}{\beta_2}+\dfrac{1}{\beta}+\dfrac{1}{\beta_3}+\dfrac{1}{\beta_4}+\dfrac{1}{\beta_5}+\dfrac{1}{\beta_6}+\dfrac{1}{\mu_0}\right)\right)}; \\
P_6 = \dfrac{\lambda_0}{\beta_5}\dfrac{1}{\left(1+\lambda_0\left(\dfrac{1}{\beta_1}+\dfrac{1}{\beta_2}+\dfrac{1}{\beta}+\dfrac{1}{\beta_3}+\dfrac{1}{\beta_4}+\dfrac{1}{\beta_5}+\dfrac{1}{\beta_6}+\dfrac{1}{\mu_0}\right)\right)}; \\
P_7 = \dfrac{\lambda_0}{\beta_6}\dfrac{1}{\left(1+\lambda_0\left(\dfrac{1}{\beta_1}+\dfrac{1}{\beta_2}+\dfrac{1}{\beta}+\dfrac{1}{\beta_3}+\dfrac{1}{\beta_4}+\dfrac{1}{\beta_5}+\dfrac{1}{\beta_6}+\dfrac{1}{\mu_0}\right)\right)}; \\
P_8 = \dfrac{\lambda_0}{\mu_0}\dfrac{1}{\left(1+\lambda_0\left(\dfrac{1}{\beta_1}+\dfrac{1}{\beta_2}+\dfrac{1}{\beta}+\dfrac{1}{\beta_3}+\dfrac{1}{\beta_4}+\dfrac{1}{\beta_5}+\dfrac{1}{\beta_6}+\dfrac{1}{\mu_0}\right)\right)}.
\end{cases}
\tag{2}
$$

С помощью исходных данных из таблицы, являющихся экспертными оценками, были проведены контрольные решения оценки вероятности успешного воздействия атаки от времени идентификации уязвимости при различном времени, затрачиваемом на выбор атаки.

Вероятность успешного воздействия $P_в = P_0 + P_8$, так как состояние 8 является реализацией угрозы «Хищение данных», а в состоянии 0 нарушитель выбирает другую цель для атаки, поэтому возвращается к нему. В связи с этим вероятность успешного воздействия будет вычисляться с помощью выражения

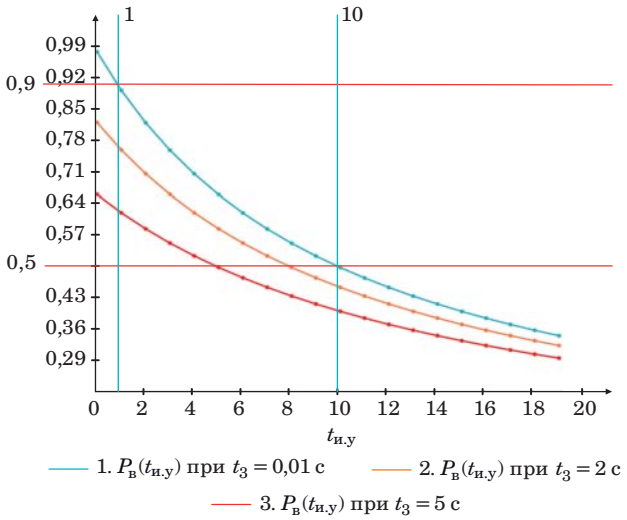
$$
P_в = \frac{\mu_0\lambda_0}{\mu_0\left(1+\lambda_0\left(\dfrac{1}{\beta_1}+\dfrac{1}{\beta_2}+\dfrac{1}{\beta}+\dfrac{1}{\beta_3}+\dfrac{1}{\beta_4}+\dfrac{1}{\beta_5}+\dfrac{1}{\beta_6}+\dfrac{1}{\mu_0}\right)\right)}. \tag{3}
$$

С применением (3) построим график зависимости вероятности успешного воздействия атаки от времени идентификации уязвимости при различном времени, затрачиваемом на выбор атаки (рис. 4). Из графика можно сделать вывод, что нарушитель должен действовать максимально быстро в ИВС при реализации всех атак, в которых присутствует этап идентификации уязвимости, поскольку любые задержки значительно увеличивают риск его обнаружения, поэтому необходимо обладать средствами, которые превосходят его по скорости.

■ Параметры процесса реализации угрозы типа DDoS-атаки с реализацией эксплуатации уязвимостей

■ Parameters of the process of implementing a threat of the type DDoS attack with the implementation of the attack Exploitation of vulnerabilities

Состояние S_n	Описание состояния S_n	Интенсивность перехода в состояние S_m	Параметр	Описание параметра	Значения параметров, с
S_0	Исходное состояние	$\lambda_0 = \frac{1}{\bar{t}_0}$	$\bar{t}_0$	Среднее время принятия решения на КА	10
S_1	Формирование ботнета	$\beta_1 = \frac{1}{\bar{t}_1}$	$\bar{t}_1$	Среднее время формирования ботнета	0,01
S_2	Подготовка к атаке	$\beta_2 = \frac{1}{\bar{t}_2}$	$\bar{t}_2$	Среднее время подготовки к атаке	0,01
S_3	Идентификация уязвимостей	$\beta = \frac{1}{\overline{t_{\text{и.у}}}}$	$\overline{t_{\text{и.у}}}$	Среднее время идентификации уязвимостей	25
S_4	Выбор метода атаки	$\beta_3 = \frac{1}{\bar{t}_3}$	$\bar{t}_3$	Среднее время выбора метода атаки	0,01 2 5
S_5	Приведение цифрового устройства или программы в состояние готовности к использованию атаки	$\beta_4 = \frac{1}{\bar{t}_4}$	$\bar{t}_4$	Среднее время приведения цифрового устройства или программы в состояние готовности к использованию атаки	0,01
S_6	Нарушение доступности к цели атаки	$\beta_5 = \frac{1}{\bar{t}_5}$	$\bar{t}_5$	Среднее время нарушения доступности к цели атаки	0,01
S_7	Проникновение в ИВС	$\beta_6 = \frac{1}{\bar{t}_6}$	$\bar{t}_6$	Среднее время проникновения в ИВС	0,01
S_8	Хищение данных	$\mu_0 = \frac{1}{\bar{t}_7}$	$\bar{t}_7$	Среднее время, затрачиваемое на хищение данных	0,01



■ ***Рис. 4.*** Зависимость вероятности успешного воздействия атаки от времени идентификации уязвимости при различном времени, затрачиваемом на выбор атаки

■ ***Fig. 4.*** Dependence of the probability of successful impact of an attack on the time of vulnerability identification at different times spent on the selection of an attack

Метод активной защиты информационно-вычислительной сети от кибератаки нарушителя

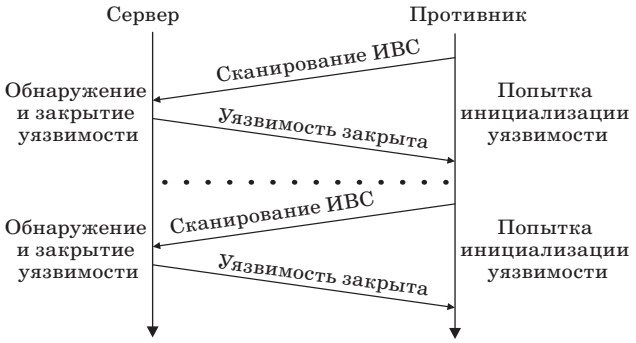
Анализ действий нарушителя [18] приводит к выводу, что самая ранняя возможность обнаружения и прерывания КА — это состояние 3, когда нарушитель пытается провести сканирование ИВС. Пример реализации прерывания цепочки КА на этапе сканирования представлен на рис. 5.

Для реализации прерывания цепочки КА предложен метод активной защиты ИВС (рис. 6) от всех типов КА, в которых нарушителю требуется предварительно проводить сканирование сети для обнаружения уязвимостей.

Суть алгоритма заключается в следующем.

1. Обнаруживается попытка сканирования нарушителем ИВС. Собирается информация о нарушителе и отправляется в центр управления ИБ [19, 20].

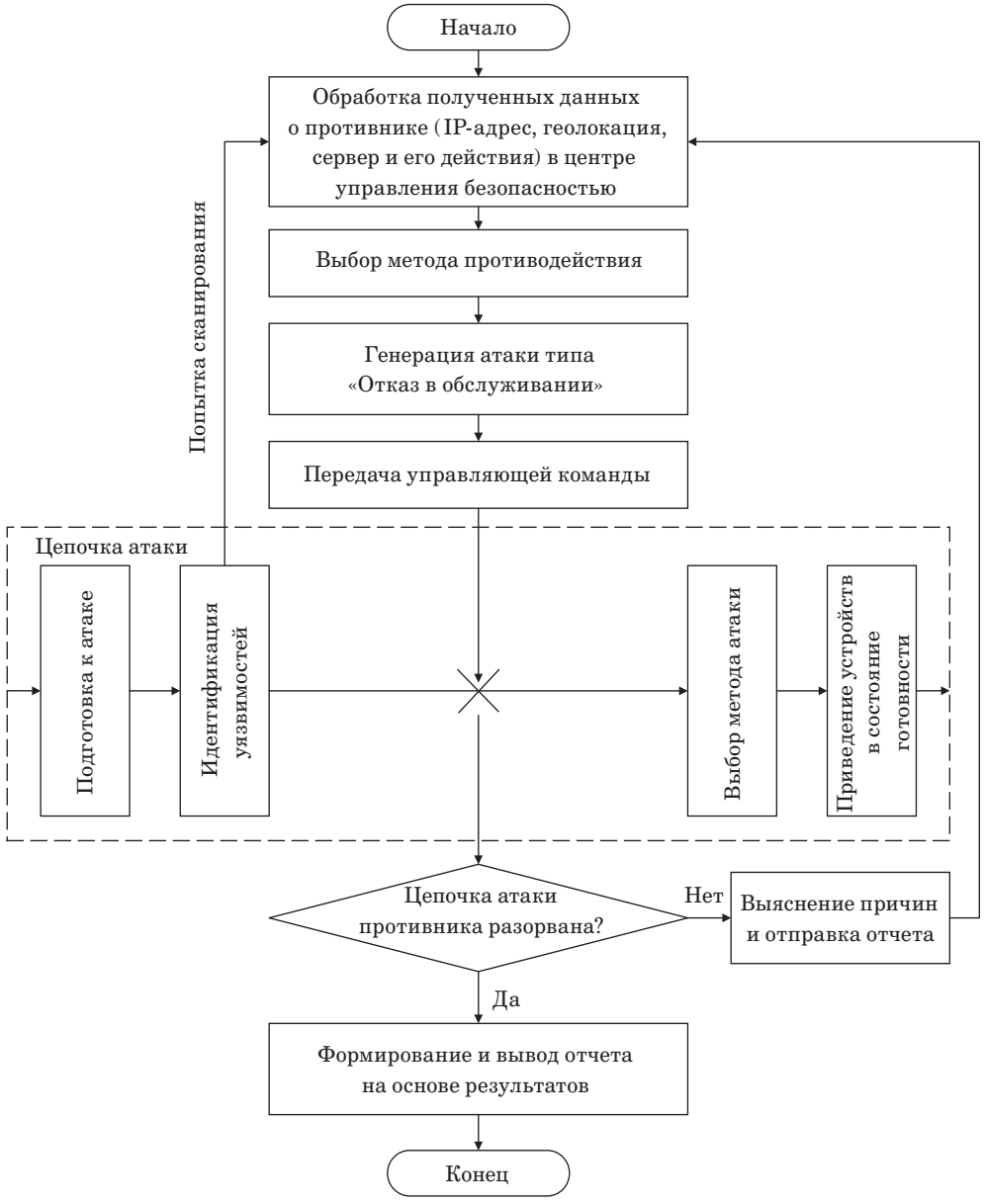
2. Полученная информация обрабатывается в центре управления ИБ. Происходит выбор и



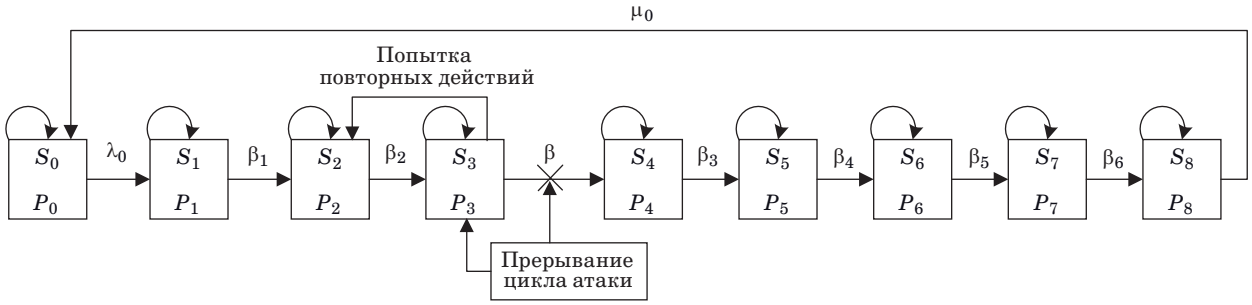
■ ***Рис. 5.*** Пример прерывания КА нарушителя
■ ***Fig. 5.*** Example of interrupting an attacker's cyberattack

передача управляющих команд на реализацию ответных мер для прерывания КА.

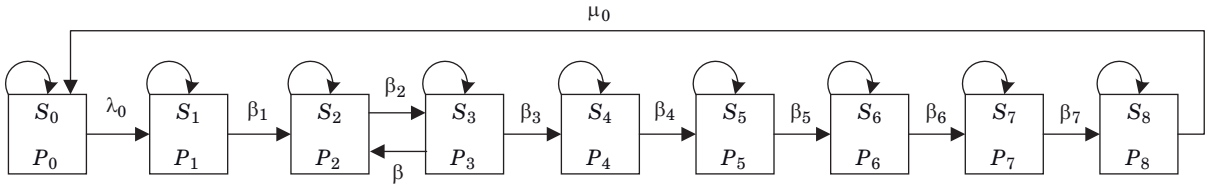
3. В конце проводится проверка, при которой выясняется, прервана КА или нет. Если атака не прервана, то анализируется неудачное прерывание атаки и отправляется отчет в центр ИБ. После принимается повторное решение об анализе сложившейся ситуации с дальнейшим выбором других ответных мер. При успешном же разрыве цепочки КА формируется отчет о результатах прерывания и отправляется администратору центра управления ИБ [21–23]. Граф прерывания КА нарушителя представлен на рис. 7.



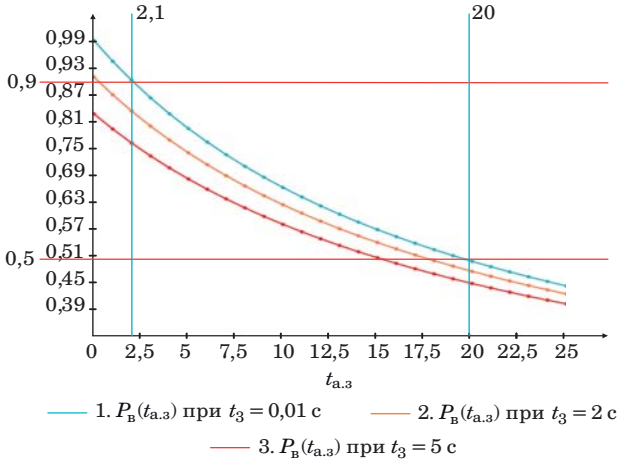
■ ***Рис. 6.*** Алгоритм прерывания КА нарушителя
■ ***Fig. 6.*** Flow diagram of the algorithm for interrupting an attacker's cyberattack



■ ***Рис. 7.*** Граф прерывания КА
■ ***Fig. 7.*** Cyberattack interrupt graph



■ ***Рис. 8.*** Граф состояний нарушителя с условием прерывания КА
■ ***Fig. 8.*** Graph of the state of the attacker with the condition of aborting the cyberattack



■ ***Рис. 9.*** Зависимость вероятности реализации нарушителем угрозы от времени активной защиты КА при различном времени, затрачиваемом на выбор атаки
■ ***Fig. 9.*** Dependence of the probability of the threat implementation by the intruder on the time of successful interruption of the active defense of a cyber attack with different times spent on choosing an attack

Если при попытке нарушителя по сканированию ИВС прервать его КА, то ему придется откатиться до состояния 2, в котором проводится подготовка новой атаки. Граф состояний нарушителя с условием прерывания КА представлен на рис. 8.

С использованием графа на рис. 8 и теории марковских случайных процессов получена зависимость вероятности реализации нарушителем угрозы от времени успешного прерывания КА

$$P_{в} = \frac{\mu_0\lambda_0}{\mu_0\left(1+\lambda_0\left(\frac{1}{\beta_1}+\frac{\beta+\beta_3}{\beta\beta_3}+\frac{1}{\beta_3}+\frac{1}{\beta_4}+\right.\right.} \\ \left.\left.+\frac{1}{\beta_5}+\frac{1}{\beta_6}+\frac{1}{\beta_7}+\frac{1}{\mu_0}\right)\right). \quad (4)$$

С учетом исходных данных, представленных в таблице, построен график зависимости вероятности реализации нарушителем угрозы от времени активной защиты КА при различном времени, затрачиваемом на выбор атаки (рис. 9).

Как видно из рис. 9, нарушителю необходимо выбрать метод КА за две секунды, чтобы повысить вероятность успешной реализации КА (кривая *1*). В случае, когда нарушитель принимает решение на реализацию КА за время более двух секунд, вероятность успешной реализации КА резко уменьшается (кривые *2* и *3*). Из вышесказанного можно сделать вывод, что система активной защиты должна реагировать на КА нарушителя в предельно короткий промежуток времени ($t_3 \le 2$ с).

Заключение

Разработана и исследована модель КА нарушителя на объект КИИ, отражающая обобщенный цикл атаки, вероятностно-временные характеристики и этапы реализации, а также предложен активный метод защиты ИВС. Работа направлена на повышение защищенности ИВС от КА.

В ходе работы достигнуты следующие результаты.

1. Структурирован процесс реализации КА, который включает в себя три основные фазы: анализ и внедрение, активное воздействие и завершение с выводом данных. Каждая фаза дополнительно разбита на этапы, такие как разведка, начальная эксплуатация, выполнение команд и повышение привилегий. Эта структура позволяет детализировать действия нарушителя и выделить ключевые точки для противодействия.

2. Проведен анализ современных типов атак, включая DDoS, SQL-инъекции, эксплуатацию уязвимостей и атаки на IoT-устройства. Проанализированы программы, которые использует нарушитель, такие как Nmap, Metasploit, Havij, Mirai и др.

3. Разработана временная диаграмма цикла атак, а также граф состояний действий нарушителя, который отражает ключевые этапы подготовки и реализации угроз, что позволяет рассчитывать вероятностно-временные характеристики успешного воздействия.

4. Разработан алгоритм прерывания КА на этапе сканирования сети, который предполагает:

— выявление попыток сканирования ИВС нарушителем;

— сбор и обработку информации о действиях нарушителя;

— формирование ответных мер для прерывания атаки, включая активные действия, направленные на затруднение последующих попыток проникновения;

— автоматизированный контроль эффективности прерывания атаки с возможностью корректировки мер защиты.

5. Проведено моделирование процесса реализации активной защиты ИВС от КА, которое продемонстрировало эффективность предложенного метода защиты. На основании моделирования можно утверждать, что система защиты должна работать в предельно короткий промежуток времени ($t \leq 2$ с), так как задержка даже в несколько секунд значительно увеличивает шансы нарушителя на успех при реализации угрозы типа DDoS-атаки с эксплуатацией уязвимостей.

Новизной предлагаемого метода является внедрение в алгоритм обеспечения защиты объектов КИИ активных мер противоборства, основанных на реализации процесса прерывания воздействия нарушителя, заключающегося в выявлении попытки сканирования ИВС нарушителем, сборе и обработки информации о нем и формировании ответных мер противодействия.

Практическая значимость заключается в том, что предложен и протестирован метод прерывания КА на этапе сканирования, который может быть внедрен в современные системы ИБ, а также использован для создания автоматизированных систем мониторинга и защиты, способных действовать в режиме реального времени. Представленные результаты обеспечивают повышение устойчивости ИВС к КА, снижая вероятность их успешной реализации.

Литература

1. **Osipov V., Kuleshov S., Zaytseva A., Levonevskiy D., Miloserdov D.** Neural network forecasting of news feeds. *Expert Systems with Applications*, 2021, vol. 169, iss. 4, Article 114521. doi:10.1016/j.eswa.2020.114521
2. **Котенко И. В., Саенко И. Б., Лаута О. С., Юрьев А. С., Запруднов М. С.** Протестное программное обеспечение: анализ и подход к обнаружению, основанный на машинном обучении. *Вопросы кибербезопасности*, 2024, № 6, с. 42–52. doi:10.21681/2311-3456-2024-6-42-52, EDN: VCJRVW
3. **Котенко И. В., Саенко И. Б., Бортникер П. В.** Методика обнаружения сетевых вторжений на основе интеграции методов вейвлет-анализа и математической статистики. *Правовая информатика*, 2024, № 4, с. 23–31. doi:10.24412/1994-1404-2024-4-23-31, EDN: TYLWLM
4. **Amma N. G. B., Selvakumar S., Velusamy R. L.** A statistical approach for detection of denial-of-service attacks in computer networks. *IEEE Transactions on Network and Service Management*, 2020, vol. 17, no. 4, pp. 2511–2522. doi:10.1109/TNSM.2020.3022799
5. Федеральный закон от 26.07.2017 № 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации». https://gossopka.ru/doc/npa/federalnye-zakony/federalnyy-zakon-ot-26072017-n-187-fz-o-bezopasnosti-kriticheskoy-informacionnoy-infrastruktury-ross-37407.html (дата обращения: 12.10.2024).
6. **Kumar A., Singh R., Sharma P.** A Comprehensive study on web application firewalls: Technologies, challenges, and future directions. *International Journal of Web Information Systems*, 2021, vol. 17, iss. 4, pp. 393–415. doi:10.1108/IJWIS-01-2021-0005
7. Пат. 2758997 C1 Российская Федерация, МПК H04L 29/02, G06F 21/50. Способ защиты информа-

ционно-вычислительной сети от вторжений, С. С. Чайковский. № 2021107965; заявл. 25.03.2021; опубл. 11.08.2021.

