DOI: https://doi.org/10.15276/aait.08.2025.13
UDC 004.056.5

A hybrid method for detecting anomalous traffic in computer networks

Yurii P. Klots[1)]
ORCID: https://orcid.org/0000-0002-3914-0989; klots@khmnu.edu.ua. Scopus Author ID: 6504043018
Nataliia S. Petliak[1)]
ORCID: https://orcid.org/0000-0001-5971-4428; npetlyak@khmnu.edu.ua. Scopus Author ID: 57786856200
Vira Yu. Titova[1)]
ORCID: https://orcid.org/0000-0001-8668-4834; titovav@khmnu.edu.ua. Scopus Author ID: 57786263500
[1)] Khmelnytskyi National University, 11, Instytuts'ka Str., Khmelnytskyi, 29016, Ukraine

ABSTRACT

This study addresses the increasing difficulty of detecting anomalies in network traffic caused by growing threats to information and communication systems. Traditional intrusion detection systems often fail to adapt to new threats, particularly when analyzing outbound traffic, which may signal internal compromise. To overcome these limitations, the study proposes a hybrid detection method aimed at improving anomaly identification accuracy. The method integrates three components. First, traffic is classified using a signature-based approach with predefined sets of allowed and prohibited signatures. Second, self-similarity analysis with the Hurst coefficient detects long-term traffic patterns. Third, fuzzy logic is applied to interpret uncertain traffic characteristics, such as port numbers, protocols, intensity, and packet sizes, using linguistic variables and fuzzy rules. The research presents formalized models of both legitimate and malicious user behavior and a composite packet signature model for comprehensive traffic analysis. This approach enhances adaptability and reduces the proportion of unclassified traffic. Experimental validation using real and synthetic data confirms improved detection accuracy and a lower false positive rate compared to conventional methods. The scientific novelty lies in combining deterministic classification with fuzzy logic within a single detection pipeline, with a special emphasis on outbound traffic monitoring. The practical value of the proposed system is its suitability for integration into existing cybersecurity frameworks, contributing to more effective threat detection and reduced operational risks in evolving network environments.

Keywords: Anomaly detection; hybrid method; network traffic; packet signature; fuzzy logic; self-similarity; behavioral model; outbound traffic; intrusion detection; cybersecurity

***For citation*: Klots Yu. P., Petliak N. S., Titova V. Yu. "A hybrid method for detecting anomalous traffic in computer networks". *Applied Aspects of Information Technology.* 2025; Vol. 8 No. 2: 191–201. DOI: https://doi.org/10.15276/aait.08.2025.13**

INTRODUCTION

In today's digital environment, cybersecurity threats are becoming increasingly relevant and complex. Every day, Internet users face numerous attacks on information systems, which can lead to financial losses, loss of confidential data, and privacy violations. According to statistics, annual losses from cyberattacks amount to billions of dollars, and this figure continues to grow. Among the most common threats are DDoS attacks, phishing, malware, and password mining.

Traditional intrusion detection systems have a number of limitations, including limited effectiveness in a changing environment and difficulty adapting to new, previously unknown types of attacks [1], [2].

They either tend to miss malicious activity (the second type of error) or, on the contrary, wrongly classify legitimate users as intruders (the first type of error). One of the reasons for this contradiction is subjectivity in determining the criteria for malicious behavior.

For example, a user who repeatedly entered a password incorrectly, downloaded a document for the wrong purpose, or connected a third-party device may be mistakenly identified as an attacker.

To overcome these limitations, it is important to implement modern methods of network traffic processing based on machine learning, fuzzy logic, and deep data analysis. Such approaches help to improve the accuracy of traffic classification, detect anomalous activities in a timely manner, and adapt to new types of threats. With the growing number of users, the volume of transmitted data, and the complexity of network

architectures, it is of particular importance to improve traffic monitoring and analysis tools to ensure the stable and secure operation of information systems.

