Module for Detecting and Analyzing Cyber Threats in Information and Communication Systems of Civil Aviation Facilities*

Anna Ilyenko[1,†], Sergiy Gnatyuk[1,*,†], Sergii Ilyenko[1,†], Olena Prokopenko[1,†] and Valentyna Teliushchenko[1,†]

[1] *The State University "Kyiv Aviation Institute", 1 Liubomyra Huzara ave., 03058 Kyiv, Ukraine*

Abstract

This paper focuses on the problems of ensuring cybersecurity of civil aviation information and communication systems, that are critical for flight safety and the functioning of the industry overall. With the growth of information flows, the complexity of communication systems, and the increasing frequency of cyber threats, it is imperative to develop effective methods for detecting and neutralizing attacks. The paper highlights the current cybersecurity challenges in the aviation sector, including DoS, DDoS, XSS, CSRF, and ransomware attacks. The paper analyses threat detection methods, including signature-based, anomaly analysis, machine learning, and combined approaches. The architecture of the module for detecting and analyzing cyber threats using combined machine learning models, such as Random Forest, MLP, and SVM, is provided. This allows the system to increase its accuracy and sensitivity, reduce the number of false positives, and ensure adaptability to new threats. In particular, the results of the study showed an increase in attack detection accuracy by 1.05–1.48 times, sensitivity by 1.02 times, and predictive ability by 2.18 times. The module is based on an integrated approach that includes real-time data collection, pre-processing, analysis based on machine learning algorithms, and automatic threat response. The proposed system demonstrates high efficiency in real-time conditions, protecting the surface-to-air and air-to-air channels. It can be adapted for other critical industries. Thus, the proposed solution is a promising tool for ensuring cybersecurity in civil aviation and critical infrastructure in general.

Keywords

cybersecurity, information and communication systems, civil aviation, machine learning, cyber threat detection, combined models

1. Introduction

Civil aviation is one of the most important critical infrastructures that not only transports passengers and cargo but also supports global economic and social processes [1, 2]. With the development of technology, most of its information and communication systems and processes are heavily dependent on digital tools, such as air traffic control systems, navigation systems, and the processing and transmission of information between onboard and ground services. This trend significantly increases the level of vulnerability to cyberattacks. With the growth of information flows and the complexity of civil aviation communication systems, ensuring the security of information flows in the information and communication systems of civil aviation facilities is becoming critically important.

* *CPITS 2025: Workshop on Cybersecurity Providing in Information and Telecommunication Systems, February 28, 2025, Kyiv, Ukraine*

* Corresponding author.

† These authors contributed equally.

anna.ilienko@npp.nau.edu.ua (A. Ilyenko); serhii.hnatiuk@npp.nau.edu.ua (S. Gnatyuk); serhii.ilienko@npp.nau.edu.ua (S. Ilyenko); olena.prokopenko@npp.nau.edu.ua (O. Prokopenko); valentyna.teliushchenko@npp.nau.edu.ua (V. Teliushchenko)

0000-0001-8565-1117 (A. Ilyenko); 0000-0003-4992-0564 (S. Gnatyuk); 0000-0002-0437-0995 (S. Ilyenko); 0000-0001-9895-888X (O. Prokopenko); 0000-0001-6026-5105 (V. Teliushchenko)

The issue of safe operation of civil aviation facilities is becoming increasingly urgent in terms of the negative impact of ever-growing cyber threats and the deterioration of cyber security in the aviation industry overall. The protection status of the ground-to-air and air-to-air channels in such aviation systems is at different levels and directly depends on the activities of all components of aviation activities (information network traffic control, etc.) Some communication channels are currently not protected at all and are open, which provokes an inevitable increase in cyberattacks and requires the introduction and use of modern information and communication technologies in such communication channels. According to research, attacks can affect the following: air traffic control systems; navigation systems; onboard and ground communication systems. These systems are key to flight safety, and even the slightest interference can lead to serious consequences, such as disruption of flight control, navigation failures, or loss of control over the aircraft. Information flows in the information and communication systems of civil aviation facilities cover a wide range of data: from flight information to technical reports and navigation data. This requires constant protection both during data transmission and during storage and processing. Different types of data have different levels of criticality, so special methods are needed to ensure their security. Today, civil aviation information and communication systems have a complex life cycle, which includes design; deployment; operation; and modernization. At each of these stages, information flows may be exposed to various threats. Existing cybersecurity solutions do not always consider the specifics of civil aviation information and communication systems.

The modern development of information technology and the growth of cyber threats make the security of information and communication systems at civil aviation facilities a critical task. The aviation industry is one of the most attractive targets for cyberattacks, as it is integrated into the global infrastructure and has a huge impact on passenger safety and the functioning of the economy. Attacks targeting aviation communication channels can have significant consequences, including disruption of flight control systems, loss of confidential data, or even disruption of airport operations.

Over the past few years, cyberattacks on the aviation industry have been increasing in both frequency and sophistication. In particular, from 2022 to 2024, the number of attacks on critical infrastructure, including aviation, increased significantly. In 2023, attacks using the MOVEit ransomware became widely known, causing data breaches at airlines such as British Airways and Aer Lingus. The vulnerability allowed the attackers to gain access to employees' data, including payment and other confidential information. In the period from 2022 to 2023, there were also numerous attempts to compromise aviation systems using botnets, which makes it difficult to identify the source of attacks and can lead to significant operational disruptions. Another problem is the growing number of DDoS attacks that disrupt airports and airlines. DDoS attacks on flight control systems lead to flight delays and financial losses namely, a cyberattack on Japan Airlines (December 26, 2024), and two Italia airports (December 28, 2024).