8. **Липатников В. А., Задбоев В. А., Мелехов К. В., Шевченко А. А.** Метод повышения защищенности информационно-телекоммуникационной сети с учетом использования средств определения геолокации нарушителя. *Труды учебных заведений связи*, 2023, № 9(4), с. 86–96. doi:10.31854/1813-324X-2023-9-4-86-96, EDN: FWQHUC
9. Методика оценки угроз безопасности информации. https://fstec.ru/dokumenty/vse-dokumenty/spetsialnye-normativnye-dokumenty/metodicheskij-dokument-ot-5-fevralya-2021-g.html (дата обращения: 05.11.2024).
10. **Задбоев В. А., Беседин М. Д., Антонов А. С.** Определение цикла атаки противника в информационно-вычислительной сети. Проблемы стандартизации, унификации и метрологии систем и средств связи специального назначения в аспекте их цифровой трансформации: науч.-техн. сб./под ред. В. А. Гаврина. СПб., ВАС, 2024, с. 47–51.
11. **Липатников В. А., Шевченко А. А., Мелехов К. В., Ткачев Д. Ф.** Методика повышения защищенности сети передачи данных объектов критической информационной инфраструктуры при многоэтапных атаках. *Информационно-управляющие системы*, 2024, № 1, с. 44–55. doi:10.31799/1684-8853-2024-1-44-55, EDN: MVWIFR
12. **Липатников В. А., Мелехов К. В., Мелихов И. А.** Особенности сетевого контроля в сети передачи данных в условиях многоэтапных атак. *Технологии. Инновации. Связь: материалы научно-практической конференции*, Санкт-Петербург, 12 апреля 2023 г. СПб., 2023, с. 171–174.
13. **Sarhan M., Layeghy S., Moustafa N., Portmann M.** *Netflow datasets for machine learning-based network intrusion detection systems*. Big Data Technologies and Applications. Springer International Publishing, 2021, pp. 117–135.
14. **Новикова Е. С., Голубев С. А.** Подход к обнаружению вторжений на основе федеративного обучения. *Международная научная конференция по проблемам управления в технических системах: материалы конференции*, Санкт-Петербург, 26–28 сентября 2023 г. СПб., 2023, с. 200–203. doi:10.1109/CTS59431.2023.10289088, EDN: XPCAUO
15. **Patil N. V., Krishna C. R., Kumar K.** Distributed frameworks for detecting distributed denial of service attacks: A comprehensive review, challenges and future directions. *Concurrency and Computation: Practice and Experience*, 2021, vol. 33, no. 10, p. e6197. doi:10.1002/cpe.6197
16. **Shukla P., Krishna C. R., Patil N. V.** Kafka-Shield: Kafka streams-based distributed detection scheme for IoT traffic-based DDoS attacks. *Security and Privacy*, 2024, vol. 7, no. 12. doi:10.1002/spy2.416
17. **Nasereddin M., ALKhamaiseh A., Qasaimeh M.** A systematic review of detection and prevention techniques of SQL injection attacks. *Information Security Journal: A Global Perspective*, 2023, vol. 32, no. 4, pp. 252–265. doi:10.1080/19393555.2021.1995537
18. **Jiang Y., Wu S., Yang H., Luo H., Chen Z., Yin S., Kaynak O.** Secure data transmission and trustworthiness judgement approaches against cyber-physical attacks in an integrated data-driven framework. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022, vol. 52, no. 12, pp. 7799–7809. doi:10.1109/TSMC.2022.3164024
19. **Gong C., Zheng Z., Wu F., Shao Y., Li B., Chen G.** To store or not? Online data selection for federated learning with limited storage. *ACM Web Conference*, 2023, pp. 3044–3055. doi:10.1145/3543507.3583426
20. **Mulay A., Gaspard B., Naidu R., Gonzalez-Toral S., Semwal T., Manish Agrawal A.** FedPerf: A practitioners' guide to performance of federated learning algorithms. *NeurIPS 2020 PMLR*, 2021, vol. 148, pp. 302–324.
21. **Saputra F. A., Salman M., Hasim J. A. N., Nadhori I. U., Ramli K.** The next-generation nids platform: Cloud-based snort nids using containers and big data. *Big Data and Cognitive Computing*, 2022, vol. 6, iss. 1, p. 19. doi:10.3390/bdcc6010019
22. **Chen J., Gao X., Deng R., He Y., Fang C., Cheng P.** Generating adversarial examples against machine learning-based intrusion detector in industrial control systems. *IEEE Transactions on Dependable and Secure Computing*, 2022, vol. 19, no. 3, pp. 1810–1825. doi:10.1109/TDSC.2020.3037500
23. **Мурашов Д. А., Ушакова В. А.** Постановка и анализ путей решения задачи синтеза программ управления и параметров информационно-вычислительной сети на основе полимодельного описания. *Авиакосмическое приборостроение*, 2022, № 8, с. 23–32. doi:10.25791/aviakosmos.8.2022.1293, EDN: UJXAWV

UDC 004.056.53
doi:10.31799/1684-8853-2025-2-37-49
EDN: PVFXDP

Active protection method against cyberattacks for the objects of critical information infrastructure based on the interruption of the process of an intruder's impact

V. A. Lipatnikov[a], Dr. Sc., Tech., Professor, orcid.org/0000-0002-3736-4743, lipatnikovanl@mail.ru
A. A. Shevchenko[a], PhD, Tech., Senior Researcher, orcid.org/0000-0001-9113-1089
K. V. Melekhov[a], Post-Graduate Student, orcid.org/0009-0007-3474-412X
V. A. Zadboev[a], Junior Researcher, orcid.org/0009-0003-9362-1307
[a]S. M. Budenny Military Academy of Communications, 3, Tikhoretskii Pr., 190064, Saint-Petersburg, Russian Federation

Introduction: The development of IT-technologies and the specifics of dynamic interaction of the parties within the conflict between critical information infrastructure objects and intruders lead to the emergence of new cyberattacks. **Purpose:** To develop a new approach to monitoring, analyzing and interrupting the attack chain even before it achieves the goal of invasion at early stages of attacks, with the results of confrontation modeling taken into account. **Results:** We structure the process of implementing a cyberattack, including the main phases: analysis and implementation, active impact and completion with data output. The features of modern methods of multi-stage attacks and the programs used by an intruder are taken into account. We develop a time model of a multi-stage attack and a state graph of an intruder's actions, which makes it possible to calculate the probabilistic and time characteristics of a successful intruder's impact on the network. We propose an algorithm for interrupting a cyberattack at the network scanning stage, which involves identifying attempts to scan the network, collecting and processing information about the intruder, as well as forming countermeasures to interrupt the attack and implementing an automated control of the system's effectiveness with the ability to adjust protection measures. As a result, we model the process of implementing active protection against cyberattacks for the information and computing network, which demonstrates the effectiveness of the proposed protection method. **Practical relevance:** The use of the method of interrupting the cyberattack cycle at critical information infrastructure objects will increase the efficiency of suppressing the impact of a cyberattack at early stages of penetration.

Keywords — information and computing network, cyberattack, threat modeling, attack interruption.

For citation: Lipatnikov V. A., Shevchenko A. A., Melekhov K. V., Zadboev V. A. Active protection method against cyberattacks for the objects of critical information infrastructure based on the interruption of the process of an intruder's impact. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2025, no. 2, pp. 37–49 (In Russian). doi:10.31799/1684-8853-2025-2-37-49, EDN: PVFXDP

References

1. Osipov V., Kuleshov S., Zaytseva A., Levonevskiy D., Miloserdov D. Neural network forecasting of news feeds. *Expert Systems with Applications*, 2021, vol. 169, iss. 4, Article 114521. doi.org/10.1016/j.eswa.2020.114521
2. Kotenko I. V., Saenko I. B., Lauta O. S., Yuryev A. S., Zaprudnov M. S. Protest software: analysis and detection approach based on machine learning. *Cybersecurity Issues*, 2024, no. 6, pp. 42–52 (In Russian). doi:10.21681/2311-3456-2024-6-42-52, EDN: VCJRVW
3. Kotenko I. V., Saenko I. B., Bortniker P. V. Methodology for detecting network intrusions based on the integration of wavelet analysis methods and mathematical statistics. *Legal Informatics*, 2024, no. 4, pp. 23–31 (In Russian). doi:10.24412/1994-1404-2024-4-23-31, EDN: TYLWLM
4. Amma N. G. B., Selvakumar S., Velusamy R. L. A Statistical approach for detection of denial-of-service attacks in computer networks. *IEEE Transactions on Network and Service Management*, 2020, vol. 17, no. 4, pp. 2511–2522. doi:10.1109/TNSM.2020.3022799
5. *Federal'nyj zakon ot 26.07.2017 N 187-FZ "O bezopasnosti kriticheskoj informacionnoj infrastruktury Rossijskoj Federacii"* [Federal Law of 26.07.2017 N 187-FZ "On the Security of Critical Information Infrastructure of the Russian Federation"]. Available at: https://gossopka.ru/doc/npa/federalnye-zakony/federalnyy-zakon-ot-26072017-n-187-fz-o-bezopasnosti-kriticheskoy-informacionnoy-infrastruktury-ross-37407.html (accessed 12 October 2024).
6. Kumar A., Singh R., Sharma P. A comprehensive study on web application firewalls: Technologies, challenges, and future directions. *International Journal of Web Information Systems*, 2021, vol. 17, iss. 4, pp. 393–415. doi:10.1108/IJWIS-01-2021-0005
7. Tchaikovsky S. S., *Sposob zashchity informacionno-vychislitel'noj seti ot vtorzhenij* [Method of protecting an information and computing network from intrusions]. Patent Russian Federation, no. 2758997, 2021.
8. Lipatnikov V. A., Zadboev V. A., Melekhov K. V., Shevchenko A. A. Method of increasing the security of information and telecommunications networks taking into account the use of means for determining the geolocation of an intruder. *Proceedings of Telecommunication Universities*, 2023, vol. 9, no. 4, pp. 86–96 (In Russian). doi:10.31854/1813-324X-2023-9-4-86-96, EDN: FWQHUC
9. *Metodika ocenki ugroz bezopasnosti informacii* [Information security threats assessment methodology]. Available at: https://fstec.ru/dokumenty/vse-dokumenty/spetsialnye-normativnye-dokumenty/metodicheskij-dokument-ot-5-fevralya-2021-g.html (accessed 5 November 2024).
10. Zadboev V. A., Besedin M. D., Antonov A. S. *Opredelenie cikla ataki protivnika v informacionno-vychislitel'noj seti*. In: *Problemy standartizacii, unifikacii i metrologii sistem i sredstv svyazi special'nogo naznacheniya v aspekte ih cifrovoj transformacii* [Determining the adversary attack cycle in the information and computing network. In: Scientific and technical collection "Problems of standardization, unification and metrology of special-purpose communication systems and means in terms of their digital transformation"]. Eds. V. A. Gavrin. Saint-Petersburg, Military Academy of Communications Publ., 2024, pp. 47–51 (In Russian).
11. Lipatnikov V. A., Shevchenko A. A., Melekhov K. V., Tkachev D. F. Methodology for increasing the security of the data transmission network of critical information infrastructure objects during multi-stage attacks. *Informacionno-upravlyayushchie sistemy* [Information and Control Systems], 2024, no. 1, pp. 44–55. doi:10.31799/1684-8853-2024-1-44-55, EDN: MVWIFR
12. Lipatnikov V. A., Melekhov K. V., Melikhov I. A. Features of network control in a data transmission network under conditions of multi-stage attacks. *Materialy nauchno-prakticheskoj konferencii "Tekhnologii. Innovacii. Svyaz'"* [Materials of the scientific and practical conference "Technologies. Innovations. Communications"]. Saint-Petersburg, 2023, pp. 171–174 (In Russian).
13. Sarhan M., Layeghy S., Moustafa N., Portmann M. *Netflow datasets for machine learning-based network intrusion detection systems*. In: *Big Data Technologies and Applications*. Springer International Publishing, 2021, pp. 117–135.
14. Novikova E. S., Golubev S. A. A federated learning approach to intrusion detection. *Materialy konferencii "Mezhdunarodnaya nauchnaya konferenciya po problemam upravleniya v tekhnicheskih sistemah"* [Conference materials "International Scientific Conference on Control Problems in Technical

Systems"]. Saint-Petersburg, 2023, pp. 200–203 (In Russian). doi:10.1109/CTS59431.2023.10289088, EDN: XPCAUO
15. Patil N. V., Krishna C. R., Kumar K. Distributed frameworks for detecting distributed denial of service attacks: A comprehensive review, challenges and future directions. *Concurrency and Computation: Practice and Experience*, 2021, vol. 33 no. 10, p. e6197. doi:10.1002/cpe.6197
16. Shukla P., Krishna C. R., Patil N. V. Kafka-Shield: Kafka streams-based distributed detection scheme for IoT traffic-based DDoS attacks. *Security and Privacy*, 2024, vol. 7, no. 12. doi:10.1002/spy2.416
17. Nasereddin M., ALKhamaiseh A., Qasaimeh M. A systematic review of detection and prevention techniques of SQL injection attacks. *Information Security Journal: A Global Perspective*, 2023, vol. 32, no. 4, pp. 252–265. doi:10.1080/19393555.2021.1995537
18. Jiang Y., Wu S., Yang H., Luo H., Chen Z., Yin S., Kaynak O. Secure data transmission and trustworthiness judgement approaches against cyber-physical attacks in an integrated data-driven framework. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022, vol. 52, no. 12, pp. 7799–7809. doi:10.1109/TSMC.2022.3164024
19. Gong C., Zheng Z., Wu F., Shao Y., Li B. Chen G. To store or not? Online data selection for federated learning with limited storage. *ACM Web Conference*, 2023, pp. 3044–3055. doi:10.1145/3543507.3583426
20. Mulay A., Gaspard B., Naidu R., Gonzalez-Toral S., Semwal T., Manish Agrawal A. FedPerf: A practitioners' guide to performance of federated learning algorithms. *NeurIPS 2020 PMLR*, 2021, vol. 148, pp. 302–324.
21. Saputra F. A., Salman M., Hasim J. A. N., Nadhori I. U., Ramli K. The next-generation nids platform: Cloud-based snort nids using containers and big data. *Big Data and Cognitive Computing*, 2022, vol. 6, iss. 1, p. 19. doi:10.3390/bdcc6010019
22. Chen J., Gao X., Deng R., He Y., Fang C., Cheng P. Generating adversarial examples against machine learning-based intrusion detector in industrial control systems. *IEEE Transactions on Dependable and Secure Computing*, 2022, vol. 19, no. 3, pp. 1810–1825. doi:10.1109/TDSC.2020.3037500
23. Murashov D. A., Ushakova V. A. Statement and analysis of ways to solve the problem of synthesizing control programs and parameters of an information and computing network based on a polymodel description. *Aerospace Instrument-Making*, 2022, no. 8, pp. 23–32 (In Russian). doi:10.25791/aviakosmos.8.2022.1293, EDN UJXAWV

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

При наличии положительной рецензии статья рассматривается редакционной коллегией. Принятая в печать статья направляется автору для согласования редакторских правок. После согласования автор представляет в редакцию окончательный вариант текста статьи.

Процедуры согласования текста статьи могут осуществляться как непосредственно в редакции, так и по e-mail (ius.spb@gmail.com).

При отклонении статьи редакция представляет автору мотивированное заключение и рецензию, при необходимости доработать статью — рецензию.

Редакция журнала напоминает, что ответственность за достоверность и точность рекламных материалов несут рекламодатели.

УДК 004.056
doi:10.31799/1684-8853-2025-2-50-67
EDN: QFQTPU

Научные статьи
Articles

Автоматизированная система анализа слабоструктурированных данных кибербразведки с использованием больших языковых моделей

А. И. Луцкович[а], *ассистент, orcid.org/0009-0000-7257-6947*
В. И. Васильев[а], *доктор техн. наук, профессор, orcid.org/0000-0002-6105-5481*
А. М. Вульфин[а], *доктор техн. наук, профессор, orcid.org/0000-0001-5857-2413, vulfin.am@ugatu.su*
А. Д. Кириллова[а], *канд. техн. наук, доцент, orcid.org/0009-0000-4164-2526*
А. Е. Сулавко[б], *доктор техн. наук, профессор, orcid.org/0000-0002-9029-8028*
[а]*Уфимский университет науки и технологий, К. Маркса ул., 12, Уфа, 450008, РФ*
[б]*Омский государственный технический университет, Мира пр., 11, Омск, 644050, РФ*

Введение: *актуальной проблемой является недостаточная эффективность и высокая трудоемкость анализа и управления слабоструктурированными и структурированными данными кибербразведки, собираемыми из различных источников.* ***Цель:*** *повышение эффективности (сокращение трудозатрат экспертов, обеспечение полноты и оперативности обновления базы знаний) анализа данных кибербразведки с помощью методов искусственного интеллекта.* ***Результаты:*** *проведенный анализ подходов к построению систем управления данными кибербразведки показал, что перспективным направлением является применение больших языковых моделей в составе вопросно-ответной системы поддержки принятия решений с механизмом расширенной дополненной генерации. Разработаны структурная схема и функциональная модель системы интеллектуального анализа данных кибербразведки на основе больших языковых моделей с применением конвейера расширенной дополненной генерации, алгоритм семантической разметки и извлечения информации из слабоструктурированных данных, исследовательский прототип (компонентная модель, микросервисная архитектура) программного обеспечения. Отличительной особенностью системы является комплексирование сведений из нескольких источников с помощью конвейера расширенной дополненной генерации. Вычислительный эксперимент на подготовленном наборе данных показал согласованность экспертных ответов с ответами, предложенными системой, на уровне 3,98 балла из пяти. Значение метрики BERTScore F1 составило 0,89, метрики «корректность ответов» (Answer Correctness) фреймворка RAGAS — 0,822.* ***Практическая значимость:*** *применение больших языковых моделей обеспечивает возможность аккумулировать знания об актуальных сценариях атак в рамках единой базы знаний. Это позволяет повысить эффективность обработки данных с целью оказать значимую поддержку специалистам по защите информации в ходе анализа актуальных угроз, уязвимостей и сценариев реализации компьютерных атак за счет снижения трудозатрат (времени сбора и обработки) и тем самым повысить защищенность корпоративных информационных систем.* ***Обсуждение:*** *дальнейшее повышение эффективности анализа данных кибербразведки возможно на основе построения гетерогенных «комитетов экспертов» из больших языковых моделей и мультиагентных систем.*

Ключевые слова *— информационная безопасность, кибербразведка, анализ уязвимостей, индикаторы компрометации, большие языковые модели, дополнительный расширенный поиск информации.*

Для цитирования: Луцкович А. И., Васильев В. И., Вульфин А. М., Кириллова А. Д., Сулавко А. Е. Автоматизированная система анализа слабоструктурированных данных кибербразведки с использованием больших языковых моделей. *Информационно-управляющие системы*, 2025, № 2, с. 50–67. doi:10.31799/1684-8853-2025-2-50-67, EDN: QFQTPU
For citation: Lutskovich A. I., Vasilyev V. I., Vulfin A. M., Kirillova A. D., Sulavko A. E. Automated system for analyzing weakly structured threat intelligence data using large language models. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2025, no. 2, pp. 50–67 (In Russian). doi:10.31799/1684-8853-2025-2-50-67, EDN: QFQTPU

Введение

Одним из ключевых трендов в обеспечении информационной безопасности (ИБ) корпоративных информационных систем (ИС) является все более широкое использование автоматизированных систем мониторинга событий и инцидентов ИБ, осуществляющих своевременное обнаружение аномалий в работе ИС, оповещение и предупреждение специалистов о возможных компьютерных атаках или инцидентах ИБ. Как правило, эти системы функционируют в составе центров мониторинга ИБ (Security Operation Center, SOC), собирающих и анализирующих информацию об угрозах, уязвимостях и сценариях реализации атак в ИС с использованием методов кибербразведки (Threat Intelligence, TI). Для автоматизации этих процессов создаются специальные системы управления данными кибербразведки (СУДК), при построении и функционировании которых используются перспективные технологии интеллектуального анализа данных и комплексирование информации, поступающей в SOC из различных источников слабоструктурированных данных [1].

Под кибербразведкой при этом понимается TI-сервис, владелец которого получает возмож-

ность следить за преступной деятельностью злоумышленников (хакерских группировок), анализировать используемые ими методы предварительного анализа информации, внедрения и закрепления в ИС, а также перемещения внутри взломанной инфраструктуры и совершения вредоносных действий, связанных с хищением данных, нарушением работы ИС жертвы, запуском программы-вымогателя и т. п.

Основное достоинство кибербразведки в том, что она позволяет компаниям действовать на опережение, предотвращая возникающие угрозы, оперативно реагируя на выявленные атаки, используя при этом накопленную аналитиками компании информацию о применяемых методах и характере действий киберпреступников.

Процесс кибербразведки включает в себя следующие основные этапы [2, 3]:

1) планирование (устанавливаются цели, требования к получаемой информации и расставляются приоритеты в дальнейшем анализе);

2) сбор информации;

3) обработку собранных данных (их интерпретация, трансляция, унификация и интеграция);

4) подготовку данных (уточнение и слияние обработанной информации);

5) распространение информации конечным потребителям (включая как внешних потребителей, так и собственные службы ИБ компании).

Для автоматизации этих процессов используются специализированные коммерческие платформы (Threat Intelligence Platform, TIP), предлагаемые на рынке рядом российских и международных компаний: Group-IB, Лабораторией Касперского, Positive Technologies, R-Vision, Anomali, ElectricIQ, Threat Connect, Threat Quotinet и др. Достаточно широкую известность также получили TIP-решения с открытым исходным кодом: MISP (Malware Information Sharing Platform) и Yeti (Your Everyday Threat Intelligence).

Источниками, или каналами («фидами»), получения данных кибербразведки могут быть [1] средства массовой информации, социальные сети, бюллетени рассылок, мессенджеры (платные и бесплатные, открытые), хакерские форумы, DarkNet (нелегальные), частные и государственные организации, предоставляющие информацию о киберугрозах. К наиболее крупным источникам данных TI относятся государственные структуры НКЦКИ (Национальный координационный центр по компьютерным инцидентам) и ФинЦЕРТ (Центр взаимодействия и реагирования на компьютерные атаки в кредитно-финансовой сфере) Банка России.

Основными типами данных, собираемых сервисами TIP, являются [1, 4]:

— индикаторы компрометации (Indicators of Compromise, IoC) — наборы данных о вредоносных объектах или подозрительных действиях (анализируются на оперативном уровне);

— тактики, техники и процедуры (Tactics, Techniques, Procedures, TTP) — описания способов действий злоумышленников, содержащие различную детализацию (анализируются на тактическом уровне);

— индикаторы атак (Indicators of Attacks, IoAs) — правила, содержащие описание подозрительного поведения в системе, которое может быть признаком целевой атаки (анализируются на стратегическом уровне).

Наибольшую ценность для кибербразведки представляют прежде всего IoC. Именно на выявлении и анализе подобных индикаторов сегодня сфокусированы исследования в области проактивного поиска и предупреждения угроз [5]. IoC [5] — это также признаки (технические данные), такие как IP- и URL-адреса, доменные имена, адреса электронной почты, темы писем, ссылки и вложения, ключи реестра, имена и хеш-суммы файлов и т. п., которые можно использовать для идентификации действий или инструментов атакующих. Помимо этих данных, IoC могут включать в себя также информацию о контексте, т. е. содержать дополнительное описание рассматриваемой угрозы (например, название вредоносного ПО, названия группировок атакующих, ранее пострадавших компаний и т. д.).

В то же время, как отмечается в работе [5], использование IoC на практике нередко встречается с определенными трудностями. Содержащиеся в них данные являются разнородными; как правило, рассредоточены в различных источниках (фидах), где представлены в виде фрагментов слабоструктурированного текста; могут содержать неполную или, наоборот, избыточную информацию, местами могут противоречить друг другу и т. п., что является предметом заботы и необходимостью ручной обработки данных с привлечением специалистов высокой квалификации.