Modern network infrastructures are increasingly dynamic and distributed, incorporating cloud technologies, mobile devices, and IoT components. These factors introduce additional vectors for attacks and complicate the process of traffic inspection. Security systems must be scalable and capable of

processing large volumes of heterogeneous data in real time. The integration of intelligent monitoring solutions makes it possible to detect deviations from normal behavior patterns, thereby identifying both known and previously unseen attack scenarios.

At the same time, an important aspect of cybersecurity strategy is continuous monitoring of the internal and outbound traffic of information systems. While perimeter protection is essential, many modern attacks exploit lateral movement or data exfiltration techniques that are not detected by perimeter-focused tools.

ANALYSIS OF LITERARY DATA

Choosing the right methods for detecting anomalies is an important factor in achieving high efficiency and accuracy of analysis. Anomalies may indicate information security incidents, such as equipment failure, unauthorized access, or cyber-attacks, so their detection requires considering the specifics of the analyzed data and system features [3]. Depending on the nature of the anomalies, the type of data, and the available computing resources, different methods are used. These methods are classified according to a number of criteria: the type of training (supervised, semi-supervised, or unsupervised), the processing approach (statistical, machine learning, or graph analysis), and the mode of operation (real-time or offline). Statistical approaches based on probabilistic models and analysis of deviations from the normative distribution is well suited for predictive data, but may not be effective in complex or dynamic environments.

Among the characteristics used for anomaly detection are traffic frequency and intensity, distribution of values in the feature space, inter-parameter relationships, and noise level. High data quality and pre-processing, such as cleaning and normalization, significantly improve detection accuracy. When choosing parameters, you should also consider the specifics of the task. For example, the focus on detecting one-time deviations or new types of anomalies.

The process of developing detection methods should consider the possible risks associated with false positives and false negatives, because in critical areas such as cybersecurity or complex systems management, the balance between sensitivity and specificity is crucial.

Automation of procedures and the ability of systems to adapt to changes in data behavior through self-learning also increase the effectiveness of anomaly detection. The study by Aklil Kiflay et al [4] proposes a method for detecting network attacks based on a multimodal system using machine learning. It is based on two separate models with the Random Forest algorithm that analyze the stream metadata and the first 32 bytes of the protocol, and the results are combined according to the principle of soft voting. Updating the rules or algorithms is only possible through retraining, and the limitation is the difficulty of processing large volumes of traffic.

Paper [5] presents an approach to intrusion detection using autonomous fuzzy logic and the concept of “soft prototypes”. The proposed system Soft Prototype-based Autonomous Fuzzy Inference System (SPAFIS) has the ability to adapt to new data in real time and change its structure without external intervention. The system automatically adds new prototypes, eliminates outdated or duplicate ones, thereby optimizing resource utilization. The method ensures efficient work with streaming traffic and scales well for large networks while maintaining high performance. Although the configuration requires the participation of specialists, the basic parameters can be easily adapted to specific conditions.

Paper [6] discusses a hybrid system for detecting network intrusions that combines machine learning methods with a self-healing mechanism. The basic concept is to integrate signature analysis and anomaly detection, which increases the efficiency and accuracy of attack detection. An important feature of the system is its ability to automatically update the signature database, including adding new rules for unknown threats. Optimized algorithms reduce the load on resources, as well as provide scalability and adaptability to user needs.

Paper [7] describes an approach to detecting and preventing zero-day attacks in networks using machine learning. The analysis of network traffic combined with the application of Benford's law allows detecting atypical parameters that may indicate abnormal activity. The proposed methods are characterized by adaptability, scalability, and compatibility with various operating systems and security systems. Particular attention is paid to semi-automated model training on limited sets of labeled data, which simplifies the implementation of new detection rules.

The study by Mahmoud Said El Sayed et al. [8] proposed an anomaly detection method to counter Distributed Denial of Service (DDoS) attacks in the environment of software-defined networks. The focus is on the use of deep learning methods that analyze traffic behavior to detect threats. Efficiency is ensured through the use of feature selection procedures such as Information Gain and Random

Forest, which reduce the number of parameters without losing accuracy.