The authors' analysis of cyber threats in civil aviation over the past five years shows a significant increase in attacks. Key trends: Ransomware: The number of attacks increased from 20% in 2019 to 40% in 2023.DDoS attacks: The share of attacks increased from 15% to 25%, indicating a growing interest of attackers in destabilizing aviation online services. Data threats: A significant increase from 25% to 45% was noted, highlighting the critical importance of protecting personal and corporate data. Malware: The number of incidents increased from 10% to 20%, indicating the active use of malware to compromise systems. Social engineering: Although its share is smaller, it is gradually increasing—from 5% to 15%, remaining a powerful tool for cybercriminals.

The chart shows the percentage distribution of major cyber threats in aviation over the last 5 years (Fig. 1).

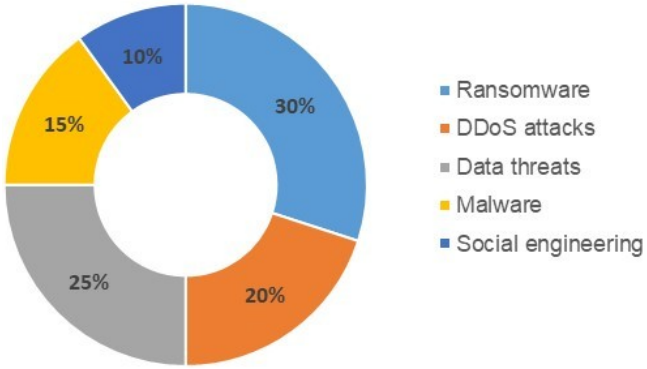

Figure 1: Cyber incidents of civil aviation facilities over the last 5 years

The data highlights the need to strengthen cybersecurity measures, including data protection, communication protocols, and flight management systems. It is also important to raise staff awareness and implement modern technologies to monitor and prevent attacks.

2. Problem statement

One of the most promising approaches to improving the security of aviation communication channels is the use of machine learning to detect anomalous behavior in network traffic [3, 4]. However, single machine learning models have limitations in accuracy and ability to adapt to new types of attacks. In this regard, it is proposed to use a combination of several machine learning models, such as Random Forest, Multi-Layer Perceptron (MLP), and Support Vector Machine (SVM), to create a more reliable and resilient cyberattack detection system.

The purpose of this study is to design and evaluate the effectiveness of a combined machine-learning system for real-time threat detection in aviation communication channels. The paper shows how the combination of models can increase the accuracy of attack detection, improve the system's sensitivity to new threats, and ensure classification stability under challenging network conditions. The test results confirm that the proposed solution is an effective tool for ensuring cybersecurity in the aviation sector. The designed software module combines the signature method with machine learning methods such as Random Forest, MLP, and SVM, ensuring increased efficiency of cyber threat detection. This model provides the ability to adapt to new types of attacks, improve detection accuracy, and reduce threat processing time. In a dynamic information and communication systems environment, this is crucial for ensuring the reliability and uninterrupted operation of systems. The module aims to improve detection accuracy using combined models. The results of the study showed an increase in accuracy, sensitivity, and a reduction in false positives, making this approach particularly effective in the context of civil aviation cybersecurity.

3. Modern methods of cyber-attack detection

Various methods and approaches are used to effectively detect cyberattacks, each with its characteristics, advantages, and disadvantages. Understanding these methods is key to building robust cybersecurity systems that can detect and respond to threats promptly [5–12].
The methods for detecting cyber-attacks are as follows:

1. Signature-based—this method describes a cyberattack using a set of rules or a formal model, which can be a character string, a semantic expression in a special language, etc. The essence of this method is using a specialized database of cyberattack patterns (signatures) to search for actions that fall under the definition of 'cyberattack'. The signature method can protect against a virus or hacker cyberattack when its signature is already known and entered the database. If the network is subjected to the first attack from the outside, the

first infection is still unknown, and the database simply does not have enough signatures to search for it, so the system will not be able to signal danger, as it considers the attacking activity to be legitimate.

2. Anomaly-based detection—this method consists of finding and identifying elements, events, or observations that deviate from the expected behavior or other elements of the data set. In the context of detecting network abuse or intrusion, unexpected activity spikes are of primary interest, not just rare events. These spikes do not meet the common statistical definition of outliers as rare objects, so many outlier detection methods, especially unsupervised algorithms, are ineffective without appropriate data collection.
3. Machine learning and AI—this method is very effective nowadays because AI can analyze huge amounts of data from various sources, identifying patterns of activity within an organization. It can track the time and place of login, the volume of network traffic, as well as the devices and cloud applications used. By understanding what actions are typical for employees, AI can detect anomalies that may indicate potential threats and require further investigation. To ensure confidentiality, data from one organization is not used to train AI in other companies. Instead, AI uses synthesized global threat intelligence gathered from many organizations to improve the efficiency of its analysis [8, 10, 13, 14].
4. Combined—this method of cyber-attack detection includes different approaches to improve detection efficiency and reduce the number of false positives. This method takes advantage of several techniques, such as signature analysis, anomaly analysis, and machine learning, to create a more reliable and accurate threat detection system. From signature analysis, this method uses databases of known attack patterns or signatures to compare current activity with known attacks. With anomaly analysis, this method detects deviations from normal system or network behavior by creating a model of normal activity and tracking any anomalies. This helps to detect new, previously unknown attacks. With machine learning and AI, this method automatically detects anomalies and attacks by analyzing large amounts of data and identifying complex patterns.