Выходом из данной ситуации является применение современных методов и технологий обработки естественного языка (Natural Language Processing, NLP), интеллектуального анализа текстов (Text Mining) и больших языковых моделей (Large Language Models, LLMs) [6]. В настоящей статье рассмотрены подходы к автоматизации сбора и обработки слабоструктурированных текстовых данных TI с использованием LLMs и технологии расширенного поиска релевантной информации (Retrieval Augmented Generation, RAG). Представлены результаты разработки прототипа автоматизированной системы анализа данных кибербразведки на основе технологии RAG, а также результаты экспериментов по

оценке эффективности разработанной системы с использованием натуральных данных TI из источников ФинЦЕРТ и НКЦКИ.

Обработка слабоструктурированных данных каналов распространения IoC с помощью LLMs и технологии RAG

Большие языковые модели — это большие нейронные сети на основе архитектуры трансформеров с миллиардами параметров, обученные на очень больших корпусах текстовых данных, взятых из самых различных источников (веб-сайтов, энциклопедий, книг, статей, материалов конференций и др.). Появление LLMs вызвало настоящую революцию в сфере обработки слабоструктурированной текстовой информации, представленной на естественном языке. Использование методов глубокого обучения позволяет им понимать сложные грамматические конструкции, смысл слов и предложений в зависимости от контекста, генерировать тексты (отчеты, рефераты, переводы и др.) в различных предметных областях. Поэтому понятен тот интерес, который проявляется к этим моделям со стороны специалистов по ИБ, и в том числе аналитиков в области киберразведки. Будучи дополнительно дообученными (этап fine-tuning) на специально отобранных профильных текстах, относящихся к сфере TI, что требует значительно меньше времени по сравнению с предварительным обучением моделей (этап pre-traing), LLMs оказываются способны извлекать ключевые факты (такие как IoC) из «сырых» слабоструктурированных данных, выявляя закономерности и аномалии, указывающие на вредоносную активность, предоставляя информацию о потенциальных угрозах и методах реализации атак, предлагая стратегии смягчения последствий от возможных атак [6–8].

В литературе приводится много примеров успешного применения LLMs в задачах сбора и обработки данных TI (киберразведки). Так, в работе [9] представлены результаты разработки прототипа ПО для оценки актуальных угроз безопасности ИС, позволяющего сопоставить и ранжировать угрозы нарушения безопасности информации для выявленного специалистом перечня уязвимостей ИС из банка данных угроз (БДУ) безопасности информации ФСТЭК России, автоматизировать подбор тактик и техник для построения возможных сценариев реализации угроз. В основе предложенного подхода используется алгоритм сопоставления множества выявленных уязвимостей, соответствующих им тактик, техник и релевантных угроз безопасности информации путем оценки метрик семантической близости из текстовых описаний с использованием технологии обработки естественного языка (Word Embedding) и нейросетевой модели трансформера BERT3 компании Google. Применение предложенного подхода позволяет упростить процедуру оценки актуальных угроз безопасности информации, а также автоматизировать процесс построения возможных сценариев их реализации, сокращая временные затраты на проведение анализа.

В работе [4] выполнен подробный анализ использования больших языковых моделей для поиска угроз кибербезопасности в целях повышения защищенности ИС. Раскрыты особенности и выполнена классификация направлений применения LLM в задачах обеспечения кибербезопасности, определены трудности внедрения LLM в процессы поиска угроз. В работе [10] сформулированы признаки, характеризующие относительно низкую эффективность использования LLM для извлечения именованных сущностей в области кибербезопасности.

В работе [11] предложен подход к построению и обучению LLM для решения задач обнаружения и классификации компьютерных атак, а также извлечения из текста именованных сущностей (IoC). В качестве LLM-системы используется нейросетевая модель трансформера BERT. Предварительно обученная нейронная сеть дообучалась в ходе проведенных экспериментов на большом корпусе текстовых описаний атак из MITRE ATT & CK, CAPEC, учебников, статей из Википедии, материалов конференций по проблематике ИБ, описаний уязвимостей из CVE и CWE. Отмечается высокая точность обработки и классификации слабоструктурированных данных TI, обеспечиваемая с помощью дообученной LLM.

В [12] приводятся результаты разработки агента искусственного интеллекта, позволяющего автоматизировать извлечение важной информации (и в первую очередь IoC) из больших объемов текстовых данных (например, отчетов TI) с использованием LLMs. Помимо этого, разработанный агент позволяет на основе полученной информации устанавливать зависимости между компонентами IoC, генерировать регулярные выражения (шаблоны для поиска фрагментов в тексте), фильтровать выходные реакции LLMs в ответ на запрос пользователя, чтобы добиться низкого уровня ошибочных решений. Подчеркивается, что использование данного агента позволяет аналитикам SOC лучше понять модели атак и сократить время реагирования на атаки, высвобождая в конечном итоге время специалиста на решение более творческих задач.

В [13] рассмотрены возможности применения LLMs, обученных на большом объеме неразмеченных данных, для построения плана поведения интеллектуальных агентов в многоагентных системах.

В работе [14] авторы оценивают перспективы применения больших языковых моделей и подчеркивают необходимость особого подхода со стороны регулирующих органов, так как LLMs являются глобальными моделями, и для них может потребоваться новая нормативная категория (согласно национальной стратегии развития искусственного интеллекта, при внедрении любых систем искусственного интеллекта должен соблюдаться принцип технологического суверенитета, включая преимущественное использование отечественных разработок).

В работах рассматриваются вопросы обеспечения безопасности интеграции LLMs в корпоративные информационные системы [15, 16].

В исследовании [17] авторы предлагают методическое, алгоритмическое и программное обеспечение для создания при ограниченных вычислительных ресурсах на основе больших языковых моделей автоматических систем обработки данных, в которых отсутствуют катастрофическое забывание, риск переобучения, галлюцинации.

Вместе с тем, несмотря на указанные преимущества LLMs, их применение на практике встречается с рядом вопросов. В работе [18] показано, что при решении практических задач в различных проблемных областях (доменах знаний) возникает множество трудностей, вызванных неоднородностью и сложностью обрабатываемых данных, вариабельностью целей анализа и синтеза решений с помощью LLM и разнообразными ограничениями. Кроме того, необходимо поддерживать доступ к конфиденциальным корпоративным знаниям и реализовывать механизм обновления базы знаний на основе LLM [19]. В ряде случаев LLM может генерировать ошибочные ответы и неверные утверждения, выдавая их как вполне обоснованные, — то, что называется галлюцинациями («связным бредом»). Это обычно происходит, когда запрос пользователя требует наличия знаний, которые выходят за рамки дополнительного обучения LLM (т. е. модель «этому не учили»).

На сегодня предложено [18, 20] несколько способов устранения указанных проблем (табл. 1).

В работе [21] представлен подход к созданию и дообучению русскоязычных LLMs (XGLM, LLaMA, ruGPT-3.5) для применения в различных доменах знаний. Приведено сравнение двух основных методик дообучения: дообучение всех параметров модели и дообучение слоев с использованием LoRA. Авторами показано, что качество предоставляемых для обучения данных существенно влияет на возможности модели, оцениваемые с помощью автоматических метрик качества MT-BENCH и MMLU.

Одним из наиболее гибких и эффективных способов борьбы с галлюцинациями LLM стало

■ ***Таблица 1.*** Способы адаптации LLM для решения прикладных задач специфической предметной области
■ ***Table 1.*** Methods of adapting LLM to solve applied problems of a specific subject area

Метод адаптации LLM	Характеристика	Особенности
Передача дополнительного контекста через запрос к LLM (prompt)	Разработка наиболее эффективных запросов для генерации ответов требуемой структуры и содержания	Не изменяет параметры LLM
Тонкая настройка модели (fine tuning, FT)	Адаптация параметров LLM в ходе дообучения на небольшом (по сравнению с исходным обучающим набором) объеме специально подготовленных данных: – адаптация нескольких (выходных) слоев модели; – адаптация всех параметров модели	Обновление весовых коэффициентов всех слоев LLM может существенно ухудшить способность к генерации ответов Необходимы существенные вычислительные ресурсы Возможно существенное возрастание качества генерируемых ответов
Генерация с расширенным поиском (RAG)	Сочетание достоинств традиционных информационно-поисковых систем с возможностями LLM	Не изменяет параметры LLM
Использование LLM для построения графа знаний и применения графа знаний как источника данных для LLM	Сочетание возможностей построения структурированной интерпретируемой модели внешней среды (графа знаний) и извлечения закономерностей из имеющихся данных с помощью LLM	Перспективное направление, позволяющее повысить эффективность анализа данных кибербразведки

появление новой технологии RAG — генерации ответов с расширенным поиском, впервые предложенной в 2020 г. в работе [22] (подробнее см. обзоры [23–25]).

Суть данной технологии заключается в формировании такой среды искусственного интеллекта, которая объединяет силу традиционных информационно-поисковых систем с возможностями современных LLMs.

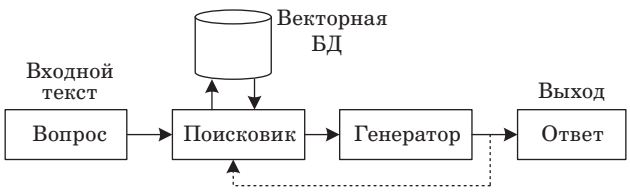
Архитектура RAG-системы в общем виде (рис. 1) состоит из двух основных частей — поисковика и генератора. Процесс ее работы начинается с ввода вопроса (или запроса) пользователя в поисковую систему. Поисковый модуль (поисковик) преобразует текст в последовательность числовых векторов (операция векторного вложения слов, Word Embedding), после чего производится семантический поиск релевантных документов в векторной базе данных, где хранятся документы компании, касающиеся в данном случае вопросов обеспечения ее ИБ. Из каждого документа извлекается текст, который используется для построения векторного вложения. Текст и вложение хранятся в базе данных. Поисковик на основе оценки семантической близости выбирает из базы данных несколько (Top-k) наиболее близких к сути вопроса (запроса) пользователя документов, содержащих максимально полезную контекстную информацию, необходимую для последующей обработки вопроса (запроса) модулем генератора.

Модуль генератора обычно реализуется на базе LLM трансформера типа GPT или BERT. Информация, полученная с поисковика (вопрос + найденный контекст), при этом также подвергается определенным NLP-операциям (токенизации, нормализации, стеммизации, лемматизации и др.) в целях приведения входных данных LLM к формату, более удобному для их обработки. Генератор генерирует ответ (выход) или инициализирует продолжение ответа, используя для этого предварительно обученную модель трансформера. Привлечение дополнительной релевантной информации (контекста) позволяет LLM дать более точный и обоснованный ответ на поступивший в систему вопрос (запрос).

Известны отдельные публикации, относящиеся к применению RAG-систем в области TI. Так, в работе [26] рассматриваются вопросы интеграции RAG с платформами TI в целях автоматизации процедур анализа и интерпретации киберугроз. Обсуждаются пути использования RAG для синтеза и контекстного обогащения данных TI из различных источников (фидов), включая логи, фиды угроз, отчеты об инцидентах ИБ, что позволяет более глубоко понимать причины и ландшафт угроз безопасности информации, повышать оперативность и эффективность принятия решений по предотвращению потенциальных угроз и смягчению последствий от их возникновения. Подчеркивается синергетический эффект от взаимодействия специалистов-профессионалов и применения технологии RAG, что является предпосылкой к достижению безопасной цифровой среды.

В [27] предложен метод автоматизированного построения графов атак (названный авторами Crystal Ball — «хрустальный мяч»), основанный также на применении технологии RAG. Поисковик RAG-системы при этом извлекает релевантную текстовую информацию из базы данных уязвимостей CVE, содержащей описания уязвимостей в виде цепочки основных этапов их реализации (предусловия, постусловия). В качестве источников данных TI могут также использоваться отчеты об угрозах безопасности информации. Обогащенный полученным контекстом, запрос пользователя поступает на вход LLM, которая строит граф компьютерной атаки, реализующей ту или иную уязвимость ИС, в форме графического представления путей атаки (действий злоумышленника), которые могут привести его к достижению поставленных целей. Приводятся примеры построения графов атак для конкретных уязвимостей с использованием различных LLMs (GPT-3.5, GPT-4, BARD), отмечается высокая эффективность, точность и релевантность обработки информации TI, в том числе возможность обработки данных в режиме реального времени, с использованием предложенного подхода.

В [28] предложен оригинальный подход к построению динамичных и адаптируемых сценариев учений по кибербезопасности для специалистов и сотрудников организаций. В основе разработанной автоматизированной системы генерации сценариев используется ансамбль из нескольких LLMs, одна из которых играет роль «эксперта по кибербезопасности», в функцию которой входит оценка текущего ландшафта киберугроз, тогда как другие LLMs выступают в качестве «исполнительных руководителей» организации, отвечающих за финансы, управление персоналом, инфраструктуру информационных технологий и т. д. Обогащение контекстом первоначального запроса на создание сценария



■ ***Рис. 1.*** Архитектура RAG-системы
■ ***Fig. 1.*** RAG system architecture

учений реализуется с использованием технологии RAG. Вместе с тем, как отмечают авторы, если одним из предназначений RAG обычно считается борьба с галлюцинациями LLMs, то в данном случае, наоборот, это вредное качество LLMs признается полезным свойством, позволяющим LLM генерировать более разнообразные (возможно, на первый взгляд, даже необычные) сценарии реализации угроз и выбора соответствующих стратегий и мер защиты. Однако окончательное решение в пользу выбора того или иного сгенерированного варианта сценария учений остается за человеком — профессионалом в области кибербезопасности. Авторами рассмотрено несколько примеров построения с использованием RAG детализированных сценариев проведения учений с учетом таких факторов, как выбор множества киберугроз и атак, особенностей информационной инфраструктуры организации, имеющихся технических средств защиты информации, распределения ролей участников в обеспечении кибербезопасности и т. д., что делает предложенные системой сценарии вполне обоснованными и реалистичными.

Повышение эффективности анализа данных кибербразведки и управления информацией о киберугрозах возможно [4] на основе перспективных подходов построения «комитетов экспертов» LLM, а также мультиагентных систем [29].

Разработка структурно-функциональной организации системы интеллектуального анализа данных кибербразведки

Существующие [30] подходы к обмену данными кибербразведки в машиночитаемых форматах характеризуются разнообразием стандартов, среди которых наиболее часто используются MISP и STIX. В то же время существенное количество данных, размещаемых на открытых платформах, представлено в структурированных текстовых форматах (plain text, CSV) и в виде слабоструктурированных отчетов на естественном языке. Как отмечалось ранее, ценным источником данных кибербразведки (но уже не в машиночитаемом формате) являются тематические блоги, каналы в мессенджерах и специализированные онлайн-ресурсы. Многообразие источников и форматов представления данных осложняет их преобразование в машиночитаемый формат конкретной платформы TI, так как необходимо сохранить семантический смысл исходных данных и контекст в рамках схемы хранения, определяемой выбранным форматом.

Развитие платформ TI тесно связано с обеспечением интерпретируемости данных — повышением эффективности преобразования форматов представления данных «человек-машина» и «машина-машина» — т. е. с простотой подключения новых источников данных, извлечением и сохранением максимального объема полезной информации при преобразовании форматов, организацией семантического поиска по анализируемым данным и диалогового интерфейса для специалистов.

Следовательно, для актуализации возможностей существующих СУДК значимо внедрение подсистем и модулей, позволяющих:

— предоставить интерфейс человеко-машинного взаимодействия в режиме «вопрос-ответ» для оперативного анализа собираемых данных;

— автоматизировать структурирование собираемых данных в конкретный формат платформы TI.

Возможным решением является разработка моделей извлечения знаний и автоматического построения онтологии и графа знаний проблемно-ориентированного корпуса для конкретной предметной области. Однако этот процесс требует трудоемкой разметки корпуса текстовых документов. Например [31], для неразмеченного корпуса текстов предложенное решение демонстрирует удовлетворительную работоспособность ввиду выделения атомарных фрагментов, включаемых в автоматически формируемую онтологию.

Другим рассматриваемым подходом для актуализации возможностей СУДК является применение RAG, позволяющей использовать открытые предобученные большие языковые модели общего назначения без необходимости дообучения для конкретной предметной области, а также без предварительного создания онтологии.

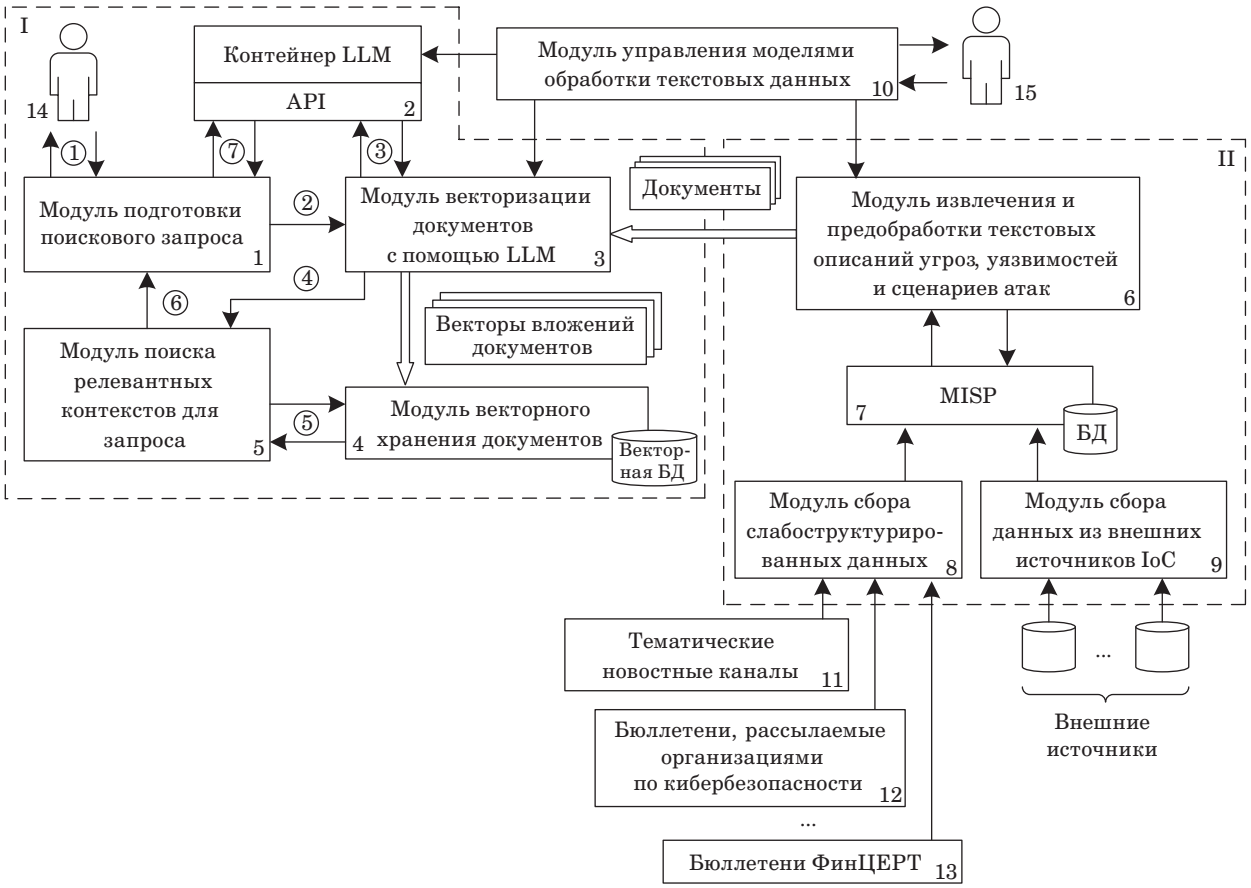
Таким образом, целью создания системы является повышение эффективности анализа слабоструктурированных и структурированных данных из различных источников с помощью развернутой платформы MISP [27].

Структурная схема представляемой системы анализа данных кибербразведки, включающей LLM с механизмом RAG в составе СУДК, представлена на рис. 2.

Обсуждаемый подход позволяет построить вопросно-ответную систему поддержки принятия решений для ранжированного набора собранных фрагментов документов, направленную на повышение оперативности анализа тактических и стратегических данных кибербразведки.

Предлагается выделить две основные подсистемы:

— подсистема I предназначена для обеспечения поддержки принятия решений на основе интеллектуального анализа данных с помощью конвейера RAG (цепочки предобработки и подготовки данных, поиска контекста и генерации ре-



■ ***Рис. 2.*** Структурная схема системы анализа данных киберразведки в составе SOC
■ ***Fig. 2.*** Structural diagram of the TI data analysis system as part of the SOC

шения) путем реализации человеко-машинного взаимодействия в режиме «вопрос-ответ»;

— подсистема II реализует текущий вариант сбора и обработки данных на основе платформы MISP в составе действующей СУДК и предназначена для автоматизации структурирования собираемых данных по определенным шаблонам в целях последующего использования.

На рис. 2 введены следующие обозначения: *1* — подготовка поискового запроса на естественном языке; *2* — предобработанный поисковой запрос эксперта; *3* — векторное представление текстового запроса (embedding) эксперта; *4*, *5* — поиск релевантных запросу фрагментов хранимых документов на основе меры сходства в пространстве векторных представлений; *6* — передача релевантных фрагментов документов для формирования контекста ответа; *7* — передача контекста и запроса в LLM для формирования ответа.

Работа модуля *1* начинается с подготовки поискового запроса. На этом этапе запрос проходит предобработку, включающую нормализацию и очистку текстовых данных для обеспечения качества и корректности дальнейшей обработки. После предобработки запрос преобразуется *3* в векторное представление с использованием модели LLM *2*. Этот этап критически важен, так как векторное представление запроса используется для поиска релевантных фрагментов документов в векторной базе предобработанного массива собранных данных *4* и последующего формирования контекста.

После создания векторного представления запроса начинается поиск *5* релевантных фрагментов документов. Поиск осуществляется на основе меры сходства в пространстве векторных представлений, что позволяет найти фрагменты, максимально соответствующие запросу пользователя. Найденные фрагменты извлекаются из векторной базы данных и передаются на следующий этап для формирования контекста ответа.

Формирование контекста *5* включает сбор и подготовку релевантных фрагментов, которые будут использованы для ответа на запрос пользователя. Контекст и исходный запрос передаются в контейнер LLM *2*. На этом этапе LLM использует предоставленный контекст для генерации окончательного ответа, который затем возвращается пользователю.

Модуль *10* управления моделями обработки текстовых данных координирует работу различных моделей, обеспечивая их интеграцию в общий процесс. Модуль *3* векторизации документов с помощью LLM преобразует текстовые документы в векторные представления. Модуль *6* извлечения и предобработки текстовых описаний угроз, уязвимостей и сценариев атак обрабатывает и подготавливает данные для анализа. Модуль *5* поиска релевантных контекстов проводит поиск на основе векторных представлений запросов. Модуль *4* векторного хранения документов обеспечивает хранение документов в векторной форме.