The authors of [9] propose to improve intrusion detection systems by integrating signature and anomaly approaches. Such a system analyzes traffic in real time, classifying packets as normal or malicious. Administrators have access to detailed reports and logs, which allows them to respond quickly to detected security incidents.

The study by Sivasankari Nitiynandan [10] considers methods for preventing man-in-the-middle attacks in Internet of Things (IoT) networks by applying regression analysis. Three approaches are proposed: linear regression, multivariate linear regression, and regression based on Gaussian processes. To model traffic, the Network simulator tool was used, which allowed generating both normal and malicious traffic.

Publication [11] describes a hybrid intrusion detection system that combines fuzzy inference, artificial neural networks, and genetic algorithms. The system classifies incoming traffic in real time, which allows for the rapid detection of abnormal behavior patterns. The solution is compatible with various operating systems and can be integrated with other cybersecurity tools.

Radivilova [12] analyzed statistical approaches to detecting anomalies in telecommunications traffic. Methods of time series analysis, decision trees, clustering, and entropy analysis were considered. To assess the effectiveness, such indicators as detection probability, false positive rate, and processing speed were considered.

The authors of [13] have developed a system for detecting and counteracting anomalies in software-defined networks based on a combination of Long Short-Term Memory (LSTM) models for predicting normal traffic and fuzzy logic for identifying anomalies. The approach is based on the formation of a digital signature of network segments based on the characteristics of flows. The system works with current traffic without the need for historical data and determines the limits of normality using the Bienamé-Chebyshev inequality. Thanks to fuzzy logic, an adaptive assessment of anomalous deviations is possible, which allows detecting, for example, DDoS or port scanning attacks. To mitigate the impact of attacks, a dynamic flow control policy is applied to block or redirect suspicious packets while maintaining network stability.

The method [14] based on machine learning for network traffic classification has both significant advantages and limitations. The best performing random forests provide a balanced accuracy of over 87%.

In [15], the FuzHD++ method is described, which combines the recovery of lost data and the detection of anomalous nodes in the IoT environment. Recovery is performed using a matrix profile that identifies repeated patterns in the data between nodes, and gaps are filled using the k-nearest neighbor algorithm. Abnormal nodes are detected using fuzzy rules based on expert knowledge and hidden patterns. At the same time, the method is limited in its ability to identify only certain types of attacks; in particular, data integrity attacks Factitious disorder imposed on another (FDIA).

Liangchen Chen [16] proposed a cluster-based DPC-GS-MND approach for detecting anomalies in network traffic, based on a peak density clustering algorithm improved by considering the mutual degree of neighborhood. This allows for more accurate identification of cluster centers based on local connections between data. Automated selection of centers reduces the influence of the human factor, but the method has difficulty processing unbalanced data and increases computation time due to the need for neighborhood analysis.

In [17], a method for detecting anomalies in cyber-physical systems is presented that integrates the least squares algorithm with entropy analysis of time windows. The least squares algorithm is modified with Gaussian functions and fractional functions to achieve sparsity of state parameters and avoid overfitting. When an anomaly is detected, the data is analyzed in reverse order using an entropy approach to identify the source of the disturbance.

The method presented in [18] uses behavioral analysis to model normal network activity and identify potential anomalies. The solution is based on a combination of forecasting algorithms: simple and double exponential smoothing and the Holt-Winters method, which are integrated into an adaptive model that can adjust to seasonal changes in traffic.

Spinareva and co-authors [19] proposed an approach based on a cascade of deep neural networks to detect and classify network attacks. The architecture includes a hybrid model with a convolutional and recurrent network LSTM for the detection phase. The method is tested on the UNSW-NB15 dataset and real networks. Despite its effectiveness, the model is difficult to set up, requires large computing resources, and does not always provide real-time processing.

A hybrid system for detecting and analyzing Zero-Day threats is described in [20]. It combines

components of automatic decision-making, static analysis with a multiscanner, and dynamic analysis in an isolated environment Sandbox. The approach includes traffic analysis, memory dumps, and the creation of custom rules for detecting anomalies. Integration with antivirus systems provides parallel analysis, but the high cost of resources for running an isolated environment and the complexity of configuration can limit its use. Additionally, system efficiency may be reduced due to the limitations of Yara rules when detecting polymorphic threats.