Considering these aspects, it becomes clear that effective cyberattack detection requires a comprehensive approach that combines the advantages of several classes of solutions. Thus, to further improve cyberattack detection systems, it is proposed to develop a software solution that integrates the best practices from each of the analyzed classes. In particular, it is important to pay attention to the use of artificial intelligence to automate threat detection, which will increase the effectiveness of responding to new and complex attacks.

A combined method is the best choice for aviation communication channels that require maximum accuracy, adaptability, and speed, namely a combination of signature-based methods and machine learning algorithms. It allows the detection of both known and new threats, reducing risks to critical infrastructure (Table 1).

Table 1
Comparative analysis of cyberattack detection methods

Method	Advantages	Disadvantages	Detectable attacks
Signature-based	High accuracy for known attacks	Failure to detect new attacks need for updates	DoS, SQLi, XSS
Anomaly-based	Detection of new attacks	False positives need for adaptation	DDoS, intrusion attempts
Machine learning and AI	Automation, adaptation to new threats	Complexity and resource intensity	Combined attacks, XSS + CSRF
Combined	Maximum accuracy, reduction of false alarms	High complexity and costs	All the attacks mentioned

4. Description of the proposed solution for detecting cyberattacks using machine learning

Modern cybersecurity systems are faced with an ever-increasing number of rapidly evolving threats. In particular, as the complexity and frequency of cyberattacks increase, existing defense methods are proving insufficiently effective. Traditional threat detection systems are largely based on static rules and signatures, making them vulnerable to new, unknown types of attacks [15].

Thus, there is an urgent need for new approaches that can adapt to changing conditions and provide more flexible and effective detection of cyberattacks.

Consider further the main machine learning algorithms that will be used in the developed system to detect cyber threats. These algorithms are responsible for analyzing data, identifying potential threats, and making decisions on further actions. Through the use of machine learning, the system can adapt to new challenges in the field of cybersecurity, improving its methods and approaches in real time [14, 16–19].

4.1. Random forest

Random Forest is an ensemble machine learning method that consists of many decision trees that work together. Each tree in the random forest is generated from a random sample of data, and the result is determined by the voting of all the trees. This method is a powerful tool for classification and regression and is especially effective in situations where data has many features [16].

Formula for node splitting calculation (Gini(D)):

$$Gini(D)=1-\sum_{i=1}^{C} p_i^2, \tag{1}$$

where $Gini(D)$ is the measure of node D impurity; C is number of classes; p_i is the probability of class i in the node D.

Random Forest is an ensemble method that combines many individual decision trees to improve the accuracy and resilience of the model. This algorithm was chosen because of its ability to work efficiently with large amounts of data and many features, which is typical for cybersecurity tasks. Random Forest effectively processes data containing noise or missing values, which reduces the likelihood of false positives. In addition, this method is highly interpretable, making it easy to explain the model's decisions and identify the key features that influenced the result.

4.2. Support vector machine

SVM is an algorithm for data classification that finds the hyperplane that best separates classes in the feature space. Choosing a hyperplane that maximizes the distance (margin) between different classes helps SVM achieve high classification accuracy [20–22].

The formula for loss function (L):

$$L(w,b,x_i,y_i)=max\ (0,1-y_i(w\times x_i+b)), \tag{2}$$

where w is weight vector; b is offset; x_i is feature vector; y_i is class label (+1 or –1).

The basis for this model is the concept of margin maximization, which is used to achieve optimal class separation in the feature space. The loss function L minimizes classification errors, ensuring optimal class separation.

Support Vector Machine is a powerful method for data classification, especially when it involves tasks with many features and a small number of training examples. SVM was chosen because of its ability to work effectively in high-dimensional feature space and provide good generalization on new data. This method also demonstrates high accuracy in anomaly detection, which is an important aspect of cybersecurity tasks where correct threat recognition is critical.

4.3. Neural networks

Neural networks are a type of deep learning algorithm that simulates the human brain to process complex patterns and features in large data sets. Neural networks consist of different layers of neurons that transform input data, learn from it, and make decisions based on their processing [16, 17, 23]. The formula for activation function:

$$ReLU(x)=max\ (0,x), \tag{3}$$

where $ReLU(x)$ is the output of the neuron; x is the input value (the sum of the weighted inputs of the neuron).

The basis for this model is the concept of activation functions that determine the output of neurons in the network. ReLU (Rectified Linear Unit) is one of the most popular activation functions that provides fast and efficient data processing in deep neural networks. The use of machine learning in cyber threat detection systems for civil aviation is critical due to the unique challenges faced in this area. The aviation infrastructure is associated with large flows of real-time data, high-security requirements, and the need to quickly adapt to new types of attacks.