Платформа MISP *7* используется для сбора и обмена информацией о вредоносных программах, обеспечивая интеграцию данных из различных источников. Модули сбора слабоструктурированных данных *8* и сбора данных из внешних источников *9* собирают информацию из тематических новостных каналов и бюллетеней, обеспечивая актуальность и полноту базы данных. IoC используются для идентификации угроз, включая данные о вредоносных программах и уязвимостях.

Функциональная модель процесса анализа данных киберразведки приведена на рис. 3.

Далее при построении прототипа системы применяется схема Advanced RAG [32], использующая увеличенное контекстное окно анализа и методы суммаризации и сжатия контекста (рис. 4). Схема поясняет блок *4* «Поиск релевантного контекста и его обработка» функциональной модели и в свою очередь включает следующие блоки.

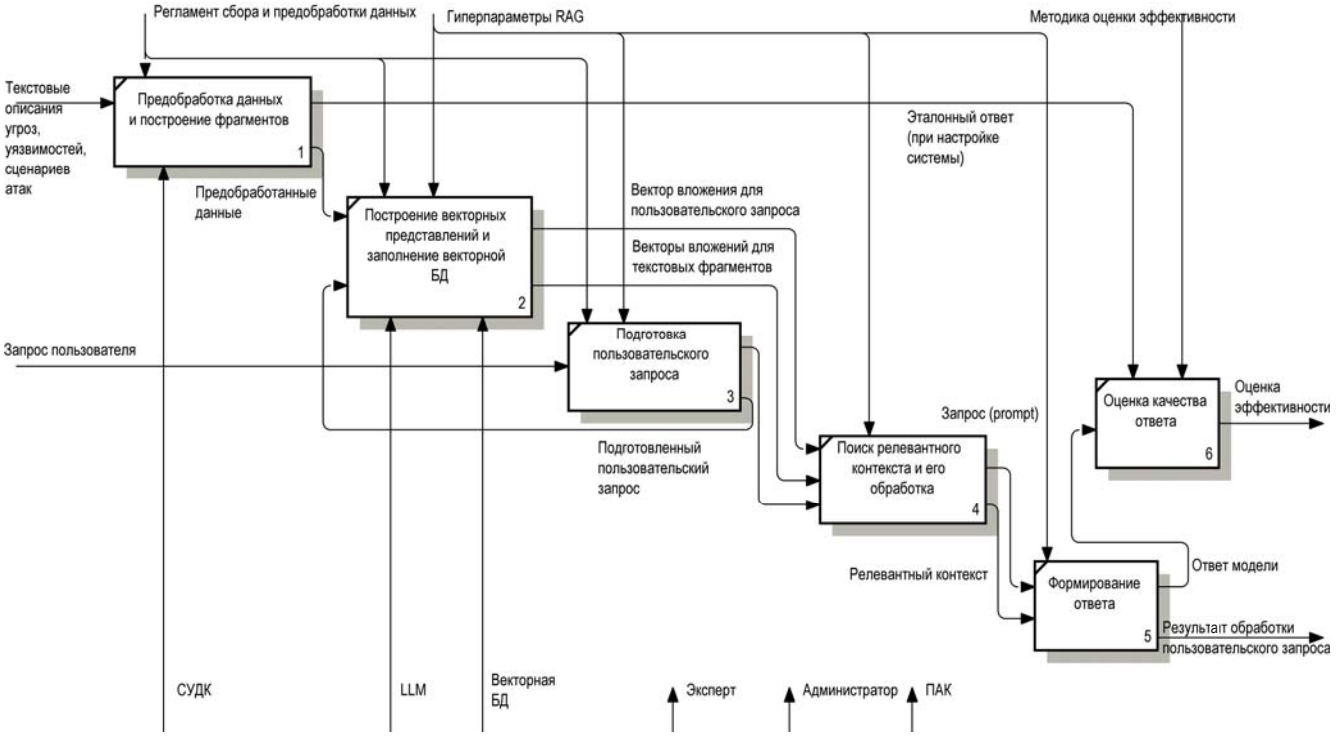
VectorStore (хранилище векторных представлений) — база данных, хранящая векторные представления документов. Используется для быстрого поиска релевантной информации на основе векторных запросов.

Search Sparse + Dense Vector (поиск с разреженными и плотными векторами) — выполняет комбинированный поиск на основе разреженного вектора признаков текстового описания (Sparse) (например, с помощью TF-IDF) и семантический поиск с помощью векторных (Dense) вложений (например, полученных с помощью LLM).

Eliminate Redundant Embeddings (удаление избыточных векторов вложений) — исключение избыточных векторов вложений, уменьшает количество нерелевантных или дублирующихся данных.

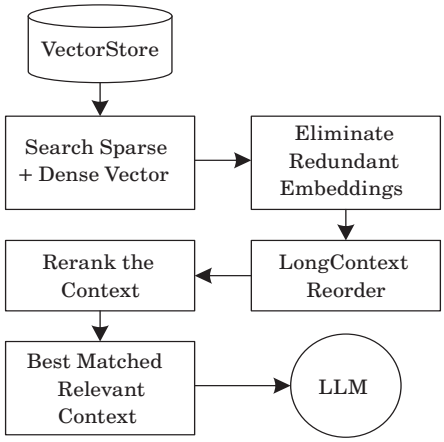
Rerank the Context (переранжирование контекста) — найденные текстовые фрагменты упорядочиваются по степени релевантности с помощью моделей ранжирования на основе LLM.

LongContext Reorder (перераспределение длинного контекста) — найденные фрагменты



■ ***Рис. 3.*** Функциональная модель процесса анализа данных киберразведки: *ПАК* — программно-аппаратный комплекс

■ ***Fig. 3.*** Functional model of the TI data analysis process



■ ***Рис. 4.*** Конвейер обработки данных Advanced RAG с суммаризацией контекста

■ ***Fig. 4.*** Advanced RAG pipeline with context summarization

переструктурируются так, чтобы их подача была наиболее логичной для генерации ответа.

Best Matched Relevant Context (наиболее релевантный контекст) — после ранжирования и перераспределения извлекается финальный контекст, который будет передан в LLM.

Алгоритм на рис. 5 иллюстрирует ключевые шаги, необходимые для реализации эффективной системы анализа семантической разметки и извлечения информации из слабоструктурированных данных киберразведки с точки зрения последующей программной реализации.

Схема описывает процесс извлечения и генерации текстовой информации с использованием больших языковых моделей и технологии RAG. LLM_1 применяется для векторизации документов, LLM_2 — для подготовки ответа на пользовательский запрос.

■ ***Рис. 5.*** Алгоритм семантической разметки и извлечения информации из слабоструктурированных данных

■ ***Fig. 5.*** The algorithm for semantic tagging and information extraction from semi-structured data

Проектирование исследовательского прототипа программного обеспечения анализа данных киберразведки

Компонентная модель исследовательского прототипа программного обеспечения для анализа данных киберразведки представлена на рис. 6. Экспертная оценка возможных кандидатов для использования в качестве ядра системы приведена в табл. 2. С целью обеспечить конфиденциальность обрабатываемых данных принято решение использовать развернутую локально LLM.

Переобученная LLM с параметрами 7.3B построена на основе Mistral, где ключевыми компонентами являются saiga_mistral_7b_gguf. Это обновление версии Mistral-7B-v0.2, которая занимала 52-е место в рейтинге Chatbot Arena, что выше, чем у GPT-3.5-Turbo-1106.

Алгоритмы предобработки текстовых данных детально проанализированы в работе авторов [33].

Механизм RAG реализован с использованием компонентных фреймворков LangChain [34] и OpenAI [35].

Работа программного прототипа начинается с получения данных из различных источников, таких как *MISP* и локальные файлы (например, текстовые файлы и PDF-документы, собранные из внешних источников). Данные из PDF-файлов извлекаются с помощью компонента *MISP PDF OCR Parser* (извлечение текстового слоя или оптическое распознавание символов). Извлеченные данные передаются в *LangChain Text Splitter*, который разделяет текст на фрагменты для дальнейшей обработки.

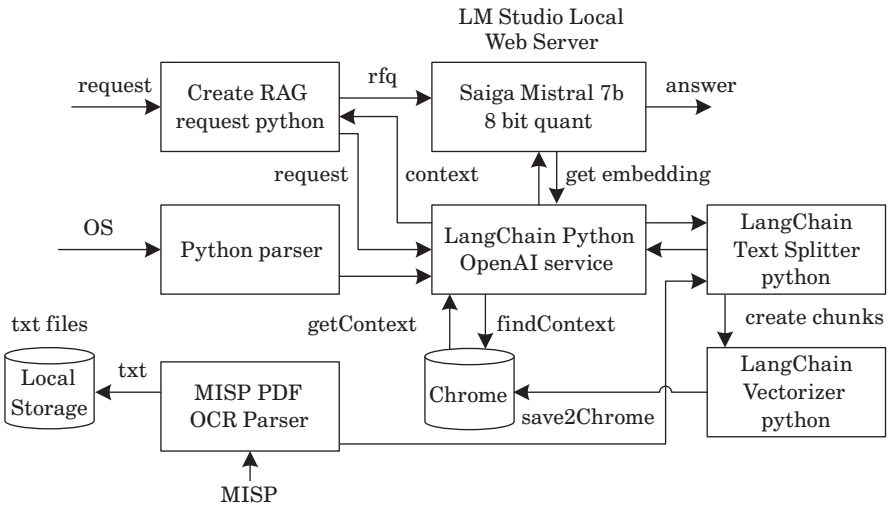
Затем данные проходят этап векторизации в *LangChain Vectorizer*, где текстовые фрагменты преобразуются в векторное представление. Для мультиязычных данных были использованы LLMs:

– sentence-paraphrase-multilingual-mpnet-base-v2;

■ ***Таблица 2.*** Оценка возможностей использования LLM в задаче обработки данных киберразведки

■ ***Table 2.*** Evaluation of the possibilities of using LLM in the task of processing TI data

Модель	Возможность локального запуска	Условия использования	Уровень адаптации к русскому языку	Быстродействие модели
ChatGPT	Нет	Платный API, нет доступа в РФ	Высокий	Высокое
Llama 13B	Да	Бесплатная	Низкий	Среднее
YandexGPT	Нет	Бесплатный API	Высокий	Высокое
Mistral 7B	Да	Бесплатная	Низкий	Среднее
Gemini	Нет	Платный API, нет доступа в РФ	Низкий	Высокое
Llama-3-Saiga	Да	Бесплатная	Высокий	Высокое



■ ***Рис. 6.*** Компонентная модель исследовательского прототипа программного обеспечения анализа данных киберразведки

■ ***Fig. 6.*** Component model of a research prototype of cyber intelligence data analysis software

— DeepPavlov/rubert-base-cased;
— multilingual-e5-large.

Векторные представления текстовых фрагментов сохраняются в векторной базе данных *Chroma*, обеспечивая быстрый доступ и поиск данных.

Для генерации запросов используется компонент *Create RAG request*, который отправляет запросы на обработку в *Saiga Mistral 7b 8 bit quant* — предобученную LLM.

Запросы пользователя и контексты обрабатываются с использованием *Python parser* и фреймворков *LangChain* и *OpenAI*, которые обеспечивают предобработку данных и взаимодействие с LLM. Полученные данные сохраняются и обрабатываются локально с помощью *Local Storage* и *LM Studio Local Web Server*, обеспечивая интеграцию всех компонентов системы и доступ к функционалу модуля через веб-интерфейс.

Микросервисная архитектура исследовательского прототипа (рис. 7) обеспечивает возможности вертикальной и горизонтальной масштабируемости.

Развертывание прототипа осуществлялось высокопроизводительным сервером со следующими параметрами:

— 20-ядерный процессор Intel Xeon E5-2698 v4 2,2 ГГц;
— 256 Гб ОЗУ RDIMM DDR4;
— четыре видеоускорителя Tesla V100 с суммарным объемом видеопамяти 128 Гб.

Каждая из моделей в изолированном окружении использовала выделенные графические процессоры, обмен данными между моделями осуществлялся через каналы в ОЗУ.

Проведение вычислительного эксперимента на собранном наборе данных

Цель вычислительного эксперимента — оценка работоспособности программного прототипа при реализации:

— вопросно-ответного диалогового взаимодействия эксперта с базой подготовленных документов об угрозах, уязвимостях, тактиках и техниках их реализации;

— структурирования текстового документа в формат MISP для последующего использования в СУДК.

В ходе разработки системы был создан набор данных на основе анализа более 20 источников, включающих тематические TI-каналы Telegram, отчеты о ландшафте угроз ведущих вендоров в области защиты информации (F.A.C.C.T., Лаборатория Касперского, Cisco и т. п.). С помощью действующей СУДК на основе платформы MISP были собраны данные за период 2022–2023 гг. из 10 основных каналов, включая, бюллетени ФинЦЕРТ и НКЦКИ. Данные представ-

■ ***Рис. 7.*** Микросервисная архитектура исследовательского прототипа программного обеспечения
■ ***Fig. 7.*** Microservice architecture of a research software prototype

лены в открытых источниках, собраны в текстовом формате и далее предобработаны.

Размер окна анализа текстового документа после предобработки составил 1000 токенов с перекрытием 300 токенов (1 токен = 2–4 символа текста). Общая база данных текстовых описаний и векторов вложений включала более 10 тыс. записей.

Вопросы к модели задавались в соответствии с эталонным списком подготовленных запросов — для каждого документа группой экспертов были подготовлены пять вопросов и эталонных ответов. Результаты генерации модели добавлялись в базу для дальнейшего сравнения с эталонными ответами эксперта. Итоговый набор текстов включал 100 предобработанных документов, для каждого из которых было подготовлено пять эталонных вопросов и ответов.

Модель способна находить правильные фрагменты текста и составлять по ним ответ на поставленный вопрос. В лучших итерациях ответ модели может практически полностью совпадать с эталонным ответом.

Анализ полученных результатов и оценка эффективности предложенного решения

Для оценки результатов работы системы были использованы следующие метрики.

1. Экспертная оценка: оценка по пятибалльной шкале (от единицы — худшее совпадение до пяти — полное совпадение), предоставленная экспертом на основе субъективного восприятия качества сгенерированных ответов. Эта оценка важна для понимания того, насколько результаты модели соответствуют ожиданиям специалистов предметной области.

2. BertScore: оценка (безразмерная величина в диапазоне [0, 1]) семантического сходства сгенерированного ответа и эталонного ответа. Использует модели BERT для вычисления векторных представлений текста и вычисления косинусного сходства между ними. Включает оценки точности (Precision), полноты (Recall) и гармонического среднего (F1-score):

— BertScore_pre: точность (Precision);

— BertScore_rec: полнота (Recall).

Высокие значения этих метрик показывают, что текст модели семантически близок к эталонному.

Результаты расчета среднего значения каждой из метрик приведены в табл. 3.

Метрики, рассчитанные индивидуально для каждого ответа модели, представлены в табл. 4.

Необходимо оценивать как эффективность работы модулей поиска и генерации по отдельности, так и работу конвейера RAG в целом. Помимо экспертной оценки результатов работы, существуют и иные метрики: ROUGE [36], BLEU [37] и RAGAS [38].

Применение фреймворка RAGAS (Retrieval Augmented Generation Automated Scoring) [38] позволяет автоматизировать оценку качества ответов модели в виде набора метрик (табл. 5). Состав предобработанного набора данных показан на рис. 8.

Использование как можно меньшего количества данных, аннотированных экспертом вручную, позволит снизить стоимость разработки систем и повысить оперативность оценки эффективности системы. Реализация фреймворка RAGAS позволяет проводить оценку эффективности системы в каждом из этих сценариев (с использованием аннотированных экспертами данных и без, когда в качестве эксперта выступает отдельная LLM-«критик»). Все значения метрик находятся в диапазоне [0, 1], при этом более высокие значения указывают на лучшую производительность. Для собранного набора данных ре-

■ ***Таблица 3.*** Результаты подсчета метрик и их анализ

■ ***Table 3.*** Results of metrics calculation and their analysis

Обозначение метрики	Среднее значение	Значение метрики	Анализ результата
Экспертная оценка	3,98	Высокие значения экспертной оценки говорят о том, что, с точки зрения профессионалов в данной области, результаты считаются качественными и удовлетворительными	Оценка показывает, что ответы модели находятся на высоком уровне и близки к эталонным. Это свидетельствует о том, что модель успешно генерирует релевантные и полезные ответы на вопросы
BERTScore_pre	0,9	Высокие баллы BERTScore означают, что ответы языковой модели семантически и синтаксически схожи с рефератами	Высокие значения BERTScore показывают, что модель хорошо справляется с задачами, где важна семантическая точность, способна генерировать тексты, близкие по смыслу к эталонным
BERTScore_rec	0,88		
BERTScore F1	0,89		

■ ***Таблица 4.*** Рассчитанные для ответов метрики
■ ***Table 4.*** Calculated metrics for responses

Экспертная оценка	BERTScore_pre	BERTScore_rec	BERTScore F1
5	0,99	0,94	0,96
5	0,76	0,45	0,57
5	0,93	0,88	0,9
5	0,94	0,94	0,94
5	0,92	0,88	0,9
5	0,94	0,94	0,94
5	0,92	0,92	0,92

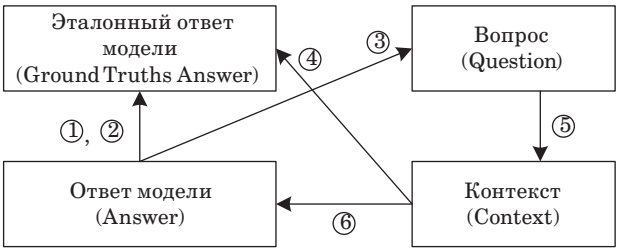
■ ***Таблица 5.*** Метрики RAGAS для полного набора данных
■ ***Table 5.*** RAGAS metrics for the full dataset

Метрика	Характеристика	Оцениваемая часть конвейера RAG	Значение
Точность контекста	Оценивает соотношение сигнала к шуму извлеченного контекста, вычисляется с использованием вопроса и контекстов	Отдельные этапы анализа RAG	0,803
Полнота контекста	Измеряет, была ли извлечена вся необходимая для ответа на вопрос информация. Эта метрика вычисляется на основе эталонных данных — ground_truth (это единственная метрика в фреймворке, которая опирается на аннотированные человеком метки истинности) и контекстов	То же	0,794
Релевантность контекста	Оценивает релевантность извлеченного контекста, определяемую на основе как вопроса, так и контекстов	–//–	–
Верность	Измеряет фактическую согласованность сгенерированного ответа по отношению к заданному контексту. Количество правильных утверждений из заданных контекстов делится на общее количество утверждений в подготовленном ответе (на основе вопросов, контекстов и ответов)	–//–	0,831
Релевантность ответа	Оценивает, насколько релевантен ответ, подготовленный моделью, соответствует вопросу. Более низкий балл присваивается ответам, которые являются неполными или содержат избыточную информацию, а более высокие баллы указывают на более высокую релевантность. Для оценки метрики используются вопросы и ответы	–//–	0,813
Семантическое сходство ответов	Оценивает семантическое сходство между сгенерированным ответом и истинным значением	Сквозная оценка конвейера RAG	–
Корректность ответов	Включает в себя измерение точности сгенерированного ответа по сравнению с истинным значением	То же	0,822

зультаты оценки с помощью RAGAS приведены в табл. 5.

В ходе эксперимента было установлено, что модель показывает высокие результаты по большинству метрик. Экспертная оценка 3,98 подтверждает, что генерируемые ответы воспринимаются как качественные и полезные. Метрики BERTScore показывают, что модель генерирует ответы, которые семантически схожи с эталонными, что также поддерживает высокую оценку качества модели. Совокупность метрик RAGAS также подтверждает способность системы корректно отвечать на поставленные вопросы, что делает ее полезным инструментом для анализа неструктурированных и структурированных данных кибербразведки.

Для оценки эффективности системы (снижение трудоемкости обработки данных и повышение оперативности анализа) эксперт в области кибербразведки анализировал открытые новостные сообщения тематического канала и структурировал информацию вручную, во втором сцена-



① корректность ответа (answer correctness)
② семантическое сходство ответов (answer similarity)
③ релевантность ответа (answer relevancy)
④ полнота контекста (context recall)
⑤ точность контекста (context precision)
⑥ верность (faithfulness)

■ ***Рис. 8.*** Основные сущности RAGAS
■ ***Fig. 8.*** Main entities of RAGAS

■ ***Таблица 6.*** Результаты эксперимента
■ ***Table 6.*** Results of the experiment

Сценарий и длительность	Ручная обработка одним экспертом	Автоматическая обработка
Сбор информации за один день	20 записей киберразведки из 10 источников за 30 мин	400 записей киберразведки из 10 источников за 30 мин
Анализ и структурирование информации за один день	20 записей из 10 источников за 150 мин	400 записей из 10 источников за 33 мин
Сбор информации за неделю	140 записей за 250 мин	2100 записей за 250 мин
Анализ и структурирование информации за неделю	140 записей за 1050 мин	2100 записей за 210 мин

рии та же задача выполнялась с использованием предлагаемой системы (табл. 6).

Таким образом, на сбор данных для одной записи в базу TI эксперт тратил в среднем 1,5 мин, при использовании предлагаемой системы — 0,075 мин. При структурировании одной записи в формат MISP эксперт тратил 7,5 мин, при использовании системы — 0,083 мин.

Заключение

Рассмотрены способы повышения эффективности (повышение оперативности и снижение трудоемкости) анализа данных киберразведки с помощью методов искусственного интеллекта:

— разработана структурно-функциональная организация (структурная схема и функциональная модель) системы интеллектуального анализа данных киберразведки на основе больших языковых моделей с применением конвейера RAG;

— разработан алгоритм семантической разметки и извлечения информации из слабоструктурированных данных киберразведки для последующего анализа с помощью LLM в составе вопросно-ответной системы поддержки принятия решений;

— разработан исследовательский прототип (компонентная модель, диаграммы вариантов использования, микросервисная архитектура) программного обеспечения для анализа данных киберразведки с помощью больших языковых моделей, оценена эффективность предложенного решения в задаче обработки слабоструктурированных данных из каналов TI.

Проведенный эксперимент показал согласованность экспертных ответов по набору собранных документов с ответами, предложенными моделью, на уровне 3,98 балла из пяти. Значение метрики BERTScore F1 составило 0,89, метрик RAGAS «корректность ответов» (Answer Correctness) — 0,822. Временные затраты на сбор и подготовку данных для СУДК снизились более чем в 20 раз.

Выявлены особенности реализации систем интеллектуального анализа данных TI:

— сложность проектирования и реализации с использованием современных фреймворков программного прототипа системы, использующей конвейер RAG для управления данными киберразведки, ниже, чем сложность настройки гиперпараметров отдельных этапов обработки и формирования набора метрик для непрерывной оценки эффективности анализа и адекватности ответов модели;

— оценка эффективности системы требует высококачественного набора проверочных данных, подготовленного экспертами, что является сложным, трудоемким и дорогостоящим процессом.