Study [21] proposes a centralized distributed system for detecting attacks in corporate networks based on multifractal analysis. The method of maximum wavelet transform modules is used to identify signal features and calculate the Hurst coefficient to assess self-similarity. Limitations include the dependence of accuracy on the amount of available data and the need to reconfigure the self-similarity parameters when changing the scale of analysis, which complicates implementation in a dynamic network environment.

Modern information and communication systems are constantly exposed to threats, in particular due to abnormal network traffic, which may be the result of cyber incidents, malicious actions, or malfunctioning of system components. Existing anomaly detection systems are mostly based on fixed signatures or statistical models, which does not allow for effective detection of new, unknown or modified threats, including Zero-Day attacks. A significant problem is the limited ability of such systems to adapt to traffic dynamics and changes in user behavior. In addition, the high number of false positives makes it difficult to respond quickly. Given these challenges, it is important to develop more flexible and reliable methods that combine different approaches to anomaly detection and consider the behavioral characteristics of both legitimate users and potential intruders.

PURPOSE AND OBJECTIVES OF THE STUDY

The aim of the work is to develop a hybrid method for detecting anomalous traffic in information and communication systems, considering the behavioral characteristics of both legitimate users and potential attackers in order to increase the reliability of detecting anomalies in the network environment.

The study envisages developing a formalized model of the network packet signature that reflects the key parameters for traffic identification; implementing the process of classifying network traffic by features, the process of classifying traffic based on the analysis of its self-similarity, developing a process of fuzzy anomaly detection; and proposing a comprehensive analysis method that combines all the previous components into a single hybrid detection system.

A HYBRID METHOD FOR DETECTING ANOMALOUS NETWORK TRAFFIC AND EXPERIMENTAL BASE

The analysis of attacks allowed us to identify the parameters of traffic packets that should be used to identify them. However, a large number of parameters slows down traffic analysis and increases the load on the equipment, so the most significant parameters were selected using the Pareto principle. Based on the data obtained, we will form a packet signature to detect anomalous traffic.

Let's represent it as a multi-component tuple [22]

$$d =< d^1, d^2, \dots, d^k >, \quad (1)$$

where k is the number of components in the tuple; d^1 is the source IP address that sends a request for connection and information exchange; d^2 is the destination IP address, i.e., to which IP address requests are sent; d^3 is the source port used to establish the connection; d^4 is the destination port; d^5 is the protocol used for data transmission; d^6 is the traffic intensity, which is determined in bits/s, d^7 is the time the packet arrives for inspection, d^8 is the MAC address of the device that sends data from the network, d^9 is the packet size.

The model of traffic classification by features is represented as follows:

$$MO =< D, D^G, D^B, D^U >, \quad (2)$$

where D is the set of incoming traffic; D^G is the set of allowed signatures; D^B is the set of forbidden signatures, and D^U is the set of undefined signatures.

The set of incoming traffic is represented as follows:

$$D = \bigcup_{i=0}^{N_d} \{d_i\}, \quad (3)$$

where d_i is the element of the set, the input generated signature.

The process of traffic classification based on signature analysis is presented in the form of a partition of the set D. According to this partitioning, if the signature of the analyzed packet contains source and destination IP addresses that are in the set D^G, then such a signature is transferred to the set D^G:

$$D^G = \begin{Bmatrix} If\ \exists d_i \in D^G\ such\ that d_i^1 = \\ = d_j^1 \wedge d_i^2 = d_j^2 \end{Bmatrix}, \quad (4)$$

where d_j is the signature of the incoming traffic being analyzed.