All these algorithms can adapt to new threats by learning from the latest data. For example, in the aviation sector, these algorithms can be integrated into systems for traffic control, analyzing the behavior of onboard software, and assessing the cybersecurity risks of airport information systems. In addition, this approach allows for the integration of machine learning into complex incident response systems. For example: automated detection of suspicious requests to the flight control system; detection of deviations in aircraft routes that may indicate intrusion into navigation systems; and monitoring of the behavior of IoT devices at the airport to prevent them from being used as points of attack.

Thus, the use of several methods in combination makes it possible to build a multi-level defense system that not only identifies threats but also adapts to new challenges. This approach makes the system as resilient and reliable as possible, which is crucial for ensuring cybersecurity in critical industries such as civil aviation.

Table 2
Comparative analysis of machine learning models

Method	Advantages	Disadvantages	Use for protection
Random Forest	High accuracy, noise resistance	Difficult to interpret, resource-intensive	Anomaly detection, DoS, DDoS
SVM	Efficient in high-dimensional spaces	Time-consuming to learn, non-adaptive	SQL Injection, XSS
MLP	Flexibility, ability to model complex tasks	Risk of relearning, resource-intensive	Multi-phase attacks, anomalies

The focus of the module for detecting and analyzing cyber threats in the information and communication systems of civil aviation facilities is to determine network traffic metrics (Tables 3 and 4). Such data can be part of training sets for machine learning models in cybersecurity tasks. Network traffic metrics can be an effective tool for detecting and preventing cyber threats in aviation communication channels. Their use will provide an additional layer of protection for critical communications systems against cyberattacks and contribute to the comprehensive security of aviation operations. Metrics can be used to protect aviation communications because they are universal indicators of network traffic and can help detect anomalies or cyber threats such as intrusion attempts, network scans, or DoS/DDoS attacks. The metrics can be used in the context of protecting aviation communication channels:

1. Real-time traffic monitoring. The use of metrics such as Flow Packets/s, Flow Bytes/s, and Flow Duration can detect abnormal traffic volumes or unusual behavior, such as a sudden increase in data volume, which may indicate an attack.
2. Detecting suspicious patterns in the network. The SYN Flag Count, FIN Flag Count, and RST Flag Count metrics help identify attempts to start or end suspicious connections. For example, a large number of SYN packets may indicate a SYN flood attack.
3. Protection against attacks on aviation systems. ACK Flag Count and PSH Flag Count can be useful for detecting attacks that attempt to exploit incomplete connections. Subflow Fwd Packets and Subflow Bwd Packets provide data flow tracking, which is important in detecting hidden data channels or exfiltration attempts.
4. Communication Protocol Behaviour Anomalies Metrics related to TCP flags (e.g. URG Flag Count, PSH Flag Count) can detect the usage of non-standard or unusual flags in traffic, which can be a sign of transport layer attacks.

Table 3
Important metrics for aviation communication channels

Metric	Description	Reason of less importance
Subflow Fwd Packets	Packets in the forward subflow	Less impactful on overall analysis, useful for in-depth analysis
Subflow Bwd Packets	Packets in the reverse subflow	Important for ex-post analysis, but not critical in real time
FIN Flag Count	Number of FIN packets	More related to session completion, less commonly used
PSH Flag Count	Number of PSH packets	Indicates data transfer immediately, but not always important
URG Flag Count	Number of URG packets	Rarely used in attack practice, minimal value
Fwd Packets/s	Packet rate in the forward direction	Less informative than the overall transmission rate

These metrics can be used as features for training machine learning models to create an early warning system that automatically detects suspicious activity in the airline channel network, including attacks such as DoS, DDoS, port scanning, etc. Such data can be a part of training sets for machine learning models in cybersecurity tasks. It is important to consider the limitations of system response time and computing resources while processing network traffic in real time. In the context of aviation communication channels, where high reliability, minimal latency, and security are important, it is critical to quickly detect anomalies that may indicate cyber threats or disruptions. Therefore, the authors propose to classify metrics into those that are critical for rapid attack detection and those that can be neglected without significant performance degradation.

Table 4
Secondary metrics (less critical for real-time)

Metric	Description	Reason of importance
Flow Packets/s	Packet transfer rate in the stream	Can detect DoS/DDoS attacks and abnormal changes in traffic
SYN Flag Count	Number of SYN packets	Detect SYN floods and hacking attempts through session establishment
RST Flag Count	Number of RST packets	Detects attempts to force termination of sessions
Flow Duration	Duration of the stream	Can indicate suspiciously long sessions or delays
ACK Flag Count	Number of ACK packets	Monitor connection completion and data transfer confirmations
Flow Bytes/s	Byte transfer rate in the stream	Displays the amount of data transferred, which is important for leak detection
Bwd Packets/s	Packet rate in the opposite direction	Detection of abnormal activity in response to requests
Total Fwd Packets	Total number of transmitted packets	Important for analyzing the volume and type of transmission
Total Backward Packets	Total number of packets in the opposite direction	Detecting anomalous responses to requests

The authors determine that critical metrics for aviation communication channels (Table 3) are those that can rapidly detect denial-of-service attacks (DoS/DDoS), interception attempts, or data integrity violations (SYN, RST, ACK flags), while secondary metrics are more suitable for offline analysis or improving detailed models, but they are not key to detecting threats in real-time (Table 4). These metrics can be effectively used to analyze the network flows of aviation communication channels. They allow the detection of anomalies and cyber threats in real time, which is critical to ensuring the security of flight control systems, data transmission, and communications between airports and aircraft.