Новизна предлагаемых алгоритма и модели основана на алгоритмах поисковой дополненной генерации с помощью больших языковых моделей, отличительной особенностью является комплексирование сведений из нескольких источников распространения IoC и последующее извлечение информации из слабоструктурированных накапливаемых данных с помощью моделей машинного обучения в составе конвейера RAG. Это позволяет повысить эффективность анализа данных киберразведки для оказания значимой поддержки специалистам по защите информации в ходе анализа актуальных угроз и

уязвимостей за счет снижения трудозатрат (времени сбора и обработки данных).

Несмотря на наблюдающиеся тенденции увеличения контекстного окна LLM, механизм RAG позволяет реализовать разделение прав и управление доступом к модели, что невозможно реализовать только лишь на уровне LLM.

Перспективным направлением дальнейших исследований является построение «графов знаний» (Knowledge Graphs) с метаданными (Knowledge Maps), описывающими структуру хранения и связи между сущностями предметной области (Graph Retrieval Augmented Generation).

Финансовая поддержка

Работа выполнена в ОмГТУ в рамках государственного задания Минобрнауки России на 2023–2025 гг. № FSGF-2023-0004.

Литература

1. **Вульфин А. М.** Система управления данными кибербразведки. *Моделирование, оптимизация и информационные технологии*, 2021, т. 9, № 1 (32). doi:10.26102/2310-6018/2021.32.1.020, EDN: RDVDTE
2. **Ramsdale A., Shiaeles S., Kolokotronis N.** A comparative analysis of cyber-threat intelligence sources, formats and languages. *Electronics (MDPI)*, 2020, vol. 9, Article 824. doi:10.3390/electronics9050824. https://www.mdpi.com/journal/electronics (дата обращения: 10.12.2024).
3. **Ainslie S., Thompson D., Maynard S., Ahmad A.** Cyber-threat intelligence for security decision-making: A review and research agenda for practice. *Computers & Security*, 2023, vol. 132, pp. 103352. doi:10.1016/j.cose.2023.103352
4. **Котенко И. В., Абраменко Г. Т.** Использование больших языковых моделей для поиска угроз кибербезопасности на основе методов глубокого обучения: анализ современных исследований. *Правовая информатика*, 2024, № 3, с. 32–42. doi:10.24682/1994-1404-2024-3-32-42
5. **Мещеряков Р. В., Исхаков С. Ю.** Исследование индикаторов компрометации для средств защиты информационных и киберфизических систем. *Вопросы кибербезопасности*, 2022, № 5, с. 82–99. doi:10.21681/2311-3456-2022-5-82-99
6. **Gholami Y.** Large Language Models (LLMs) for cybersecurity: A systematic review. *World Journal of Advanced Engineering Technology and Sciences*, 2024, vol. 13(01), pp. 057–069. doi:10.30574/wjaets.2024.13.1.0395
7. **Chen Y., Cui M., Wang D., Cao Y., Yang P., Jiang B., Lu Zh., Liu B.** A survey of large language models for cyber threat detection. *Computers & Security*, 2024, vol. 145, iss. C, pp. 104016. doi:10.1016/j.cose.2024.104016, EDN: NVEXHQ
8. **Hasanov I., Virtanen S., Hakkala A., Isoaho J.** Application of Large Language Models in cybersecurity: A systematic literature review. *IEEE Access*, 2024, vol. 12, pp. 176751–176778. doi:10.1109/ACCESS.2024.3505983
9. **Васильев В. И., Вульфин А. М., Кучкарова Н. В.** Оценка актуальных угроз безопасности информации с помощью технологии трансформеров. *Вопросы кибербезопасности*, 2022, № 2 (48), с. 27–38. doi:10.21681/2311-3456-2022-2-27-38
10. **Bolla A., Talentino F.** *Threat Hunting Driven by Cyber Threat Intelligence*. Diss. Polytechnic University of Turin, 2022. 162 p.
11. **Park Y., You W.** A pretrained language model for cyber threat intelligence. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Industry Track*, December 6–10, 2023, pp. 113–122. doi:10.18653/v1/2023.emnlp-industry.12
12. **Kucsván Z. L., Caselli M., Peter A., Continella A.** Inferring recovery steps from cyber threat intelligence reports. *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessmen*, LNCS, 2024, vol. 14828, pp. 330–349.
13. **Ковалев А. К., Панов А. И.** Применение предобученных больших языковых моделей в задачах воплощенного искусственного интеллекта. *Доклады Российской академии наук. Математика, информатика, процессы управления*, 2022, т. 508, с. 94–99. doi:10.31857/S268695432207013X
14. **Андрейченко А. Е., Гусев А. В.** Перспективы применения больших языковых моделей в здравоохранении. *Национальное здравоохранение*, 2023, т. 4, № 4, с. 48–55. doi:10.47093/2713-069X.2023.4.4.48-55, EDN: MPDTJU
15. **Евглевская Н. В., Казанцев А. А.** Обеспечение безопасности сложных систем с интеграцией больших языковых моделей: анализ угроз и методов защиты. *Экономика и качество систем связи*, 2024, № 4 (34), с. 129–144. EDN: CJEAAZ. https://journal-ekss.ru/wp-content/uploads/2024/12/129-144.pdf (дата обращения: 10.12.2024).
16. **Мударова Р. М., Намиот Д. Е.** Противодействие атакам типа инъекция подсказок на большие языковые модели. *International Journal of Open Information Technologies*, 2024, т. 12, № 5, с. 39–48. EDN: LVERMB. http://injoit.org/index.php/j1/article/view/1845/1682 (дата обращения: 10.12.2024).
17. **Самонов А. В., Бурова И. О.** Методика разработки автоматизированных средств генерации программного кода посредством настройки больших языковых моделей. *Вопросы кибербезопасности*, 2024, № 3 (61), с. 68–75. doi:10.21681/2311-3456-2024-3-68-75, EDN: NPSBQW
18. **Зеленков Ю. А.** Управление знаниями организации и большие языковые модели. *Российский жур-*

нал менеджмента, 2024, т. 22, № 3, с. 573–601. doi:10.21638/spbu18.2024.309, EDN: KKHPPZ

19. **O'Leary D. E.** The rise and design of enterprise large language models. *IEEE Intelligent Systems*, 2024, vol. 39, no. 1, pp. 60–63. doi:10.1109/MIS.2023.3345591
20. **Yu W., Zhu C., Li Z., Hu Z., Wang Q., Ji H., Jiang M.** A survey of knowledge-enhanced text generation. *ACM Computing Surveys*, 2022, vol. 54, no. 11s, pp. 1–38. doi:10.1145/3512467
21. **Косенко Д. П., Куратов Ю. М., Жарикова Д. Р.** Большие языковые модели для следования инструкциям на русском языке: модели и датасеты с открытой лицензией для коммерческого использования. *Доклады Российской академии наук. Математика, информатика, процессы управления*, 2023, т. 514, № 2, с. 262–269. doi:10.31857/S2686954323602063, EDN: GECIST
22. **Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V., Goyal N., Küttler H., Lewis M., Yih W., Rocktäschel T., Riedel S., Kiela D.** Retrieval-Augmented Generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 2020, vol. 33, pp. 9459–9474.
23. **Cai D., Wang Y., Liu L., Shi S.** Recent advances in retrieval-augmented text generation. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Madrid, Spain, July 11–15, 2022. New York, 2022, pp. 3417–3419. doi:10.1145/3477495.3532682
24. **Fan W., Ding Y., Ning L., Wang S., Li H., Yin D., Chua T.-S., Li Q.** A survey on rag meeting llms: Towards retrieval-augmented large language models. *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, Barcelona, Spain, August 25–29, 2024. New York, 2024, pp. 6491–6501. doi:10.1145/3637528.3671470
25. **Li Z., Li C., Zhang M., Mei Q., Bendersky M.** Retrieval augmented generation or long-context LLMs? A comprehensive study and hybrid approach. *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track*, Miami, Florida, November 12–16, 2024, pp. 881–893. doi:10.18653/v1/2024.emnlp-industry.66
26. **Singh J. P., Agrawal Sh.** Automating threat intelligence analysis with Retrieval Augmented Generation (RAG) for enhanced cybersecurity posture. *International Journal of Science and Research (IJSR)*, 2024, vol. 13, iss. 5, pp. 251–255. doi:10.21275/SR24502103758
27. **Zhang Y., Du T., Ma Y., Wang X., Xie Y., Yang G., Lu Y., Chang E. C.** AttacKG+: Boosting attack graph construction with Large Language Models. *Computers & Security*, 2024, vol. 150, pp. 104220. doi:10.1016/j.cose.2024.104220
28. **Yamin M. M., Hashmi E., Ullah M., Katt B.** Application of LLMs for generating cyber security exercise scenarios. *Research Square*, Feb. 2024. doi:10.21203/rs.3.rs-3970015/v1. https://www.researchsquare.com/article/rs-3970015/v1 (дата обращения: 10.12.2024).
29. **Wu Q., Bansal G., Zhang J., Wu Y., Li B., Zhu E., Jiang L., Zhang X., Zhang S., Liu J., Awadallah A. H., White R. W., Burger D., Wang C.** AutoGen: Enabling next-gen LLM applications via multi-agent conversations. *First Conference on Language Modeling*, Pennsylvania, PA, October 7–9, 2024, pp. 1–46. https://openreview.net/pdf?id=BAakY1hNKS (дата обращения: 10.12.2024).
30. **Mavroeidis V., Bromander S.** Cyber threat intelligence model: An evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence. *2017 European Intelligence and Security Informatics Conference (EISIC)*, Greece, September 11–13, 2017. IEEE, 2017, pp. 91–98. doi:10.1109/EISIC.2017.20
31. **Зулкарнеев Р. Х., Юсупова Н. И., Сметанина О. Н., Гаянова М. М., Вульфин А. М.** Методы и модели извлечения знаний из медицинских документов. *Информатика и автоматизация*, 2022, т. 21, № 6, с. 1169–1210. doi:10.15622/ia.21.6.4, EDN: ASOOVS
32. **Rajapaksha S., Rani R., Karafili E.** A RAG-based question-answering solution for cyber-attack investigation and attribution. *Workshop on Security and Artificial Intelligence 2024, 29th European Symposium on Research in Computer Security*, 2024. http://eprints.soton.ac.uk/id/eprint/493520 (дата обращения: 10.12.2024).
33. **Васильев В. И., Вульфин А. М., Кучкарова Н. В.** Автоматизация анализа уязвимостей программного обеспечения на основе технологии Text Mining. *Вопросы кибербезопасности*, 2020, № 4 (38), с. 22–31. doi:10.21681/2311-3456-2020-04-22-31, EDN: TMRYIT
34. **Topsakal O., Akinci T. C.** Creating Large Language Model applications utilizing LangChain: A primer on developing LLM apps fast. *International Conference on Applied Engineering and Natural Sciences*, 2023, vol. 1, no. 1, pp. 1050–1056. doi:https://doi.org/10.59287/icaens.1127
35. **Auger T., Saroyan E.** *Overview of the OpenAI APIs.* Generative AI for Web Development: Building Web Applications Powered by OpenAI APIs and Next.js. Berkeley, CA, Apress, 2024, pp. 87–116.
36. **Lin C. Y.** ROUGE: A package for automatic evaluation of summaries. *Proceedings of the Workshop on Text Summarization Branches Out (WAS 2004)*, Barcelona, Spain, July 25–26, 2004, pp. 74–81.
37. **Papineni K., Roukos S., Ward T., Zhu W. J.** BLEU: A method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 2002, pp. 311–318. https://doi.org/10.3115/1073083.1073135
38. **Es S., James J., Anke L. E., Schockaert S.** RAGAs: Automated evaluation of Retrieval Augmented Generation. *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, 2024, pp. 150–158. doi:10.48550/arXiv.2309.15217

UDC 004.056
doi:10.31799/1684-8853-2025-2-50-67
EDN: QFQTPU

Automated system for analyzing weakly structured threat intelligence data using large language models

A. I. Lutskovich[a], Assistant Professor, orcid.org/0009-0000-7257-6947
V. I. Vasilyev[a], Dr. Sc., Tech., Professor, orcid.org/0000-0002-6105-5481
A. M. Vulfin[a], Dr. Sc., Tech., Professor, orcid.org/0000-0001-5857-2413, vulfin.am@ugatu.su
A. D. Kirillova[a], PhD, Tech., Associate Professor, orcid.org/0009-0000-4164-2526
A. E. Sulavko[b], Dr. Sc., Tech., Professor, orcid.org/0000-0002-9029-8028
[a]Ufa University of Science and Technology, 12, K. Marx St., 450008, Ufa, Russian Federation
[b]Omsk State Technical University, 11, Mira Pr., 644050, Omsk, Russian Federation

Introduction: The lack of efficiency and the complexity of the analysis and management of semi-structured and structured Threat Intelligence data collected from various sources is a pressing issue. **Purpose:** To increase the efficiency (which implies reducing the labor costs of experts and ensuring the completeness and promptness of the knowledge base updates) of Threat Intelligence data analysis using artificial intelligence methods. **Results:** The conducted analysis of approaches to the construction of Threat Intelligence data management systems has shown as a promising direction the use of large language models as part of a "question-answer" decision support system with an extended augmented generation mechanism. We develop a structural diagram and functional model of a data mining system for the Threat Intelligence data. It is based on large language models using a retrieval augmented generation pipeline. In addition, we develop an algorithm for semantic tagging and extracting information from weakly structured data. We present a research prototype (component model, microservice architecture) of the software. A distinctive feature of the system is the integration of information from multiple sources using a retrieval augmented generation pipeline. A computational experiment on the prepared data set has shown the consistency of expert responses for the set of collected documents with the responses proposed by the system at the level of 3.98 points out of 5. The BERTScore F1 metric value is 0.89, and the RAGAS Answer Correctness metric value is 0.822. **Practical relevance:** The use of large language models provides the accumulation of knowledge about current attack scenarios within a single knowledge base, which makes it possible to enhance the efficiency of data processing to enable a significant support to information security specialists during the analysis of current threats, vulnerabilities and cyber attack scenarios by reducing labor costs (time for collecting and processing), and thereby to increase the security of corporate information systems. **Discussion:** A further improvement of the efficiency of the Threat Intelligence data analysis is possible on the basis of constructing heterogeneous "expert committees" out of large language models and multi-agent systems.

Keywords — information security, Threat Intelligence, vulnerability analysis, Indicators of Compromise, Large Language Models, Retrieval-Augmented Generation.

For citation: Lutskovich A. I., Vasilyev V. I., Vulfin A. M., Kirillova A. D., Sulavko A. E. Automated system for analyzing weakly structured threat intelligence data using large language models. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2025, no. 2, pp. 50–67 (In Russian). doi:10.31799/1684-8853-2025-2-50-67, EDN: QFQTPU

Financial support
The research was carried out in Omsk State Technical University within the State assignment of the Ministry of Science and Higher Education of Russian Federation (theme No. FSGF-2023-0004).

References

1. Vulfin A. M. Cyber threat intelligence data management system. *Modeling, Optimization and Information Technology*, 2021, vol. 9, no. 1 (32) (In Russian). doi:10.26102/2310-6018/2021.32.1.020, EDN: RDVDTE
2. Ramsdale A., Shiaeles S., Kolokotronis N. A comparative analysis of cyber-threat intelligence sources, formats and languages. *Electronics (MDPI)*, 2020, vol. 9, p. 824. doi:10.3390/electronics9050824. Available at: https://www.mdpi.com/journal/electronics (accessed 10 December 2024).
3. Ainslie S., Thompson D., Maynard S., Ahmad A. Cyber-threat intelligence for security decision-making: A review and research agenda for practice. *Computers & Security*, 2023, vol. 132, pp. 103352. doi:10.1016/j.cose.2023.103352
4. Kotenko I. V., Abramenko G. T. Using large language models for cyber threat hunting based on deep learning methods: Analysis of modern research. *Legal Informatics*, 2024, no. 3, pp. 32–42 (In Russian). doi:10.24682/1994-1404-2024-3-32-42
5. Meshcheryakov R. V., Iskhakov S. U. Study of compromet-ation indicators for improvement of information and cyber-physical systems protection facilities. *Cybersecurity Issues*, 2022, no. 5, pp. 82–99 (In Russian). doi:10.21681/2311-3456-2022-5-82-99
6. Gholami Y. Large Language Models (LLMs) for cybersecurity: A systematic review. *World Journal of Advanced Engineering Technology and Sciences*, 2024, vol. 13(01), pp. 057–069. doi:10.30574/wjaets.2024.13.1.0395
7. Chen Y., Cui M., Wang D., Cao Y., Yang P., Jiang B., Lu Zh., Liu B. A survey of large language models for cyber threat detection. *Computers & Security*, 2024, vol. 145, iss. C, pp. 104016. doi:10.1016/j.cose.2024.104016, EDN: NVEXHQ
8. Hasanov I., Virtanen S., Hakkala A., Isoaho J. Application of Large Language Models in cybersecurity: A systematic literature review. *IEEE Access*, 2024, vol. 12, pp. 176751–176778. doi:10.1109/ACCESS.2024.3505983
9. Vasilyev V. I., Vulfin A. M., Kuchkarova N. V. Assessment of current threats to information security using transformer technology. *Cybersecurity Issues*, 2022, no. 2 (48), pp. 27–38 (In Russian). doi:10.21681/2311-3456-2022-2-27-38
10. Bolla A., Talentino F. *Threat Hunting Driven by Cyber Threat Intelligence*. Diss. Polytechnic University of Turin, 2022. 162 p.
11. Park Y., You W. A pretrained language model for cyber threat intelligence. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Industry Track*, 2023, pp. 113–122. doi:10.18653/v1/2023.emnlp-industry.12
12. Kucsván Z. L., Caselli M., Peter A., Continella A. Inferring recovery steps from cyber threat intelligence reports. *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessmen*, LNCS, 2024, vol. 14828, pp. 330–349.
13. Kovalev A. K., Panov A. I. Application of pretrained large language models in embodied artificial intelligence. *Doklady Rossijskoj Akademii Nauk. Mathematika, Informatika, Processy Upravlenia*, 2022, vol. 508, pp. 94–99 (In Russian). doi:10.31857/S268695432207013X
14. Andreychenko A. E., Gusev A. V. Perspectives on the application of large language models in healthcare. *National Health Care*, 2023, vol. 4, no. 4, pp. 48–55 (In Russian). doi:10.47093/2713-069X.2023.4.4.48-55, EDN: MPDTJU
15. Evlevskaya N. V., Kazantsev A. A. Ensuring the security of complex systems with the integration of large language models: Threat analysis and protection methods. *Ekonomika i kachestvo sistem svyazi*, 2024, no. 4 (34), pp. 129–144.

EDN: CJEAAZ. Available at: https://journal-ekss.ru/wp-content/uploads/2024/12/129-144.pdf (accessed 10 December 2024) (In Russian).
16. Mudarova R. M., Namiot D. E. Countering prompt injection attacks on large language models. *International Journal of Open Information Technologies*, 2024, vol. 12, no. 5, pp. 39–48. EDN: LVERMB. Available at: http://injoit.org/index.php/j1/article/view/1845/1682 (accessed 10 December 2024) (In Russian).
17. Samonov A. V., Burova I. O. Methodology for the development of automated software code generation tools by fine-tuning large language models. *Cybersecurity Issues*, 2024, no. 3 (61), pp. 68–75 (In Russian). doi:10.21681/2311-3456-2024-3-68-75, EDN: NPSBQW
18. Zelenkov Yu. A. Knowledge management in organization and the large language models. *Russian Management Journal*, 2024, vol. 22, no. 3, pp. 573–601 (In Russian). doi:10.21638/spbu18.2024.309, EDN: KKHPPZ
19. O'Leary D. E. The rise and design of enterprise large language models. *IEEE Intelligent Systems*, 2024, vol. 39, no. 1, pp. 60–63. doi:10.1109/MIS.2023.3345591
20. Yu W., Zhu C., Li Z., Hu Z., Wang Q., Ji H., Jiang M. A survey of knowledge-enhanced text generation. *ACM Computing Surveys*, 2022, vol. 54, no. 11s, pp. 1–38. doi:10.1145/3512467
21. Kosenko D. P., Kuratov Y. M., Zharikova D. R. Accessible Russian large language models: Open-source models and instructive datasets for commercial applications. *Doklady Rossijskoj Akademii Nauk. Mathematika, Informatika, Processy Upravlenia*, 2023, vol. 514, no. 2, pp. 262–269 (In Russian). doi:10.31857/S2686954323602063, EDN: GECIST
22. Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V., Goyal N., Küttler H., Lewis M., Yih W., Rocktäschel T., Riedel S., Kiela D. Retrieval-Augmented Generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 2020, vol. 33, pp. 9459–9474.
23. Cai D., Wang Y., Liu L., Shi S. Recent advances in retrieval-augmented text generation. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2022, pp. 3417–3419. doi:10.1145/3477495.3532682
24. Fan W., Ding Y., Ning L., Wang S., Li H., Yin D., Chua T.-S., Li Q. A survey on rag meeting LLMs: Towards retrieval-augmented large language models. *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024, pp. 6491–6501. doi:10.1145/3637528.3671470
25. Li Z., Li C., Zhang M., Mei Q., Bendersky M. Retrieval augmented generation or long-context LLMs? A comprehensive study and hybrid approach. *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track*, 2024, pp. 881–893. doi:10.18653/v1/2024.emnlp-industry.66
26. Singh J. P., Agrawal Sh. Automating threat intelligence analysis with Retrieval Augmented Generation (RAG) for enhanced cybersecurity posture. *International Journal of Science and Research (IJSR),* 2024, vol. 13, iss. 5, pp. 251–255. doi:10.21275/SR24502103758
27. Zhang Y., Du T., Ma Y., Wang X., Xie Y., Yang G., Lu Y., Chang E. C. AttacKG+: Boosting attack graph construction with Large Language Models. *Computers & Security*, 2024, vol. 150, pp. 104220. doi:10.1016/j.cose.2024.104220
28. Yamin M. M., Hashmi E., Ullah M., Katt B. Application of LLMs for generating cyber security exercise scenarios. *Research Square*, Feb. 2024. doi:10.21203/rs.3.rs-3970015/v1. Available at: https://www.researchsquare.com/article/rs-3970015/v1 (accessed 10 December 2024).
29. Wu Q., Bansal G., Zhang J., Wu Y., Li B., Zhu E., Jiang L., Zhang X., Zhang S., Liu J., Awadallah A. H., White R. W., Burger D., Wang C. AutoGen: Enabling next-gen LLM applications via multi-agent conversations. *First Conference on Language Modeling*, Pennsylvania, PA, 2024, pp. 1–46. Available at: https://openreview.net/pdf?id=BAakY1hNKS (accessed 10 December 2024).
30. Mavroeidis V., Bromander S. Cyber threat intelligence model: An evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence. *2017 European Intelligence and Security Informatics Conference (EISIC).* IEEE, 2017, pp. 91–98.
31. Zulkarneev R. Kh., Yusupova N. I., Smetanina O. N., Gayanova M. M., Vulfin A. M. Method and models of extraction of knowledge from medical documents. *Informatics and Automation*, 2022, vol. 21, no. 6, pp. 1169–1210 (In Russian). doi:10.15622/ia.21.6.4, EDN: ASOOVS
32. Rajapaksha S., Rani R., Karafili E. A RAG-based question-answering solution for cyber-attack investigation and attribution. *Workshop on Security and Artificial Intelligence 2024, 29th European Symposium on Research in Computer Security*, 2024. Available at: http://eprints.soton.ac.uk/id/eprint/493520 (accessed 10 December 2024).
33. Vasilyev V. I., Vulfin A. M., Kuchkarova N. V. Automation of software vulnerabilities analysis on the basis of text mining technology. *Cybersecurity Issues*, 2020, no. 4 (38), pp. 22–31 (In Russian). doi:10.21681/2311-3456-2020-04-22-31, EDN: TMRYIT
34. Topsakal O., Akinci T. C. Creating Large Language Model applications utilizing LangChain: A primer on developing LLM apps fast. *International Conference on Applied Engineering and Natural Sciences*, 2023, vol. 1, no. 1, pp. 1050–1056. doi:https://doi.org/10.59287/icaens.1127
35. Auger T., Saroyan E. *Overview of the OpenAI APIs*. In: *Generative AI for Web Development: Building Web Applications Powered by OpenAI APIs and Next.js*. Berkeley, CA, Apress, 2024, pp. 87–116.
36. Lin C. Y. ROUGE: A package for automatic evaluation of summaries. *Proceedings of the Workshop on Text Summarization Branches Out (WAS 2004)*, Barcelona, Spain, 2004, pp. 74–81.
37. Papineni K., Roukos S., Ward T., Zhu W. J. BLEU: A method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 2002, pp. 311–318. https://doi.org/10.3115/1073083.1073135
38. Es S., James J., Anke L. E., Schockaert S. RAGAs: Automated evaluation of Retrieval Augmented Generation. *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, 2024, pp. 150–158. doi:10.48550/arXiv.2309.15217