If the signature of the analyzed packet contains source or destination IP addresses that are in set D^B, then such a signature is transferred to set D^B:

$$D^B = \begin{Bmatrix} If\ \exists d_i \in D^B\ such\ that\ d_i^1 = \\ = d_j^1 \vee d_i^2 = d_j^2 \end{Bmatrix}. \quad (5)$$

In all other cases, the signature is added to the set D^U :

$$D^U = \{If\ d_j \notin D^G \wedge d_j \notin D^B\} \quad (6)$$

Since the sets D^G, D^B, and D^U are partitions of D, the following statements are true:

$$D = D^G \cup D^B \cup D^U, \quad (7)$$

The use of self-similarity allows new signatures to be categorized, reducing the need to validate the signature according to a fuzzy process for detecting anomalous traffic. The main advantage of self-similarity is its ability to detect consistent patterns in traffic, even if it changes over time. By analyzing the time windows of the data flow, it is possible to use previously obtained data to verify new data, which allows you to adaptively update the set D without the need to completely revise the entire data set. The Hearst measure is used to quantify the degree of self-similarity of a data stream. It makes it possible to calculate the level of similarity between signatures, which allows you to accurately assess whether new packets should be assigned to the already known set of D or whether they require additional analysis. This is important for detecting cyber threats, as the speed and accuracy of traffic classification plays a significant role in cyber defense. Self-similarity can significantly reduce the amount of unidentified traffic that requires further analysis, which simplifies the process of monitoring network traffic and improves the overall performance of the anomaly detection system.

A model of the traffic classification process based on self-similarity:

$$MS = < D, D^G, D^U, f(T), t_1, t_2, t_3, M, M' >, \quad (8)$$

where $f(T)$ is the similarity function, t_1 is the beginning of the analysis time interval, t_2 is the moment that separates the previously analyzed signatures and the new ones that have just appeared, t_3 is the end of the analysis interval, M is the set of previously analyzed signatures for self-similarity, M' is the set of new signatures that have not been analyzed for self-similarity (Fig).

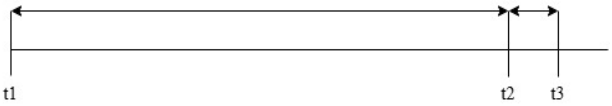

Fig. **Time intervals for traffic classification**
Source: **compiled by the authors**

Let's define the set M of signatures previously analyzed for self-similarity

$$M = \{d_i \in D^G \cup D^U \mid t_1 < d_i^7 \leq t_2\}, \quad (9)$$

and a set M' of new signatures that have not been previously analyzed for self-similarity

$$M' = \{d_i \in D^G \cup D^U \mid t_2 < d_i^7 \leq t_3\} \quad (10)$$

Let's define the similarity function

$$y = f(T), \quad (11)$$

where T is a set of signatures, y is a numerical measure of similarity, $y \in [0,1]$.

In order to determine the similarity of a packet signature to the set of previously defined packet signatures, we set the standard deviation function $S(n)$:

$$S^j(n) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(d_i^j - \overline{d^J})^2}, \quad (12)$$

where $n = |T|$ is the number of signatures in the set T, j is the number of the component in the tuple, d_i^j is the value of the jth component of the i-th tuple of the set, $\overline{d^J}$ is the average value of the jth component of the tuples of the set.

And the function $R(n)$ is the range of the accumulated deviation of the function.

$$R^j(n) = \max_{i=1,n} \sum_{m=1}^{k} d_i^j - \overline{d^J} - \min_{i=1,n} \sum_{m=1}^{k} d_i^j - \overline{d^J}, \quad (13)$$

$$k = 1, n$$

Let's define the Hurst measure for each component of the tuple:

$$H^j = \frac{\ln\left(R^j(n) / S^j(n)\right)}{\ln(n)} \quad (14)$$

We define the Hearst measure for a packet signature as the average value for each component of a tuple:

$$H = f(T) = \frac{1}{k}\sum_{j=1}^{k} H^j, \quad (15)$$

where j is the number of components in the tuple.

The process of classifying uncertain traffic based on signature similarity analysis is represented as follows:

$$D^G = \{D^G \cup M' | \, f(M \cup M') > 0{,}5\}, \quad (16)$$

$$D^U = \{D^U \backslash M' | \, f(M \cup M') > 0{,}5\}, \quad (17)$$

Upon completion of the analysis, the set D^U will contain signatures that are undefined or different from the allowed signatures and require further analysis.