The impact of using metrics: Accuracy: High accuracy is essential to minimize false positives and false negatives. In the aviation industry, it helps to avoid unnecessary alerts and focus on real threats. Precision: High accuracy ensures that many detected threats are genuine attacks, reducing the load on response systems and time spent on verifying false positives. Recall: Recall is critical in real-time networks because it ensures that as many threats as possible are detected, even if they are new or subtle types of attacks. F1-score: It provides a balance between recall and precision, which allows you to optimize the process of detecting attacks without significantly increasing false positives. Latency: For aviation channels, the speed of threat detection is important to prevent potential disruptions of critical systems.

Reducing attack detection delays is a key performance indicator for the model. Use cases: Detecting DDoS attacks: Accuracy and recall metrics provide a quick way to detect anomalies in traffic; XSS and CSRF analysis: F1-score provides accurate identification of threats, even if they are disguised as legitimate requests$ real-time anomaly monitoring: The system can respond quickly to abnormal changes in network traffic due to its recall and reduced latency.

The described metrics and combined machine learning models are effective tools for analyzing network flows of aviation communication channels. They provide high accuracy and speed of threat detection, which minimizes the risk of disruption, prevents cyberattacks, and improves the overall security of aviation infrastructure.

5. The architecture of the module for detecting and analyzing cyber threats in information and communication systems of civil aviation facilities

The basic architecture of the module consists of several components that collect, process, and analyze data using machine learning methods, and respond to detected threats. Each module performs its specific function, integrating into the overall system to ensure continuous monitoring and protection of the computer network of objects that provide effective detection, analysis, and response to cyber threats:

1. Data Collection Module. The data collection module performs the key function of threat detection by receiving information about network activity in real-time or by downloading it from pre-saved CSV files. The main scenarios of the module's operation include two modes: a) Real-time data collection: The module uses an element to monitor network activity. Psutil collects metrics such as the number of bytes and packets sent and received. This process takes place periodically, for example, every 5 seconds, and the result is saved to a data structure for further processing and analysis. b) Downloading data from CSV: The module also supports the ability to load data from CSV files, which enables the use of previously saved datasets. This module is the basis for further processing steps such as data pre-processing and analysis based on machine learning algorithms.
2. Data pre-processing module. The data pre-processing module performs key operations to clean, normalize, and reduce the dimensionality of the data before it is passed to the machine learning algorithms. Its primary task is to prepare the collected or uploaded data for further analysis, ensuring that it is correct and suitable for use in models. This method standardizes each feature, converting it to a value with a mean of 0 and a standard deviation of 1. This is necessary to improve the performance of machine learning models, especially for algorithms that depend on data scaling, such as SVMs and neural networks. PCA (Principal Component Analysis) is used to efficiently process large amounts of data and prevent model overtraining. This module provides data readiness for the next stage of analysis based on machine learning algorithms, allowing models to work correctly and efficiently.
3. AI analysis module. The AI analysis module is the core part of the system that detects anomalous activity or threats based on collected and pre-processed data. This module uses several machine learning algorithms to classify data and identify potential cyber threats. The system uses pre-trained models that have been created based on historical data with labels. The module allows both training new models and loading already trained models from files to analyze new data in real-time. The module supports three main types of models: RandomForest, SVM, and MLP. The model evaluation is performed using several metrics to understand how effectively the system detects threats. The main functions of this stage are: model training, which is performed on historical data with labels; during real-time analysis, pre-trained models are loaded; performance evaluation, using metrics to analyze classification results and model accuracy; real-time analysis, using models to detect anomalies in real-time data; various metrics are used to evaluate the performance of machine learning models, such as accuracy, precision, recall, F1-score, area under ROC curve. These metrics provide an objective assessment of the model's performance in detecting anomalies or cyber threats.

4. Response module. The Response Module is a critical part of the system that provides automatic response to detected anomalies and cyber threats. Its functionality ensures the prompt sending of notifications to administrators and the implementation of measures to minimize possible threats, such as blocking suspicious IP addresses. As soon as the system detects anomalies or a potential cyber threat, it is necessary to immediately notify the administrator or a contact person who has permission to send emails via SMTP servers. The main advantages of such a system are that alerts are sent to emails in real-time, which allows for a prompt response to potential threats, and that the use of .env files to store credentials increases the security of the notification process. Blocking suspicious IP addresses is one of the most important measures in response to cyber threats. For this purpose, a tool or corresponding commands on other operating systems are used to dynamically block IP addresses at the firewall level.
5. Data storage module. One of the key requirements of the system is to maintain an event log and store data on detected anomalies and actions taken to neutralize them.

This architecture ensures continuous monitoring, data processing, anomaly detection, and rapid response, contributing to a high level of security for civil aviation information and communication systems.