УДК 519.72
doi:10.31799/1684-8853-2025-2-68-77
EDN: MAHCAY

Научные статьи
Articles

Декодирование одиночных пакетов ошибок по минимуму длины пакета на основе информационных совокупностей

М. Н. Исаева[a], *старший преподаватель, orcid.org/0009-0007-6228-0617, imn@guap.ru*
[a]*Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ*

Введение: *для улучшения характеристик систем связи и хранения информации необходимо учитывать специфику шума, которая во многих практических ситуациях приводит к эффекту пакетирования ошибок. Декодирование независимых ошибок по информационным совокупностям является лучшим методом декодирования случайных линейных кодов, но имеет экспоненциальную сложность.* ***Цель:*** *анализ корректирующей способности и вычислительной сложности алгоритмов на основе информационных совокупностей для исправления одиночных пакетов ошибок.* ***Результаты:*** *введен критерий декодирования одиночных пакетов с использованием информационных совокупностей на основе минимума длины пакета. Доказано, что такой критерий обеспечивает исправление всех пакетов ошибок в пределах корректирующей способности линейного кода. Рассмотрены вопросы уменьшения количества необходимых для правильного декодирования информационных совокупностей, в целях чего исследованы методики построения множества информационных совокупностей ограниченного диаметра, называемых плотными. Проведен анализ, показывающий возможность снизить мощность множества информационных совокупностей с O(n) до O(1). Оценена вероятность построения множества и сформулирована оптимизационная задача по максимизации этой вероятности.* ***Практическая значимость:*** *разработанный декодер одиночных пакетов ошибок на основе критерия минимальной длины пакета, а также методика построения плотных информационных совокупностей мощностью O(1) гарантируют реализацию корректирующей способности случайных линейных кодов при исправлении одиночных пакетов ошибок.* ***Обсуждение:*** *полученные результаты подтверждают, что задача исправления однократных пакетов ошибок случайными линейными кодами является полиномиальной. Вместе с тем рассмотренные методы позволяют получать множества информационных совокупностей с достаточно высокой вероятностью лишь для кодов с не очень большой скоростью, а при использовании других, более практических классов кодов, вероятности нахождения множеств информационных совокупностей могут достаточно сильно измениться, что является направлением дальнейших исследований.*

Ключевые слова *— декодирование по информационным совокупностям, граница Рейгера, исправление пакетов ошибок, плотные информационные совокупности, критерии декодирования.*

Для цитирования: Исаева М. Н. Декодирование одиночных пакетов ошибок по минимуму длины пакета на основе информационных совокупностей. *Информационно-управляющие системы*, 2025, № 2, с. 68–77. doi:10.31799/1684-8853-2025-2-68-77, EDN: MAHCAY

For citation: Isaeva M. N. Decoding of single error bursts using minimal burst length criteria and information sets. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2025, no. 2, pp. 68–77 (In Russian). doi:10.31799/1684-8853-2025-2-68-77, EDN: MAHCAY

Введение

В современном мире информационные потоки внедрились во все сферы деятельности человека. Большую значимость имеет передача информации (например, в двоичном виде) по каналам связи, в которых могут возникать различные помехи, в том числе из-за особенностей среды передачи. Актуальной задачей является исправление ошибок, возникающих в передаваемых битовых последовательностях. Для сокращения числа ошибок используются помехоустойчивые коды, для которых необходимо разрабатывать специальные методы декодирования, позволявшие бы исправлять значительное число ошибок (обеспечивать малую вероятность ошибки декодирования) при условии ограниченных вычислительных и емкостных ресурсов.

В большинстве кодов при декодировании ошибки рассматриваются как независимые события, однако в реальности ошибки на выходе канала часто группируются в пакеты ошибок, например из-за особенностей среды передачи. Исправление таких пакетов является менее исследованной задачей в теории кодирования, чем исправление независимых ошибок, тем не менее известно, что учет особенностей векторов ошибок позволяет строить схемы кодирования и декодирования, теоретически позволяющие достичь лучших характеристик системы связи. В литературе, как правило, при решении задачи борьбы с ошибками в каналах с памятью применяется либо построение специальных кодов и традиционных декодеров [1–7], либо, достаточно редко — специальных декодеров для специфических моделей каналов связи и классических схем

кодирования [8–13]. В то же время кажется, что такая задача должна решаться совместно алгоритмами обработки информации на передатчике и приемнике.

В данной статье исследуется декодирование для исправления одиночных пакетов ошибок для самого общего класса кодов — линейных кодов, являющихся линейными векторными подпространствами и задающимися либо своим базисом (порождающей матрицей), либо базисом ортогонального пространства (проверочной матрицей). Случайные линейные коды обладают хорошей корректирующей способностью как в каналах с независимыми ошибками, так и в каналах с памятью, хотя их использование ограничено отсутствием структуры кода, которую можно использовать для повышения вычислительной эффективности процедур кодирования и декодирования.

Исправление одиночных пакетов ошибок линейными кодами изучалось, например, в статье [12] с помощью декодера со скользящим окном, однако сложность данного декодера для линейных кодов в общем случае оценивается авторами как экспоненциальная. В [14] рассматривается исправление пакетов ошибок на основе информационных совокупностей и исследуется способность некоторых кодов (с малыми параметрами) исправлять ошибки сверх корректирующей способности. В [15] анализируется экспонента сложности декодирования «почти» по максимуму правдоподобия при применении информационных совокупностей в каналах с двумя состояниями при наличии у декодера информации о состояниях канала или надежностях символов.

В данной статье предложена вероятностная методика построения множеств с малым числом совокупностей при максимизации вероятности построения множества для случайных линейных кодов.

Декодирование по информационным совокупностям для исправления пакетов ошибок

Линейный двоичный (n, k)-код — это k-мерное подпространство n-мерного линейного векторного пространства двоичных векторов. Любой базис $\mathbf{G}$ этого пространства является $(k \times n)$-матрицей ранга k и называется порождающей матрицей кода. Матрица $\mathbf{H}$ размерности $(r \times n)$, где $r = n - k$, для которой $\mathbf{GH}^{\mathrm{T}} = \mathbf{0}$, является базисом ортогонального пространства размерности r и называется проверочной матрицей кода. Сообщение $\mathbf{m}$, представленное вектором длиной k, кодируется в кодовое слово $\mathbf{a}$ с помощью преобразования $\mathbf{mG} = \mathbf{a}$. Вектор $\mathbf{S} = \mathbf{bH}^{\mathrm{T}}$ размерности r называется синдромом вектора $\mathbf{b}$ и является нулевым вектором тогда и только тогда, когда $\mathbf{b}$ — кодовое слово [16–18].

Способность кода исправлять независимые битовые ошибки (корректирующую способность кода) обычно связывают с параметром d_0 — минимальным расстоянием кода, равным минимуму из расстояний Хемминга между всеми парами кодовых слов. Код с расстоянием $d_0 = 2t + 1$ гарантирует исправление любой комбинации из не более чем t битовых ошибок, если декодирование заключается в поиске кодового слова, ближайшего к принятому из канала, где расстояние считается в метрике Хемминга. В действительности код может исправлять и значительное число комбинаций из более чем t ошибок, полный их список можно получить, если рассмотреть разбиение n-мерного векторного пространства по коду на смежные классы [14, 16]. Векторы одного смежного класса имеют одинаковый синдром (например, у всех кодовых слов он нулевой), и если выбрать из каждого смежного класса одного представителя, называемого лидером, и считать его возможным вектором ошибки, можно установить взаимно однозначное соответствие между синдромами и лидерами, что дает список ошибок, исправляемых кодом, и процедуру декодирования, использующую установленное соответствие, называемую синдромным декодированием и имеющую экспоненциальную сложность. Обычно в качестве лидеров смежных классов выбираются наиболее вероятные векторы в канале связи.

Стоит заметить, что как декодирование по минимуму расстояния (эквивалентное декодированию лидеров смежных классов), так и определение минимального расстояния кода являются NP-трудными задачами.

Если в векторе $\mathbf{e}$ длиной n первая ненулевая компонента расположена на позиции i_1, а последняя ненулевая компонента — на позиции i_2, то говорят, что вектор $\mathbf{e}$ образует однократный пакет длиной $b = i_2 - i_1 + 1$. Иногда может возникнуть неоднозначность в трактовке термина «пакет»: он может подразумевать как сам вектор $\mathbf{e}$, так и подвектор на позициях от i_1 до i_2 (так называемых позициях пакета). Конкретный смысл этого термина обычно ясен из контекста.

Таким образом, все ненулевые компоненты вектора $\mathbf{e}$ находятся на позициях внутри интервала от i_1 до i_2 или на позициях пакета. Иногда пакетом длиной b считается также вектор $\mathbf{e}$, все ненулевые компоненты которого расположены на позициях пакета, но сами позиции i_1 и (или) i_2 могут содержать нулевые элементы, мы будем в этом случае говорить, что $\mathbf{e}$ образует пакет длиной, не превышающей b.

Для определенности в дальнейшем будем считать, что позиции векторов нумеруются, начиная с единицы. Назовем позицию i_1 началом пакета, тогда, если мы рассматриваем пакеты, длина которых в точности равна b, начало пакета может принимать значения от 1 до $n - b + 1$.

В некоторых случаях, особенно при рассмотрении циклических кодов, вводится понятие циклического пакета, когда отсчет позиций пакета может происходить от i_2 до i_1 с переходом от позиции n снова к единице, однако в данной статье мы не рассматриваем циклические коды и циклические пакеты.

По аналогии с параметром t можно рассмотреть величину b_0: если все пакеты, образующие пакеты длиной $b \le b_0$, находятся в лидерах смежных классов, то любой пакет длиной, не превышающей b_0, может быть исправлен. Исправление пакетов сверх корректирующей способности кода b_0 рассматривалось в работе [14] для кодов малых длин, где возможно построение списка лидеров смежных классов. Для линейного (n, k)-кода в соответствии с границей Рейгера справедливо $b_0 \le (n - k)/2$.

Важно, что в отличие от параметра t, параметр b_0 для любого линейного кода может быть определен с полиномиальной сложностью [19], это дает основание надеяться, что и исправление одиночных пакетов может оказаться полиномиальной задачей, однако в данной статье мы ограничиваемся рассмотрением только исправления пакетов длиной до b_0, предполагая, что для любого кода это значение известно.

Если линейный (n, k)-код не обладает никакой дополнительной структурой, то лучшим подходом к исправлению ошибок является декодирование по информационным совокупностям [14, 20]. Если ошибки независимы, количество требуемых информационных совокупностей экспоненциально, однако при исправлении пакетов их количество может быть значительно снижено [14].

Пусть $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ — кодовое слово. Если $\chi = \{1 \le j_1 < j_2 < \ldots < j_k \le n\}$ — множество чисел, состоящее из сочетания (без повторений) k чисел от 1 до n, обозначим через $\mathbf{a}(\chi)$ подвектор $\mathbf{a}$, составленный из элементов на позициях с номерами из χ. Аналогично, если $\mathbf{A}$ — $(k \times n)$-матрица, обозначим через $\mathbf{A}(\chi)$ $(k \times k)$-подматрицу $\mathbf{A}$ из столбцов с номерами из χ. Множество χ называется информационной совокупностью, если любое кодовое слово $\mathbf{a}$ может быть однозначно восстановлено, $\mathbf{a}(\chi)$.

Является ли некоторое множество чисел χ информационной совокупностью, зависит не от конкретного кодового слова $\mathbf{a}$, а от кода. Определить это можно с помощью порождающей матрицы $\mathbf{G}$: множество χ является информационной совокупностью тогда и только тогда, когда $\operatorname{rank} \mathbf{G}(\chi) = k$. Также можно использовать проверочную матрицу $\mathbf{H}$, в этом случае должно выполняться условие: $\operatorname{rank} \mathbf{H}(\{1, \ldots, n\} \backslash \chi) = r$, где дополнение $\{1, \ldots, n\} \backslash \chi$ — множество чисел от 1 до n, не вошедших в χ.

Таким образом, если χ — информационная совокупность, матрица $\mathbf{G}(\chi)$ является невырожденной и к ней существует обратная (над полем GF(2)). Пусть $\mathbf{G}_\chi = \mathbf{G}^{-1}(\chi)\mathbf{G}$, тогда $\mathbf{G}_\chi$ также является базисом того же пространства, которое задается базисом $\mathbf{G}$, и, следовательно, $\mathbf{G}_\chi$ — еще одна порождающая матрица того же кода. На позициях χ матрица $\mathbf{G}_\chi$ содержит единичную подматрицу, следовательно, произведение $\mathbf{a}(\chi)\mathbf{G}_\chi$ задаст кодовое слово $\mathbf{a}$ с элементами $\mathbf{a}(\chi)$ на позициях χ. Декодирование на основе информационных совокупностей построено на поиске такого множества χ, которое не искажено ошибками, тогда кодовое слово $\mathbf{a}$ будет восстановлено правильно. Псевдокод алгоритма исправления независимых ошибок при декодировании по информационным совокупностям приведен на рис. 1.

Предполагается, что для использования алгоритма на рис. 1 необходимо сгенерировать множество информационных совокупностей X мощностью N, а также соответствующее множество порождающих матриц $\{\mathbf{G}_{\chi_1}, \ldots, \mathbf{G}_{\chi_N}\}$. К сожалению, при исправлении независимых ошибок для обеспечения малых вероятностей ошибки декодирования число N должно быть экспоненциально велико.

Версия алгоритма, приведенная на рис. 1, не определяет непосредственно отсутствие ошибок на информационной совокупности, как в некоторых других вариациях, так как это требует знания значения t, которое в общем случае для случайных линейных кодов неизвестно, и вместо этого вычисляет функцию расстояния между векторами $d(\mathbf{a}, \mathbf{b})$. Из рассмотренных кодовых слов выбирается ближайшее к принятому слову. Для независимых битовых ошибок используется расстояние Хемминга. Данный алгоритм мож-

Вход: принятое из канала слово $\mathbf{b}$, множество $X = \{\chi_1, \ldots, \chi_N\}$, множество $\{\mathbf{G}_{\chi_1}, \ldots, \mathbf{G}_{\chi_N}\}$.

Выход: решение декодера о кодовом слове $\hat{\mathbf{a}}$.

1. $d_{\min} \leftarrow \infty$.
2. Для $i = 1$ до N:
   - 2.1. $\mathbf{a} \leftarrow \mathbf{b}(\chi_i)\mathbf{G}_{\chi_i}$.
   - 2.2. Если $d(\mathbf{a}, \mathbf{b}) < d_{\min}$:
     - 2.2.1. $d_{\min} \leftarrow d(\mathbf{a}, \mathbf{b})$.
     - 2.2.2. $\hat{\mathbf{a}} \leftarrow \mathbf{a}$.
3. Возвратить $\hat{\mathbf{a}}$.

■ ***Рис. 1.*** Декодирование по информационным совокупностям

■ ***Fig. 1.*** Information set decoding

но приспособить для исправления одиночных пакетов ошибок, однако необходимо учитывать специфику шума при построении множества X, а также при вычислении функции расстояния.

Для двух векторов **a** и **b** длиной n определим

$$d_b(\mathbf{a}, \mathbf{b}) = W_b(\mathbf{a} \oplus \mathbf{b}) = W_b(\mathbf{e}), \qquad (1)$$

где $W_b(\mathbf{e})$ — функция, возвращающая длину пакета, который образует вектор **e**. Тогда если **a** — передававшееся кодовое слово, а **b** — слово, принятое из канала, то $\mathbf{e} = \mathbf{a} \oplus \mathbf{b}$ — вектор ошибки. В этом случае алгоритм на рис. 1 будет принимать решение в пользу кодового слова, которому в канале связи соответствует пакет ошибок меньшей длины. Такой критерий декодирования ориентирован на обеспечение низких вероятностей ошибки в каналах, в которых более длинные пакеты ошибок менее вероятны, это согласуется с характеристиками многих реальных каналов связи.

При построении множества X и определении его объема N необходимо учитывать возможность нахождения информационной совокупности, не затронутой ошибками. В случае, когда ошибки образуют пакеты длиной b_0 (или менее) и они имеют возможные начала от 1 до $n - b_0 + 1$, нетрудно указать способ построения X, при котором гарантировано нахождение информационной совокупности, лежащей за пределами пакета. Для этого достаточно определить информационную совокупность, не включающую позиции пакета, тогда для всех возможных начал пакета $N = n - b_0 + 1$ и количество используемых информационных совокупностей линейно зависит от длины кода n.

Тем не менее возможно дальнейшее уменьшение величины N, что будет показано ниже.

Заметим, что функция расстояния, задаваемая (1), не является метрикой, как, например, расстояние Хемминга, и это не позволяет связать корректирующую способность b_0 со значениями длин пакетов $W_b(\mathbf{a})$, образуемых кодовыми словами, и утверждать о способности предложенного алгоритма реализовывать корректирующую способность кода при исправлении пакетов ошибок, что рассмотрим в следующем разделе.

Правило декодирования для исправления пакетов информационными совокупностями

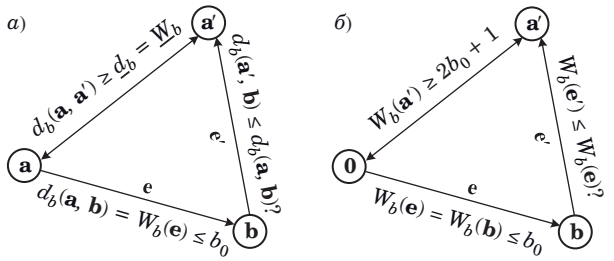
При исправлении независимых ошибок с помощью поиска кодового слова, ближайшего к принятому, в метрике Хемминга декодирование будет гарантированно правильным, если число произошедших ошибок не превышает корректирующей способности t кода. Это используется и в алгоритме, приведенном на рис. 1: никакая информационная совокупность, искаженная не более чем t ошибками, не даст кодового слова, более близкого к принятому, чем правильное, и ошибка декодирования в данном алгоритме возможна лишь в случае, когда среди N информационных совокупностей не нашлось не затронутой ошибками.

Однако функция расстояния по длине пакета, определенная в (1), не является метрикой, в частности, для нее не выполняется неравенство треугольника, следовательно, нельзя напрямую связать корректирующую способность кода и декодирование по правилу (1). В данном разделе подробнее рассмотрим использование правила (1) в алгоритме на рис. 1.

В качестве иллюстрации приведем следующую ситуацию. Пусть C — (n, k)-код, способный исправлять любые векторы ошибки, представляющие собой пакеты длиной до b_0. Это означает, что все такие векторы ошибки лежат в разных смежных классах. Пусть передается кодовое слово **a**, и в канале произошла ошибка, описываемая вектором **e** и представляющая собой пакет длиной не более b_0. Пусть из канала принято слово $\mathbf{b} = \mathbf{a} \oplus \mathbf{e}$. Рассмотрим возможность декодирования вектора **b** в неправильное кодовое слово $\mathbf{a}'$, при котором используется критерий минимума расстояния по правилу (1). Это означало бы, что декодер нашел вектор ошибки $\mathbf{e}' = \mathbf{a}' \oplus \mathbf{b}$ такой, что ошибочное кодовое слово $\mathbf{a}'$ ближе (или, возможно, не дальше) к принятому, чем правильное слово **a**, или $d_b(\mathbf{a}', \mathbf{b}) \le d_b(\mathbf{a}, \mathbf{b})$. Данная ситуация изображена на рис. 2, *а*.

Рассмотрим несколько простых соотношений для функции расстояния (1), многие из которых полностью эквивалентны таковым для метрики Хемминга. Для двоичных векторов **a**, **b**, **c** справедливо

$$d_b(\mathbf{a} \oplus \mathbf{c}, \mathbf{b} \oplus \mathbf{c}) = W_b(\mathbf{a} \oplus \mathbf{b}) = d_b(\mathbf{a}, \mathbf{b}).$$



■ ***Рис. 2.*** Оценка возможности ошибки при декодировании по критерию минимальной длины пакета в общем случае (*а*) и при передаче нулевого слова (*б*)

■ ***Fig. 2.*** Estimation of error possibility using minimal burst length decoding criteria in general case (*а*) and in case of zero codeword (*б*)

Обозначим через W_b минимальную длину пакета, которую образуют ненулевые кодовые слова кода C:

$$\underline{W}_b = \min_{\mathbf{a}\in C, \mathbf{a}\neq \mathbf{0}} \{W_b(\mathbf{a})\}.$$

Обозначим через d_b минимальное попарное расстояние среди всех кодовых слов:

$$\underline{d}_b = \min_{\substack{\mathbf{a}_i, \mathbf{a}_j \in C, \\ i\neq j}} \{d_b(\mathbf{a}_i, \mathbf{a}_j)\}.$$

Если $\mathbf{a}_i$, $\mathbf{a}_j$ — ближайшие кодовые слова, то

$$d_b = d_b(\mathbf{a}_i, \mathbf{a}_j) = W_b(\mathbf{a}_i \oplus \mathbf{a}_j) \geq W_b,$$

так как $\mathbf{a}_i \oplus \mathbf{a}_j$ также является кодовым словом. С другой стороны, для кодового слова $\mathbf{a}$ минимальной пакетной длины

$$W_b = W_b(\mathbf{a}) = d_b(\mathbf{0}, \mathbf{a}) \geq d_b,$$

следовательно: $W_b = d_b$.