In an ever-changing network environment, classical methods of detecting anomalous traffic are often not effective enough due to strict decision-making criteria. In such circumstances, there is a need for an approach that can consider uncertainty, fuzzy boundaries between normal and abnormal traffic, and subjective expert judgment. This is why the hybrid method implements a fuzzy detection model. It allows for more flexible traffic analysis, generating estimates based on linguistic variables and rules that mimic the human thinking process. This is especially important when processing uncertain or new signatures that cannot be unambiguously classified by traditional methods.

The model of fuzzy anomalous traffic detection, consisting of the sets D^G, D^B, D^U, a set of linguistic variables and rules, is presented as follows:

$$MN = < D^U, L, PR, D^G, D^B >, \quad (18)$$

where $L = \{ld^2, ld^4, ld^5, ld^6, ld^9\}$ are linguistic variables; *PR* are rules.

It should be noted that the first component of the tuple, the source IP address (d^1), the third component, the source port (d^3), the seventh component, the time of arrival of data for inspection (d^7), and the eighth component, the MAC address of the device (d^8), are not used for the fuzzy anomaly detection process model. The source IP address (d^8), which identifies the device that initiates the connection and transmits the data, is not a reliable indicator for analysis, as it can change due to the use of dynamic IP, NAT or proxy servers, as well as spoofing. The source port (d^3), which determines the incoming communication point on the sender's side, is usually randomly generated by the operating system, making it unpredictable and of little importance in the context of fuzzy anomaly classification. The packet arrival time (d^7), expressed in 24-hour format, depends on many external factors such as network delays, peak loads, and changes in traffic depending on the time of day. This creates a significant level of variability, which can make it difficult to correctly fuzzify anomalous patterns. The MAC address of the (d^8) device sending the data is also not used due to its local nature, as this address is not transmitted through routed networks and can be easily spoofed or changed at the device level. The exclusion of these parameters is aimed at improving the accuracy of the model by eliminating variables that may introduce unnecessary noise. Thus, only those characteristics that provide stability and high differentiation between normal and potentially threatening traffic remain for fuzzy analysis.

The second component in the tuple is the destination IP address (d^2), which can be displayed as a character variable that takes one of the values of the set of identifiers $ld^2 \in \{ld_1^2, ld_2^2, ld_3^2\}$, namely

ld_1^2 = "IP address of allowed traffic",
ld_2^2 = "IP address of unspecified traffic",
ld_3^2 = "IP address of abnormal traffic".

The fourth component in the tuple is the destination port (d^4), which can be displayed as a symbolic variable that takes one of the values of the set of identifiers $ld^4 \in \{ld_1^4, ld_2^4, ld_3^4\}$, namely

ld_1^4 = "port of normal typical traffic",
ld_2^4 = "port of undefined traffic",
ld_3^4 = "port of abnormal traffic".

The fifth component in the tuple is the protocol (d^5), which can be displayed as a symbolic variable that takes one of the values of the set of identifiers $ld^5 \in \{ld_1^5, ld_2^5, ld_3^5\}$, namely

ld_1^5 = "safe protocols",
ld_2^5 = "neutral protocols",
ld_3^5 = "uncertain protocols",
ld_4^5 = "suspicious protocols",
ld_5^5 = "dangerous protocols".

Traffic intensity in ICS (d^6) can be considered from two approaches: objective and subjective. The objective traffic intensity is defined as the ratio of the amount of information transmitted through the communication channel for a certain period of time to the total available network resource. This approach is based on measuring and analyzing the actual performance of the system, for example, the number of packets or bytes that pass through the system per unit of time.

Subjective traffic intensity reflects the assessment of experts or users of the system regarding the level of congestion of communication channels at a particular moment or over a certain period. It is based on perceptions or estimates of system performance, as well as on forecasts of system performance depending on the current or

expected volume of information transmitted. In situations where obtaining accurate statistics is difficult, experts use a logical-linguistic approach. In this case, the traffic intensity is represented by a linguistic variable (ld^6) with a basic term set, which allows formalizing qualitative assessments and ensuring their convenient interpretation for further analysis:

$$ld^6 = \bigcup_{i=1}^{m} ld_i^6, \quad (19)$$

where m is the number of terms for which the ratio of the order $ld_1^6 < ld_2^6 < \cdots < ld_m^6$ is valid.