6. Effectiveness evaluation of the module for detecting and analyzing cyber threats in information and communication systems of civil aviation facilities using machine learning

The evaluation of the models is performed using several metrics to understand how effectively the system detects threats. The main functions of this stage are: model training, which is performed on historical data with labels; during real-time analysis, pre-trained models are loaded; effectiveness evaluation, using metrics to analyze the classification results and accuracy of the models; real-time analysis, using models to detect anomalies in the data collected in real-time; various metrics are used to evaluate the effectiveness of machine learning models, such as accuracy, precision, recall, F1-measure and area under ROC curve. These metrics provide an objective assessment of the model's performance in detecting anomalies or cyber threats. To assess the accuracy of each model, we also calculated confusion matrices, which display the number of cases in which the model correctly identified threats (positive cases) and normal data (negative cases), and situations where it made a mistake. This tool is useful for evaluating the accuracy of the model, as it indicates how effectively it detects threats and distinguishes them from normal traffic.

Below are the mathematical formulas for each metric.

1. Accuracy. The ratio of correct predictions to the total number of predictions. This metric can be used to assess the overall performance of the model.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{4}$$

where TP is the number of true positive results (correctly classified attacks); TN is the number of true negative results (correctly classified non-attacks); FP is the number of false positives (not attacks that are mistakenly classified as attacks); FN is the number of false negatives (attacks that are mistakenly classified as non-attacks).

2. Recall. The rate of true positive cases that were correctly detected by the model. An important metric for assessing the model's ability to detect all possible threats.

$$Recall = \frac{TP}{TP + FN}, \tag{5}$$

3. Precision. The rate of correct positive predictions among all cases that the model recognized as threats. It helps to estimate the number of false positives.

$$Precision = \frac{TP}{TP + FP}, \tag{6}$$

4. F1-score. The harmonic mean between the accuracy of positive predictions and recall. This is an integrated metric that considers both false positives and false negatives.

$$F1 = 2 \times \frac{¿ision \times Recall}{Precision + Recall}, \tag{7}$$

5. The area under ROC curve (AUC-ROC). The area under the ROC curve is a measure of the model's ability to distinguish between positive and negative classes. The values range from 0.5 (random prediction) to 1 (perfect prediction).

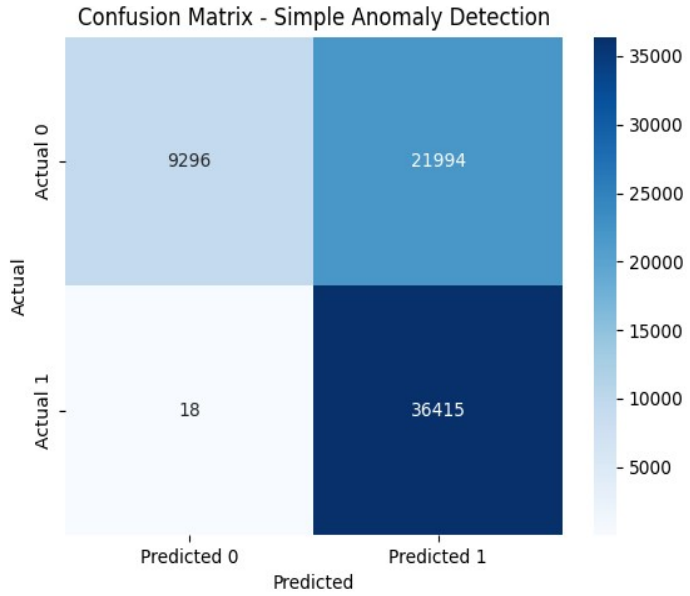

Figure 2: Confusion matrix for a threshold-based detection scheme

The proposed combination of machine learning models, such as Random Forest and MLP, has improved the accuracy and stability of the cyberattack detection system for civil aviation information networks (Figs. 2–5), i.e. the test model showed the following results: the accuracy of attack detection was increased by up to 3.35 times and the precision was increased from 1.05 to 1.48 times.