Величина d_b имеет некоторую, хотя и не полную, аналогию с минимальным расстоянием d_0 кода (в метрике Хемминга), а W_b — с минимальным весом. Из теории кодирования известно, что, если код исправляет пакеты длиной b_0 и меньше, в коде нет ненулевых кодовых слов, образующих пакет длиной $2b_0$ или меньше, таким образом:

$$W_b \geq 2b_0 + 1. \qquad (2)$$

Отличие от случая независимых ошибок вызвано тем, что расстояние (1) не является метрикой и состоит в том, что минимальное расстояние кода в метрике Хемминга и число независимых исправляемых ошибок связаны равенством $d_0 = 2t + 1$, тогда как в (2) присутствует неравенство. Таким образом, по величине b_0 нельзя точно определить W_b и наоборот.

Пусть $\mathbf{a}$ — вектор длиной n. Представим его как конкатенацию двух векторов: $\mathbf{a} = (\mathbf{a}_1|\mathbf{a}_2)$, и образуем два вектора $\mathbf{e}_1 = (\mathbf{a}_1|\mathbf{0})$ и $\mathbf{e}_2 = (\mathbf{0}|\mathbf{a}_2)$ длиной n каждый. Назовем множество $\{\mathbf{e}_1, \mathbf{e}_2\}$ декомпозицией вектора $\mathbf{a}$. Для дальнейшего рассмотрения докажем следующую лемму.

Лемма. Пусть $\mathbf{a}$ — кодовое слово. Тогда для любой его декомпозиции $\{\mathbf{e}_1, \mathbf{e}_2\}$ векторы $\mathbf{e}_1$ и $\mathbf{e}_2$ имеют одинаковый синдром.

Доказательство: Пусть $\mathbf{H}$ — проверочная матрица кода, тогда $\mathbf{aH}^{\mathrm{T}} = \mathbf{0}$. Пусть $\mathbf{S}_1 = \mathbf{e}_1\mathbf{H}^{\mathrm{T}}$ и $\mathbf{S}_2 = \mathbf{e}_2\mathbf{H}^{\mathrm{T}}$, тогда $\mathbf{S}_1 \oplus \mathbf{S}_2 = (\mathbf{e}_1 \oplus \mathbf{e}_2)\mathbf{H}^{\mathrm{T}} = \mathbf{aH}^{\mathrm{T}} = \mathbf{0}$ или $\mathbf{S}_1 = \mathbf{S}_2$, что и требовалось доказать.

Из данной леммы прямо следует выражение (2). Действительно, если ненулевое кодовое слово $\mathbf{a}$ образует пакет длиной $2b_0$ или меньше, его декомпозиция, проведенная по средней позиции пакета, образует два пакета длиной, не превышающей b_0, каждый, они будут иметь одинаковый синдром и находиться в одном смежном классе, что противоречит утверждению о том, что код исправляет пакеты длиной b_0.

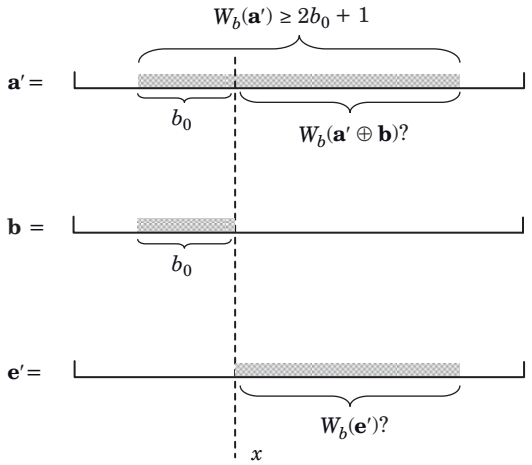
Подытоживая вышесказанное, прибавив (покомпонентно по модулю 2) ко всем векторам, изображенным на рис. 2, *а*, вектор $\mathbf{a}$, мы получим более простой для рассмотрения случай (рис. 2, *б*), когда передается нулевое слово, а векторы $\mathbf{b}$ и $\mathbf{a}'$ заново переобозначены. При этом сохраняются все попарные расстояния между рассматриваемыми векторами, принятое слово $\mathbf{b}$ образует пакет длиной не более b_0, а кодовое слово $\mathbf{a}'$ образует пакет длиной не менее $2b_0 + 1$. Сформулируем и докажем следующую теорему.

Теорема. Пусть код C исправляет все пакеты длиной до b_0. При декодировании по минимуму расстояния с использованием функции расстояния (1) любой пакет ошибок длиной до b_0 будет корректно исправлен.

Доказательство: Чтобы доказать теорему, необходимо показать, что для ситуации, изображенной на рис. 2, *б*, невозможно выполнение неравенства $d_b(\mathbf{a}', \mathbf{b}) \leq d_b(\mathbf{a}, \mathbf{b})$ и, следовательно, невозможно выполнение $W_b(\mathbf{e}') \leq W_b(\mathbf{e})$.

Рассмотрим вектор $\mathbf{a}'$ и условия, необходимые для выполнения указанных неравенств. Вектор $\mathbf{a}'$ представляет собой одиночный пакет длиной не менее $2b_0 + 1$ (рис. 3).

Вектор $\mathbf{b}$ представляет собой пакет длиной не более b_0, и нам необходимо определить, при каких условиях длина пакета вектора $\mathbf{e}' = \mathbf{a}' \oplus \mathbf{b}$ может оказаться меньше или равной длине пакета вектора $\mathbf{e}$. Другими словами, нужно определить возможность уменьшения пакета



■ ***Рис. 3.*** Уменьшение длины пакета при сложении векторов

■ ***Fig. 3.*** Decreasing of burst length from vectors addition

в векторе $\mathbf{a}'$ путем прибавления к нему вектора $\mathbf{b}$ (см. рис. 3).

Будем считать, что любой пакет начинается и заканчивается единицей, и нетрудно видеть, что максимальное уменьшение длины пакета может происходить, если пакет в векторе $\mathbf{b}$ полностью совпадает либо с левой частью пакета в векторе $\mathbf{a}'$ (что изображено на рис. 3), либо с правой, при этом длина пакета вектора $\mathbf{b}$ должна быть как можно большей, поэтому положим ее равной b_0.

Рассмотрим сокращение пакета в векторе $\mathbf{a}'$ на b_0 позиций слева. Случай сокращения справа полностью аналогичен. Предположим, что в векторе $\mathbf{b}$ последняя позиция пакета равна $x - 1$, тогда пакет в векторе $\mathbf{e}'$ может начинаться с позиции x (это обозначено пунктирной линией на рис. 3). Неопределенность состоит в том, что элемент на позиции x в векторе $\mathbf{e}'$ может оказаться нулевым, как и последующие за ним, кроме последней позиции пакета, и длина пакета, таким образом, может оказаться сколь угодно маленькой, вплоть до единичной.

Воспользуемся доказанной леммой. Проведем декомпозицию вектора $\mathbf{a}'$ по позиции x, получив векторы $\mathbf{e}_1$ и $\mathbf{e}_2$, причем $\mathbf{e}_1 = \mathbf{e}$, $\mathbf{e}_2 = \mathbf{e}'$. Очевидно: $W_b(\mathbf{e}_1) = b_0$. По лемме, вектор $\mathbf{e}_2$ имеет тот же синдром, что и вектор $\mathbf{e}_1$, следовательно, находится с ним в одном смежном классе. Так как код исправляет пакеты длиной b_0, в смежном классе может быть только один вектор длиной, не превышающей b_0, следовательно: $W_b(\mathbf{e}_2) = W_b(\mathbf{e}') \geq b_0 + 1$ и $W_b(\mathbf{e}') > W_b(\mathbf{e})$, что и требовалось доказать.

Следствие. Если длина пакета ошибок не превышает корректирующей способности кода b_0, декодирование по информационным совокупностям, использующее функцию расстояния (1), гарантированно исправит пакет ошибок, если в множестве X, заданном алгоритму на рис. 1, окажется хотя бы одна информационная совокупность, не содержащая позиций пакета. Это можно обеспечить, включив в X хотя бы одну такую информационную совокупность для каждой из $n - b_0 + 1$ начальных позиций пакета.

Анализ методик построения информационных совокупностей

В предыдущем разделе мы доказали, что алгоритм декодирования одиночных пакетов по информационным совокупностям, использующий правило (1), гарантированно обеспечивает декодирование в пределах корректирующей способности кода b_0. При этом множество информационных совокупностей X должно содержать $n - b_0 + 1 = O(n)$ информационных совокупностей. Однако, хотя в общем случае нахождение информационной совокупности на случайной позиции имеет достаточно большую вероятность, при ограничении выбираемых позиций и при снижении числа информационных совокупностей необходимо оценивать вероятность построения того или иного множества с учетом ограничений и конструкции кода.

В этом разделе рассмотрим подходы к дальнейшему уменьшению количества информационных совокупностей и оценке вероятности нахождения множества X для случайных линейных кодов.

Обратим внимание, что информационная совокупность, образованная, к примеру, числами $\{1, \ldots, k\}$, будет свободной от ошибок для любого пакета, начинающегося с позиции $k + 1$. Таким образом, останется найти только информационные совокупности для пакетов, начало которых приходится на первые k позиций. Такой подход может позволить снизить количество информационных совокупностей до некоторого значения $T \leq N = n - b_0 + 1$.

Конечно, нет никаких гарантий, что k последовательно идущих чисел образуют информационную совокупность, поэтому введем параметр $\Delta \in [0, n - k - b_0]$ и будем искать информационные совокупности в интервале длиной $k + \Delta$. Значение $k + \Delta$ назовем диаметром информационной совокупности. Ограниченные таким образом совокупности назовем плотными [14]. Нетрудно показать, что при $\Delta = n - k - b_0$ интервалы поиска информационных совокупностей будут ограничены только длинами пакетов, что совпадает с исходной процедурой поиска.

Вероятность нахождения информационной совокупности на интервале $k + \Delta$ соответствует вероятности нахождения невырожденной матрицы $\mathbf{M}$ размера $k \times k$. Известно [21], что для случайной $(k \times k + \Delta)$-матрицы эта вероятность может быть примерно оценена как

$$\Pr(\mathrm{rank}(\mathbf{M}_{k,k+\Delta}) = k) \approx Q_\Delta = \prod_{j=\Delta+1}^{\infty}\left(1 - \frac{1}{2^j}\right) \quad (3)$$

для всех натуральных k, имеющих практический интерес (больше нескольких единиц).

В частности, при $\Delta = 0$ вероятность выбора случайной невырожденной двоичной $(k \times k)$-матрицы составляет примерно 0,28 вне зависимости от значения k. Предположим, что нахождение информационных совокупностей представляет собой последовательность независимых событий, тогда из (3) следует, что вероятность нахождения множества из T информационных совокупностей будет иметь следующий вид:

$$Q_{\Delta_T} = Q_\Delta^T = \left(\prod_{j=\Delta+1}^{\infty}\left(1 - \frac{1}{2^j}\right)\right)^T. \quad (4)$$

Сопоставим некоторой плотной информационной совокупности множество начал пакетов, для которых пакет не затрагивает данную информационную совокупность. Назовем позиции из такого множества защищенными (данной информационной совокупностью). Тогда процедура построения множества X информационных совокупностей должна продолжаться до тех пор, пока защищенными не станут все возможные начала пакетов, т. е. $\{1, \ldots, n - b_0 + 1\}$.

Рассмотрим методику построения, при которой начала плотных информационных совокупностей для случайного линейного кода выбираются циклически с шагом $k + \Delta + b_0$. Если на очередном интервале длиной $k + \Delta$ информационная совокупность не найдена, будем считать, что происходит отказ процедуры поиска. На практике в этом случае можно либо сгенерировать другой случайный код, либо расширять значение Δ.

Число плотных информационных совокупностей, построенных по такой методике при заданных Δ и b_0, можно оценить с помощью выражения

$$T \geq \left\lceil \frac{n - b_0 + 1}{n - b_0 + 1 - (k + \Delta)} \right\rceil. \qquad (5)$$

Если считать, что используются коды, корректирующая способность которых соответствует границе Рейгера ($b_0 = (n - k)/2$), то выражение (5) может быть записано как

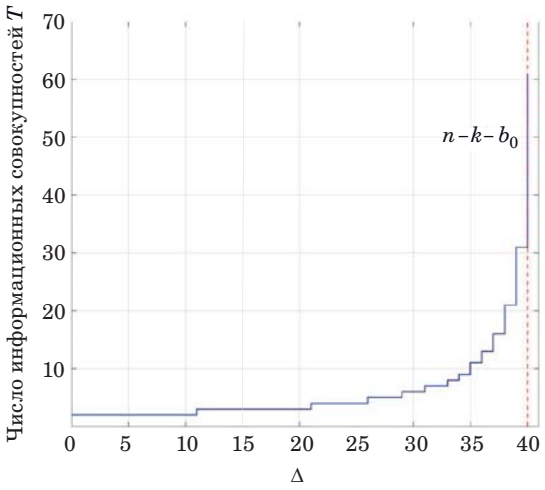
$$T \geq \left\lceil \frac{n + nR + 2}{n + nR + 2 - 2(nR + \Delta)} \right\rceil, \qquad (6)$$

где $R = k/n$ — скорость кода.

Как нетрудно убедиться, при фиксированных скорости кода и параметре Δ число T плотных информационных совокупностей оценивается как $O(1)$ от длины кода n. При фиксированных n и Δ величина T растет с ростом R. Наконец, используя выражения (4) и (6), рассмотрим зависимость величины T от значения Δ при фиксированных n и R, например, для кода со скоростью $R = 0{,}2$, длиной $n = 100$, $b_0 = 40$ (рис. 4).

График зависимости имеет ступенчатый вид, и можно наблюдать достаточно большие диапазоны значений Δ, которым соответствует одинаковое число информационных совокупностей T. Рассмотрим значения функции Q_Δ^T (4) — вероятности нахождения множества плотных информационных совокупностей — при фиксированном T. Функция Q_Δ^T из выражения (4) — возрастающая, рассмотрим ее значения, например при $T = 2$, в зависимости от значений Δ (таблица).

Как можно видеть, при малых Δ вероятность построения множества X плотных информационных совокупностей не очень велика, однако с ростом Δ она приближается к единице.



■ ***Рис. 4.*** Зависимость числа плотных информационных совокупностей T от значения Δ

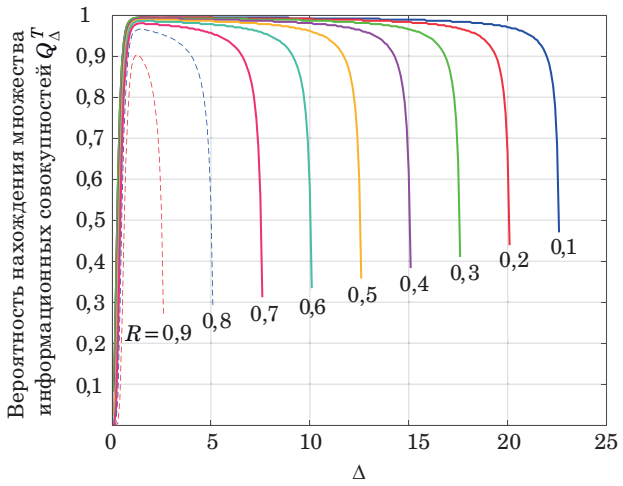
■ ***Fig. 4.*** Dependence of the number T of dense information sets from the Δ value

■ Значения функции Q_Δ^T для $T = 2$ и кода: $n = 100$, $R = 0{,}2$

■ Function values Q_Δ^T for $T = 2$ and code with parameters: $n = 100$, $R = 0.2$

Δ	Q_Δ^2	Δ	Q_Δ^2
0	0,08	6	0,96
1	0,33	7	0,98
2	0,6	8	0,99
3	0,87	9	0,99
4	0,88	10	0,99
5	0,93		

Следовательно, при построении множества плотных информационных совокупностей можно выбирать то значение Δ, при котором вероятность Q_Δ^T максимальна, т. е. максимальное для заданного числа T. Рассмотрим теперь зависимость вероятности нахождения множества информационных совокупностей от значения Δ для разных скоростей кодов с длиной $n = 100$ (рис. 5). Приведенные зависимости не являются монотонными, и существует некоторое оптимальное значение Δ, максимизирующее значение Q_Δ^T (его довольно нетрудно найти по приведенным формулам). Заметим, что с увеличением скорости кода это оптимальное значение заметно снижается, это вызвано тем, что для высокоскоростных кодов значение T достаточно велико.



■ ***Рис. 5.*** Зависимость вероятности нахождения множества информационных совокупностей Q_Δ^T от значения Δ при разных скоростях R кода

■ ***Fig. 5.*** Dependence of dense information set probability Q_Δ^T from values of Δ for different code rates R

Заключение

В статье рассмотрена задача исправления одиночных пакетов случайными линейными кодами с помощью подхода на основе информационных совокупностей. Предложен вариант декодирования по функции расстояния, соответствующей длине пакета. Доказано, что декодирование по минимуму длины пакета гарантирует исправление всех пакетов в пределах корректирующей способности кода. На основании этого доказано, что при использовании множества информационных совокупностей, покрывающих все начала пакетов, декодирование по информационным совокупностям и минимуму длины пакета реализует корректирующую способность кода. При этом число необходимых информационных совокупностей оценивается как $O(n)$.

Рассмотрена методика построения множества плотных информационных совокупностей, позволяющая уменьшить их число до $O(1)$. Проведен анализ вероятности построения множества плотных информационных совокупностей в зависимости от их диаметра и предложен способ выбора диаметра, на основе которого сформулирована оптимизационная задача, позволяющая максимизировать вероятность построения множества информационных совокупностей.

Все вероятностные расчеты проведены в предположении, что рассматривается случайный линейный двоичный код. Возможным направлением дальнейших исследований является распространение подходов, рассмотренных в статье, на конкретные классы помехоустойчивых кодов, которые обладают достаточно хорошей корректирующей способностью для исправления одиночных пакетов ошибок.

Финансовая поддержка

Работа выполнена при финансовой поддержке Российского научного фонда, грант № 22-19-00305 «Пространственно-временные стохастические модели беспроводных сетей с большим числом абонентов».

Литература

1. **Yang M., Pan Z., Djordjevic I. B.** FPGA-based burst-error performance analysis and optimization of regular and irregular SD-LDPC codes for 50G-PON and beyond. *Opt. Express*, 2023, vol. 31, no. 6, pp. 10936–10946. doi:10.1364/OE.477546
2. **Song L., Huang Q., Wang Z.** Construction of multiple-burst-correction codes in transform domain and its relation to LDPC codes. *IEEE Trans. Commun.*, 2020, vol. 68, no. 1, pp. 40–54. doi:10.1109/TCOMM.2019.2948341
3. **Vafi S.** Cyclic low density parity check codes with the optimum burst error correcting capability. *IEEE Access*, 2020, vol. 8, pp. 192065–192072. doi:10.1109/ACCESS.2020.3032837
4. **Xiao X., Vasic B., Lin S., Li J., Abdel-Ghaffar K.** Quasi-cyclic LDPC codes with parity-check matrices of column weight two or more for correcting phased bursts of erasures. *IEEE Trans. Commun.*, 2021, vol. 69, no. 5, pp. 2812–2823. doi:10.1109/TCOMM.2021.3059001
5. **Aharoni Z., Huleihel B., Pfister H. D., Permuter H. H.** Data-driven polar codes for unknown channels with and without memory. *2023 IEEE International Symposium on Information Theory (ISIT)*, Taipei, IEEE, 2023, pp. 1890–1895. doi:10.1109/ISIT54713.2023.10206663
6. **Sasoglu E., Tal I.** Polar coding for processes with memory. *IEEE Trans. Inform. Theory*, 2019, vol. 65, no. 4, pp. 1994–2003. doi:10.1109/TIT.2018.2885797
7. **Vajha M., Ramkumar V., Jhamtani M., Kumar P. V.** On the performance analysis of streaming codes over the Gilbert — Elliott channel. *2021 IEEE Information Theory Workshop (ITW)*, Kanazawa, Japan, 2021, pp. 1–6. doi:10.1109/ITW48936.2021.9611448
8. **Eckford A., Kschischang F., Pasupathy S.** Analysis of low-density parity-check codes for the Gilbert — Elliott channel. *IEEE Transactions on Information Theory*, 2005, vol. 51, no. 11, pp. 3872–3889. doi:10.1109/TIT.2005.856934
9. **Li L., Lv J., Li Y., Dai X., Wang X.** Burst error identification method for LDPC coded systems. *IEEE Com-*

mun. Lett., 2024, pp. 1–5. doi:10.1109/LCOMM.2024.3391826

10. **Fang Y., Chen J.** Decoding polar codes for a generalized Gilbert — Elliott channel with unknown parameter. *IEEE Trans. Commun.*, 2021, vol. 69, no. 10, pp. 6455–6468. doi:10.1109/TCOMM.2021.3095195
11. **Трофимов А. Н., Таубин Ф. А.** Улучшенная граница вероятности ошибки при оптимальном приеме в канале с межсимвольной интерференцией. *Информационно-управляющие системы*, 2023, № 5, с. 33–42. doi:10.31799/1684-8853-2023-5-33-42, EDN: MDHOXU
12. **Ovchinnikov A. A., Veresova A. M., Fominykh A. A.** Decoding of linear codes for single error bursts correction based on the determination of certain events. *Информационно-управляющие системы*, 2022, № 6, с. 41–52. doi:10.31799/1684-8853-2022-6-41-52, EDN: UWXZHN
13. **Veresova A., Fominykh A., Ovchinnikov A.** About usage of metrics in decoding of LDPC codes in two-state channels with memory. *2021 XVII International Symposium Problems of Redundancy in Information and Control Systems (REDUNDANCY)*, Moscow, IEEE, 2021, pp. 143–148. doi:10.1109/REDUNDANCY52534.2021.9606474
14. **Исаева М. Н., Овчинников А. А.** Исправление одиночных пакетов ошибок за пределами корректирующей способности кода с использованием информационных совокупностей. *Научно-технический вестник информационных технологий, механики и оптики*, 2024, вып. 24, № 1, с. 70–80. doi:10.17586/2226-1494-2024-24-1-70-80
15. **Sung W., Coffey J. T.** Information set decoding complexity for linear codes in bursty channels with side information. *Proceedings of 1995 IEEE International Symposium on Information Theory*, Whistler, BC, Canada, 1995, pp. 53. doi:10.1109/ISIT.1995.531155
16. **Moon T. K.** *Error Correction Coding: Mathematical Methods and Algorithms*. Second ed. Hoboken, NJ, Wiley, 2021. 992 p.
17. **Gazi O.** *Forward Error Correction Via Channel Coding*. Cham, Springer, 2020. 319 p.
18. **Lin S., Li J.** *Fundamentals of Classical and Modern Error-Correcting Codes*. Cambridge, Cambridge University Press, 2022. 840 p. doi:10.1017/9781009067928
19. **Veresova A. M., Isaeva M. N., Ovchinnikov A. A.** Estimation of independent errors and bursts correction capability of linear codes. *2024 Conference of Young Researchers in Electrical and Electronic Engineering (ElCon)*, Saint-Petersburg, IEEE, 2024, pp. 23–27. doi:10.1109/ElCon61730.2024.10468456
20. **Исаева М. Н.** Поиск информационных совокупностей при исправлении пакетов ошибок квазициклическими кодами. *T-Comm: Телекоммуникации и транспорт*, 2023, т. 17, № 7, с. 4–12. doi:10.36724/2072-8735-2023-17-7-4-12
21. **Berlekamp E. R.** The technology of error correcting codes. *Proc. IEEE*, 1980, vol. 68, pp. 564–593.