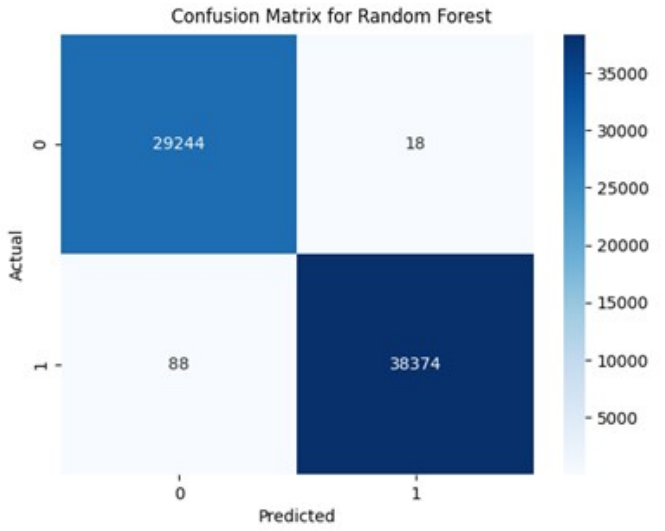

Figure 3: Confusion matrix for Random Forest

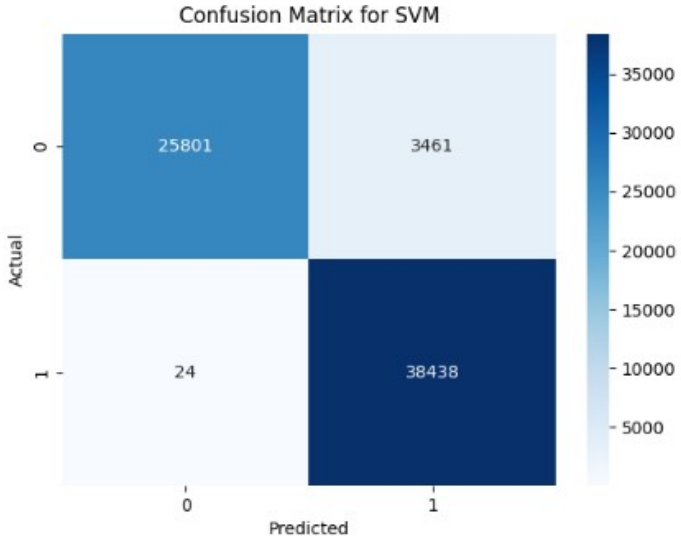

Figure 4: Confusion matrix for SVM

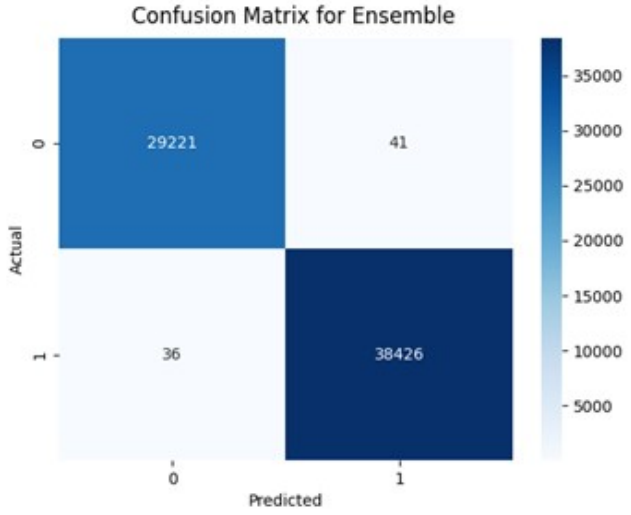

Figure 5: Confusion matrix for the combined approach Random Forest+MLP

The author's model also provides an increase in sensitivity to attack detection (recall) by up to 1.02 times and an increase in the model's predictive ability (f1-score) by 1.04 to 2.18 times. The proprietary software solution can detect various threats in real-time, such as DoS (Denial of Service), DDoS (Distributed Denial of Service), XSS (Cross-Site Scripting), CSRF (Cross-Site Request Forgery), and anomalies in network traffic. Thus, the proposed model provides high accuracy due to the combination of several models; the ability to quickly detect cyberattacks in real-time; and the model can adapt to new types of attacks (Table 5).

Table 5
Effectiveness evaluation of the proposed solution using machine learning

	Precision	Recall	F1	Accuracy
Threshold-based detection (deviation from the average)	0.2971	0.9981	0.4579	0.6750
Random Forest	0.9995	0.9977	0.9985	0.9984
MLP	0.9974	0.9977	0.9975	0.9972
SVM	0.9174	0.9994	0.9570	0.9485
Combined (Random Forest + SVM + MLP)	0.9939	0.9991	0.9968	0.9959
Combined (Random Forest + MLP)	0.9989	0.9991	0.9992	0.9998

This model allows us to effectively monitor anomalies and detect cyberattacks on aviation communication channels. Such a multi-level system makes it possible to achieve high efficiency in detecting, classifying, and preventing cyber threats in critical communication channels.

Conclusions

These combined machine learning models, in particular Random Forest and MLP, have demonstrated significant benefits in creating a cyberattack detection system for civil aviation information networks. The combination of Random Forest, MLP, and other algorithms has significantly improved the quality of attack detection due to their complementary characteristics. Random Forest is efficient in working with large data sets and has a high resistance to overtraining, while MLP can handle non-linear relationships, which is important for complex threats in the cyber environment.

The results demonstrate the following advantages: 1. Increased attack detection accuracy. The model achieved an increase in accuracy of up to 3.35 times, which indicates its ability to better identify threats. 2. Improved classification accuracy. The model's accuracy has increased from 1.05 to 1.48 times, which ensures fewer false positives and false negatives. 3. Increased system sensitivity. The sensitivity of the model, which is a critical indicator for detecting new or subtle attacks, has increased to 1.02 times. 4. Increased predictive capability The system's predictive capability has improved by a factor of 1.04 to 2.18, which confirms its ability to more accurately predict future threats. The results confirm the effectiveness of combining different machine learning algorithms to create a cyberattack detection system capable of operating in a complex and changing civil aviation environment. Such a system enhances the security of information networks by providing an effective response to threats in real-time. The modular architecture of the system allows for the integration of various data sources, pre-processing, and analysis, as well as training and updating of models to adapt to changing conditions. This is critical for aviation systems operating in a dynamic environment with high safety requirements. The integration of analysis and response modules allows for rapid identification of threats and automatic action (e.g., blocking IP addresses), minimizing the impact of cyberattacks. The proposed system architecture can be easily scaled to handle large volumes of traffic and integrated into various information and communication systems. This allows it to be used not only in aviation but also in other critical

infrastructure sectors. The implementation of such a system helps to reduce the risks of cyber-attacks on civil aviation facilities, including DoS, DDoS, XSS, CSRF, and network traffic anomalies. This significantly improves the security of information resources and communication systems, which is critical to ensuring the stable and safe operation of the aviation industry.

Thus, the developed system for detecting and analyzing cyber threats using combined machine learning models is a reliable solution for improving the level of cybersecurity in the information and communication systems of civil aviation facilities. It combines high accuracy, adaptability, and efficiency, making it an effective tool in the fight against modern cyber threats.

Declaration on Generative AI

While preparing this work, the authors used the AI programs Grammarly Pro to correct text grammar and Strike Plagiarism to search for possible plagiarism. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

References

[1] O. Mykhaylova, et al., Mobile application as a critical infrastructure cyberattack surface, in: Workshop on Cybersecurity Providing in Information and Telecommunication Systems II, CPITS-II, vol. 3550 (2023) 29–43.

[2] A. Zahynei, et al., Method for calculating the residual resource of fog node elements of distributed information systems of critical infrastructure facilities, in: Cybersecurity Providing in Information and Telecommunication Systems, CPITS, vol. 3654 (2024) 432–439.

[3] Z. B. Hu, et al., Authentication system by human brainwaves using machine learning and artificial intelligence, in: Advances in Computer Science for Engineering and Education IV (2021) 374–388. doi:10.1007/978-3-030-80472-5_31

[4] V. Zhebka, et al., Methodology for predicting failures in a smart home based on machine learning methods, in: Cybersecurity Providing in Information and Telecommunication Systems, CPITS, vol. 3654 (2024) 322–332.

[5] S. A. Mahboub, E. S. A. Ahmed, R. A. Saeed, Smart IDS and IPS for cyber-physical systems, Artificial Intelligence Paradigms for Smart Cyber-Physical Systems, 2021, 109–136. doi:10.4018/978-1-7998-5101-1.ch006

[6] A. N. Jaber, et al., The importance of IDS and IPS in cloud computing environment: Intensive review and future directions, in: Advances in Cyber Security: Second International Conference, Penang, Malaysia, 2021, 479–491. doi:10.1007/978-981-33-6835-4_32

[7] A. Aldweesh, A. Derhab, A. Z. Emam, Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues, Knowl.-Based Syst. 189 (2020) 105124. doi:10.1016/j.knosys.2019.105124

[8] S. M. Sohi, J. P. Seifert, F. Ganji, RNNIDS: Enhancing network intrusion detection systems through deep learning, Comput. & Secur. 102 (2021) 102151. doi:10.1016/j.cose.2020.102151

[9] R. Achary, Cryptography and network security: An introduction, Mercury Learning and Information, 2021.

[10] I. Anna, I. Sergii, H. Marharyta, A biometric asymmetric cryptosystem software module based on convolutional neural networks, Int. J. Comput. Netw. Inf. Secur. 9(6) (2021) doi:10.5815/ijcnis.2021.06.01

[11] U. Inayat, et al., Learning-based methods for cyber attacks detection in IoT systems: A survey on methods, analysis, and future prospects, Electron. 11(9) (2022) 1502. doi:10.3390/electronics11091502

[12] H. Ahmetoglu, R. Das, A comprehensive review on detection of cyber-attacks: Data sets, methods, challenges, and future research directions, Internet of Things, 20 (2022) 100615. doi:10.1016/j.iot.2022.100615

[13] A. Ilyenko, et al., Prospective directions of traffic analysis and intrusion detection based on neural networks, Cybersecur. Educ. Sci. Tech. 1(17) (2022) 46–56. doi:10.28925/2663-4023.2022.17.4656
[14] A. N. Ozalp, Z. Albayrak, Detecting cyber attacks with high-frequency features using machine learning algorithms, Acta Polytechnica Hungarica, 19(7) (2022).
[15] R. Kyrychok, et al., Development of a method for checking vulnerabilities of a corporate network using Bernstein transformations, Eastern-Europ. J. Enterp. Technol. 1(9(115)) (2022) 93–101. doi:10.15587/1729-4061.2022.253530
[16] A. Oyetoro, J. Mart, U. Amah, Using machine learning techniques random forest and neural network to detect cyber attacks, ScienceOpen Preprints (2023). doi:10.14293/PR2199.000059.v1
[17] A. Kumar, N. Saxena, B. J. Choi, Machine learning algorithm for detection of false data injection attack in power system, in: International Conference on Information Networking (ICOIN), 2021, 385–390. doi:10.1109/ICOIN50884.2021.9333913
[18] A. Nazir, R. A. Khan, A novel combinatorial optimization based feature selection method for network intrusion detection, Comput. Secur. 102 (2021) 102164.
[19] Z. Wang, et al., Intrusion detection methods based on integrated deep learning model, Comput. Secur. 103 (2021) 102177.
[20] D. Dwivedi, et al., Detection of malicious network traffic attacks using support vector machine, in: International Conference on Advanced Network Technologies and Intelligent Computing, 2023, 54–68.
[21] P. Semwal, A. Handa, Cyber-attack detection in cyber-physical systems using supervised machine learning, Handb. Big Data Anal. Forensics, (2022) 131–140. doi:10.1007/978-3-030-74753-4_9
[22] J. Gu, et al., A novel approach to intrusion detection using SVM ensemble with feature augmentation, Comput. Secur. 86 (2019) 53–62.
[23] R. D. Reddy, S. Katkam, C. R. S. Rao, Cyber attacks detection using machine learning, NeuroQuantology, 20(19) (2022) 4388.