UDC 519.72
doi:10.31799/1684-8853-2025-2-68-77
EDN: MAHCAY

Decoding of single error bursts using minimal burst length criteria and information sets

M. N. Isaeva[a], Senior Lecturer, orcid.org/0009-0007-6228-0617, imn@guap.ru
[a]Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation

Introduction: To improve the characteristics of data transmission and storage systems the specificity of noise should be taken into account, which tends to form the error bursts in many practical situations. Information set decoding for independent errors is the best decoding approach for random linear codes, but has exponential complexity. **Purpose:** To analyze the error-correcting capability and computational complexity of algorithms based on information sets in order to correct single error bursts. **Results:** We suggest the decoding criteria of minimal burst length for single bursts correction using information sets. We prove the correction of all error bursts within burst-correction capability of the code using this criterion. The task of decreasing the number of information sets required for correct decoding is considered, for this task the methodology of constructing the information sets with limited diameter, which are called dense sets, is investigated. We perform the analysis showing the possibility of decreasing the number of information sets from $O(n)$ to $O(1)$. We estimate the probability of information sets construction and formulate the optimization problem for maximizing this probability. **Practical relevance:** The developed decoder of single error bursts basing on minimal burst length criteria, as well as methodology of construction the collection of dense information sets with cardinality $O(1)$ guarantee the realization of burst-correction capability of random linear codes. **Discussion:** The results obtained show that the problem of single error bursts correction with random linear codes is polynomial. At the same time, the proposed methods make it possible to get a high probability construction of information sets for codes with a relatively low rate, and these probabilities can significantly change due to the use of more practical classes of codes, which may be considered as future research directions.

Keywords — information set decoding, Reiger bound, error burst correction, dense information set, decoding criteria.

For citation: Isaeva M. N. Decoding of single error bursts using minimal burst length criteria and information sets. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2025, no. 2, pp. 68–77 (In Russian). doi:10.31799/1684-8853-2025-2-68-77, EDN: MAHCAY

Financial support

The paper was prepared with the financial support of the Russian Science Foundation, project No. 22-19-00305 “Spatial-temporal stochastic models of wireless networks with a large number of users”.

References

1. Yang M., Pan Z., Djordjevic I. B. FPGA-based burst-error performance analysis and optimization of regular and irregular SD-LDPC codes for 50G-PON and beyond. *Opt. Express*, 2023, vol. 31, no. 6, pp. 10936–10946. doi:10.1364/OE.477546
2. Song L., Huang Q., Wang Z. Construction of multiple-burst-correction codes in transform domain and its relation to LDPC codes. *IEEE Trans. Commun.*, 2020, vol. 68, no. 1, pp. 40–54. doi:10.1109/TCOMM.2019.2948341
3. Vafi S. Cyclic low density parity check codes with the optimum burst error correcting capability. *IEEE Access*, 2020, vol. 8, pp. 192065–192072. doi:10.1109/ACCESS.2020.3032837
4. Xiao X., Vasic B., Lin S., Li J., Abdel-Ghaffar K. Quasi-cyclic LDPC codes with parity-check matrices of column weight two or more for correcting phased bursts of erasures. *IEEE Trans. Commun.*, 2021, vol. 69, no. 5, pp. 2812–2823. doi:10.1109/TCOMM.2021.3059001
5. Aharoni Z., Huleihel B., Pfister H. D., Permuter H. H. Data-driven polar codes for unknown channels with and without memory. *2023 IEEE International Symposium on Information Theory (ISIT)*, Taipei, IEEE, 2023, pp. 1890–1895. doi:10.1109/ISIT54713.2023.10206663
6. Sasoglu E., Tal I. Polar coding for processes with memory. *IEEE Trans. Inform. Theory*, 2019, vol. 65, no. 4, pp. 1994–2003. doi:10.1109/TIT.2018.2885797
7. Vajha M., Ramkumar V., Jhamtani M., Kumar P. V. On the performance analysis of streaming codes over the Gilbert – Elliott channel. *2021 IEEE Information Theory Workshop (ITW)*, Kanazawa, Japan, 2021, pp. 1–6. doi:10.1109/ITW48936.2021.9611448
8. Eckford A., Kschischang F., Pasupathy S. Analysis of low-density parity-check codes for the Gilbert – Elliott channel. *IEEE Transactions on Information Theory*, 2005, vol. 51, no. 11, pp. 3872–3889. doi:10.1109/TIT.2005.856934
9. Li L., Lv J., Li Y., Dai X., Wang X. Burst error identification method for LDPC coded systems. *IEEE Commun. Lett.*, 2024, pp. 1–5. doi:10.1109/LCOMM.2024.3391826
10. Fang Y., Chen J. Decoding polar codes for a generalized Gilbert – Elliott channel with unknown parameter. *IEEE Trans. Commun.*, 2021, vol. 69, no. 10, pp. 6455–6468. doi:10.1109/TCOMM.2021.3095195
11. Trofimov A. N., Taubin F. A. Improved bound on optimal reception error probability for an intersymbol interference channel. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2023, no. 5, pp. 33–42 (In Russian). doi:10.31799/1684-8853-2023-5-33-42, EDN: MDHOXU
12. Ovchinnikov A. A., Veresova A. M., Fominykh A. A. Decoding of linear codes for single error bursts correction based on the determination of certain events. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 6, pp. 41–52. doi:10.31799/1684-8853-2022-6-41-52, EDN: UWXZHN
13. Veresova A., Fominykh A., Ovchinnikov A. About usage of metrics in decoding of LDPC codes in two-state channels with memory. *2021 XVII International Symposium Problems of Redundancy in Information and Control Systems (REDUNDANCY)*, Moscow, IEEE, 2021, pp. 143–148. doi:10.1109/REDUNDANCY52534.2021.9606474
14. Isaeva M. N., Ovchinnikov A. A. Correction of single error bursts beyond the code correction capability using information sets. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 1, pp. 70–80 (In Russian). doi:10.17586/2226-1494-2024-24-1-70-80
15. Sung W., Coffey J. T. Information set decoding complexity for linear codes in bursty channels with side information. *Proceedings of 1995 IEEE International Symposium on Information Theory*, Whistler, BC, Canada, 1995, pp. 53. doi:10.1109/ISIT.1995.531155
16. Moon T. K. *Error Correction Coding: Mathematical Methods and Algorithms*. Second ed. Hoboken, NJ, Wiley, 2021. 992 p.
17. Gazi O. *Forward Error Correction Via Channel Coding*. Cham, Springer, 2020. 319 p.
18. Lin S., Li J. *Fundamentals of Classical and Modern Error-Correcting Codes*. Cambridge, Cambridge University Press, 2022. 840 p. doi:10.1017/9781009067928
19. Veresova A. M., Isaeva M. N., Ovchinnikov A. A. Estimation of independent errors and bursts correction capability of linear codes. *2024 Conference of Young Researchers in Electrical and Electronic Engineering (ElCon)*, Saint-Petersburg, IEEE, 2024, pp. 23–27. doi:10.1109/ElCon61730.2024.10468456
20. Isaeva M. N. Finding information sets when correcting error bursts with quasi-cyclic codes. *T-Comm*, vol. 17, no. 7, pp. 4–12 (In Russian). doi:10.36724/2072-8735-2023-17-7-4-12
21. Berlekamp E. R. The technology of error correcting codes. *Proc. IEEE*, 1980, vol. 68, pp. 564–593.

АБРОСИМОВ
Вячеслав
Константинович

Ведущий научный сотрудник Главного научно-исследовательского испытательного межвидового центра перспективного вооружения Минобороны России, Москва.
В 1974 году окончил Военную академию им. Ф. Э. Дзержинского по специальности «Баллистика и теория полета».
В 1974 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 187 научных публикаций, двух патентов на изобретения и пяти свидетельств на полезные модели.
Область научных интересов — искусственный интеллект, методы моделирования, методы адаптивного управления, теория полета летательных аппаратов.
Эл. адрес: avk787@yandex.ru

ВАСИЛЬЕВ
Владимир
Иванович

Профессор кафедры вычислительной техники и защиты информации Уфимского университета науки и технологий.
В 1980 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина) по специальности «Электронные вычислительные машины».
В 1990 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 400 научных публикаций и 14 свидетельств о регистрации программ для ЭВМ.
Область научных интересов — интеллектуальные системы управления, информационная безопасность.
Эл. адрес: vas0015@yandex.ru

ВО
Ньы Тхань

Заведующий кафедрой мехатроники факультета машиностроения Университета науки и технологий, Университет Дананга, Вьетнам.
В 2008 году окончил бакалавриат Университета политехники Калифорнии в Помоне по специальности «Машиностроение».
В 2017 году защитил диссертацию PhD по инженерии интеллектуальных механических систем.
Является автором более 50 научных публикаций.
Область научных интересов — системы управления, обработка изображений и звука, интеллектуальная автоматизация, приложения искусственного интеллекта.
Эл. адрес: vnthanh@dut.udn.vn

БОГАЧЕВ
Илья
Владимирович

Доцент Высшей школы кибернетики и цифровых технологий Тихоокеанского государственного университета, Хабаровск.
В 2010 году окончил Тихоокеанский государственный университет по специальности «Управление и информатика в технических системах».
В 2020 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 100 научных публикаций.
Область научных интересов — искусственные нейронные сети, цифровая обработка сигналов, сжатие информации.
Эл. адрес:
ilya.bogachev@yahoo.com

ВО
Вьет Чыонг

Студент, ассистент исследователя в группе по науке о данных на факультете информационных технологий Университета науки и технологий, Университет Дананга, Вьетнам.
Область научных интересов — искусственный интеллект, обработка изображений, методы анализа данных.
Эл. адрес:
102210383@sv1.dut.udn.vn

ВУЛЬФИН
Алексей
Михайлович

Профессор кафедры вычислительной техники и защиты информации Уфимского университета науки и технологий.
В 2008 году окончил Уфимский государственный нефтяной технический университет по специальности «Программное обеспечение вычислительной техники и автоматизированных систем».
В 2022 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 160 научных публикаций, 23 свидетельств о регистрации программ для ЭВМ и одного патента на изобретение.
Область научных интересов — интеллектуальный анализ данных, моделирование сложных технических систем.
Эл. адрес: vulfin.am@ugatu.su

ЗАДБОЕВ Вадим Александрович

Младший научный сотрудник Военной академии связи им. Маршала Советского Союза С. М. Буденного, Санкт-Петербург.
В 2017 году окончил Северо-Кавказский Федеральный университет по специальности «Информационная безопасность автоматизированных систем».
Является автором 29 научных публикаций.
Область научных интересов — информационная безопасность, системы контроля и управления безопасностью информационно-телекоммуникационных и информационно-вычислительных систем.
Эл. адрес: zadboev89@mail.ru

ИСАЕВА Мария Николаевна

Старший преподаватель кафедры инфокоммуникационных технологий и систем связи Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 2020 году окончила Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Информационная безопасность».
Является автором 26 научных публикаций.
Область научных интересов — теория кодирования, кодовая криптография, цифровая стеганография.
Эл. адрес: imn@guap.ru

ЛЕВЕНЕЦ Алексей Викторович

Профессор Высшей школы кибернетики и цифровых технологий Тихоокеанского государственного университета, Хабаровск.
В 1988 году окончил Хабаровский политехнический институт по специальности «Электронные вычислительные машины».
В 2018 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 200 научных публикаций.
Область научных интересов — цифровая обработка сигналов, сжатие информации, передача данных.
Эл. адрес: levalvi@bk.ru

ЗАЙЦЕВ Дамир Павлович

Магистрант, лаборант Высшей школы кибернетики и цифровых технологий Тихоокеанского государственного университета, Хабаровск.
В 2023 году окончил Тихоокеанский государственный университет по специальности «Информационные системы и технологии».
Является автором пяти научных публикаций и одного патента на изобретение.
Область научных интересов — искусственные нейронные сети, информационно-измерительные системы, сжатие информации.
Эл. адрес:
zautsev.damer@gmail.com

КИРИЛЛОВА Анастасия Дмитриевна

Старший преподаватель кафедры вычислительной техники и защиты информации Уфимского университета науки и технологий.
В 2015 году окончила бакалавриат по специальности «Информационная безопасность», в 2017 году — магистратуру по специальности «Информатика и вычислительная техника» Уфимского государственного авиационного технического университета.
В 2023 году защитила диссертацию на соискание ученой степени кандидата технических наук.
Является автором 50 научных публикаций и девяти свидетельств о регистрации программ для ЭВМ.
Область научных интересов — информационная безопасность, автоматизированные системы управления, управление рисками, когнитивное моделирование.
Эл. адрес:
kirillova.andm@gmail.com

ЛИПАТНИКОВ Валерий Алексеевич

Профессор, старший научный сотрудник Военной академии связи им. Маршала Советского Союза С. М. Буденного, Санкт-Петербург, заслуженный изобретатель РФ, член-корреспондент РАЕН.
В 1974 году окончил Военную академию связи им. Маршала Советского Союза С. М. Буденного по специальности «Специальная радиотехника».
В 2000 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 300 научных публикаций и 80 патентов на изобретения.
Область научных интересов — теория многоуровневой иерархической радиоэлектронной защиты, безопасности связи и информации инфотелекоммуникационных сетей.
Эл. адрес: lipatnikovanl@mail.ru

СВЕДЕНИЯ ОБ АВТОРАХ

ЛУЦКОВИЧ Андрей Иванович

Ассистент кафедры вычислительной техники и защиты информации Уфимского университета науки и технологий.
В 2001 году окончил Уфимский государственный авиационный технический университет по специальности «Организация и технология защиты информации».
Является автором семи научных публикаций.
Область научных интересов — интеллектуальные системы противодействия мошенничеству, методы машинного обучения при мониторинге информационной безопасности.
Эл. адрес:
andrey.lutskovich@gmail.com

МИХАЙЛОВА Екатерина Сергеевна

Младший научный сотрудник Главного научно-исследовательского испытательного межвидового центра перспективного вооружения Минобороны России, аспирант аспирантуры ПАО «НПО «Алмаз», Москва.
В 2023 году окончила Московский государственный технический университет им. Н. Э. Баумана по специальности «Системы летательных аппаратов».
Является автором восьми научных публикаций.
Область научных интересов — искусственный интеллект, методы моделирования, методы адаптивного управления, теория полета летательных аппаратов.
Эл. адрес:
ekaterinaolimp99@mail.ru

ФАМ Конг Тханг

Заместитель декана факультета информационных технологий Университета науки и технологий, Университет Дананга, Вьетнам.
В 2013 году окончил Тульский государственный университет по специальности «Вычислительные машины, комплексы, системы и сети».
В 2016 году защитил диссертацию на соискание ученой степени кандидата технических наук в Тульском государственном университете.
Является автором 35 научных публикаций.
Область научных интересов — обработка изображений, машинное обучение, наука о данных.
Эл. адрес: pcthang@dut.udn.vn

МЕЛЕХОВ Кирилл Витальевич

Адъюнкт Военной академии связи им. Маршала Советского Союза С. М. Буденного, Санкт-Петербург.
В 2013 году окончил Краснодарский филиал Военной академии связи им. С. М. Буденного по специальности «Комплексная защита объектов информатизации».
Является автором 38 научных публикаций и трех патентов на изобретения.
Область научных интересов — информационная безопасность, компьютерная безопасность, защита информации от несанкционированного доступа, способы контроля уязвимостей и управление безопасностью информационно-телекоммуникационных, информационно-вычислительных и распределенных информационных сетей.
Эл. адрес: kirill_melehov@bk.ru

СУЛАВКО Алексей Евгеньевич

Профессор кафедры комплексной защиты информации Омского государственного технического университета.
В 2009 году окончил Сибирскую государственную автомобильно-дорожную академию по специальности «Комплексное обеспечение информационной безопасности автоматизированных систем».
В 2023 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 200 научных публикаций и одного патента на изобретение.
Область научных интересов — распознавание образов, машинное обучение, биометрия, искусственный интеллект, защита информации, искусственные нейронные сети.
Эл. адрес: sulavich@mail.ru

ХО Ван Тхао

Студент, ассистент исследователя в группе по науке о данных на факультете информационных технологий Университета науки и технологий, Университет Дананга, Вьетнам.
Область научных интересов — искусственный интеллект, обработка изображений, методы анализа данных.
Эл. адрес:
102210376@sv1.dut.udn.vn

ХОАНГ
Нгуен Нят Минь

Студент, ассистент исследователя в группе по науке о данных на факультете информационных технологий Университета науки и технологий, Университет Дананга, Вьетнам.
Область научных интересов — искусственный интеллект, обработка изображений, методы анализа данных.
Эл. адрес:
102210361@sv1.dut.udn.vn

ШЕВЧЕНКО
Александр
Александрович

Старший научный сотрудник Военной академии связи им. Маршала Советского Союза С. М. Буденного, Санкт-Петербург.
В 2015 году окончил Рязанский государственный радиотехнический университет по специальности «Компьютерная безопасность».
В 2021 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 86 научных публикаций и шести патентов на изобретения.
Область научных интересов — компьютерная безопасность, информационная безопасность, способы контроля уязвимостей и управления безопасностью информационно-вычислительных, информационно-телекоммуникационных и распределенных информационных сетей.
Эл. адрес:
alex_pavel1991@mail.ru

ЧАН
Минь Куан

Студент, ассистент исследователя в группе по науке о данных на факультете информационных технологий Университета науки и технологий, Университет Дананга, Вьетнам.
Область научных интересов — искусственный интеллект, обработка изображений, методы анализа данных.
Эл. адрес:
102210369@sv1.dut.udn.vn

Уважаемые авторы!

При подготовке рукописей статей необходимо руководствоваться следующими рекомендациями.

Статьи должны содержать изложение новых научных результатов. Название статьи должно быть кратким, но информативным. В названии недопустимо использование сокращений, кроме самых общепринятых (РАН, РФ, САПР и т. п.).

Текст рукописи должен быть оригинальным, а цитирование и самоцитирование корректно оформлено.

Объем статьи (текст, таблицы, иллюстрации и библиография) не должен превышать эквивалента в 20 страниц, напечатанных на бумаге формата А4 на одной стороне через 1,5 интервала Word шрифтом Times New Roman размером 13, поля не менее двух сантиметров.

Обязательными элементами оформления статьи являются: индекс УДК, заглавие, инициалы и фамилия автора (авторов), ученая степень, звание (при отсутствии — должность), полное название организации, аннотация и ключевые слова на русском и английском языках, ORCID и электронный адрес одного из авторов. При написании аннотации не используйте аббревиатур и не делайте ссылок на источники в списке литературы. Предоставляйте подрисуночные подписи и названия таблиц на русском и английском языках.

Статьи авторов, не имеющих ученой степени, рекомендуется публиковать в соавторстве с научным руководителем, наличие подписи научного руководителя на рукописи обязательно; в случае самостоятельной публикации обязательно предоставляйте заверенную по месту работы рекомендацию научного руководителя с указанием его фамилии, имени, отчества, места работы, должности, ученого звания, ученой степени.

Простые **формулы** набирайте в Word, сложные с помощью редактора Mathtype или Equation. Для набора одной формулы не используйте два редактора; при наборе формул в формульном редакторе знаки препинания, ограничивающие формулу, набирайте вместе с формулой; для установки размера шрифта в Mathtype никогда не пользуйтесь вкладкой Other, Smaller, Larger, используйте заводские установки редактора, не подгоняйте размер символов в формулах под размер шрифта в тексте статьи, не растягивайте и не сжимайте мышью формулы, вставленные в текст; пробелы в формуле ставьте только после запятых при перечислении с помощью Ctrl+Shift+Space (пробел); не отделяйте пробелами знаки: + = – ×, а также пространство внутри скобок; для выделения греческих символов в Mathtype полужирным начертанием используйте Style → Other → bold.

Для набора формул в Word никогда не используйте вкладки: «Уравнение», «Конструктор», «Формула» (на верхней панели: «Вставка» — «Уравнение»), так как этот ресурс предназначен только для внутреннего использования в Word и не поддерживается программами, предназначенными для изготовления оригинал-макета журнала.

При наборе символов в тексте помните, что символы, обозначаемые латинскими буквами, набираются светлым курсивом, русскими и греческими — светлым прямым, векторы и матрицы — прямым полужирным шрифтом.

Подробнее см. http://i-us.ru/index.php/ius/author-guide

Иллюстрации:

— рисунки, графики, диаграммы, блок-схемы предоставляйте в виде отдельных исходных файлов, поддающихся редактированию, используя векторные программы: Visio (*.vsd, *.vsdx); Adobe Illustrator (*.ai); Coreldraw (*.cdr, версия не выше 15); Excel (*.xls); Word (*.docx); AutoCad, Matlab (экспорт в PDF, EPS, SVG, WMF, EMF); Компас (экспорт в PDF); веб-портал DRAW.IO (экспорт в PDF); Inkscape (экспорт в PDF);

— фото и растровые — в формате *.tif, *.png с максимальным разрешением (не менее 300 pixels/inch).

Наличие подрисуночных подписей и названий таблиц на русском и английском языках обязательно (желательно не повторяющих дословно комментарии к рисункам в тексте статьи).

В редакцию предоставляются:

— сведения об авторе (фамилия, имя, отчество, место работы, должность, ученое звание, учебное заведение и год его окончания, ученая степень и год защиты диссертации, область научных интересов, количество научных публикаций, домашний и служебный адреса и телефоны, e-mail), фото авторов: анфас, в темной одежде на белом фоне, должны быть видны плечи и грудь, высокая степень четкости изображения без теней и отблесков на лице, фото можно представить в электронном виде в формате *.tif, *.png, *.jpg с максимальным разрешением — не менее 300 pixels/inch при минимальном размере фото 40×55 мм;

— экспертное заключение;

— экспортное заключение.

Список литературы составляется по порядку ссылок в тексте и оформляется следующим образом:

— для книг и сборников — фамилия и инициалы авторов, полное название книги (сборника), город, издательство, год, общее количество страниц, doi;

— для журнальных статей — фамилия и инициалы авторов, полное название статьи, название журнала, год издания, номер журнала, номера страниц, doi;

— ссылки на иностранную литературу следует давать на языке оригинала без сокращений;

— при использовании web-материалов указывайте адрес сайта и дату обращения.

Список литературы оформляйте двумя отдельными блоками по образцам lit.dot на сайте журнала (http://i-us.ru/paperrules): Литература и References.

Более подробно правила подготовки текста с образцами изложены на нашем сайте в разделе «Руководство для авторов» – http://i-us.ru/index.php/ius/author-guide.

Контакты

Куда: 190000, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. А, ГУАП, РИЦ
Кому: Редакция журнала «Информационно-управляющие системы»
Тел.: (812) 494-70-02
Эл. почта: ius.spb@gmail.com
Сайт: www.i-us.ru