The ninth component in the tuple is the packet size (d^9).

This component can be displayed numerically or through a linguistic variable:

$$ld^9 = \bigcup_{i=1}^{y} ld_i^9, \quad (20)$$

where y is the number of terms for which the ratio of the order $ld_1^9 < ld_2^9 < \cdots < ld_y^9$.

The membership function was chosen as trapezoidal because it can be used to model different states of network traffic, which allows for more detailed analysis and classification of traffic behavior, which is important for detecting anomalies and ensuring network security. The parameters for each set of values are set based on expert opinions.

The hybrid method of detecting anomalous traffic combines the advantages of several approaches – signature, behavioral, and fuzzy analysis – to achieve high accuracy, adaptability, and reduce the number of false positives. The input data for the method are packets from the original data stream, a set D, a set D^G, a set of prohibited signatures D^B, a set D^U, time intervals of the reference and verification samples, and a list of rules.

The steps of the method are as follows.

Step 1. Data initialization.

Step 2. Formation of the packet signature d_j from the original data stream (D).

Step 3. Performing the traffic classification process by features. If the signature is classified, then the transition to step 7 occurs. Otherwise, you go to step 4.

Step 4. Perform the traffic classification process based on self-similarity. If the signature is classified, you go to step 7. Otherwise, you go to step 5.

Step 5. Performing a fuzzy method for detecting anomalous traffic. Go to step 6.

Step 6. If you want to continue checking the signatures, you go to step 2. Otherwise, go to step 7.

Step 7. Finish the work.

The structural model of the anomalous traffic detection system in an information and communication system implements a full cycle of network traffic processing – from raw traffic capture to deciding on the security of the connection. At the first stage, a specialized traffic capture module is integrated with the network environment and performs pre-processing, which includes traffic filtering by features, data format normalization and packet aggregation in a session, forming the basis for further analysis. The obtained data is transmitted to the traffic classification module by features, which performs signature identification according to the sets of allowed and prohibited connections. The signature database is dynamically updated in real time, which allows you to detect both known threats and modified or new attack variants. At the same time, a traffic classification module based on self-similarity is activated, which compares current characteristics with models of typical user and attacker behavior, detecting anomalous deviations, in particular zero-day attacks. The main element of the method is a fuzzy model for detecting anomalous traffic, which uses linguistic features and a membership function to form an integrated risk indicator. The system operates on a fuzzy logic rule base, which allows forming a generalized assessment of the threat level considering such parameters as transmission speed, number of connection attempts, signature matching, etc. The inference subsystem combines the results of the classification, self-similarity, and fuzzy logic modules, providing a comprehensive decision-making mechanism. In case of detection of uncertain or contradictory situations, the system can direct the flow for additional verification, and at an increased threat level, initiate interaction with other components of cyber protection, in particular firewalls and intrusion

prevention systems. All events recorded during traffic processing are recorded in a centralized logging system, which includes IP addresses, ports, time stamps, and protocol types. This data is used not only for audit and incident analysis, but also for feedback, which allows refining classification rules and updating self-similarity models. The proposed hybrid method for anomaly detection is characterized by a high degree of adaptability to changes in the traffic of the information and communication system, a reduction in false positives and optimization of computing resources. Special attention is paid to the analysis of outgoing traffic as a separate threat vector in order to detect internal attacks or data leaks.

RELIABILITY OF A HYBRID METHOD FOR DETECTING ABNORMAL TRAFFIC

General requirements for the efficiency of solving the problem solved by an anomaly detection system can be expressed through the following well-known and frequently used quality metrics: True Positive (TP) – the number of events that are correctly classified as anomalous and are in fact so; False Positive (FP) – the number of events that were mistakenly identified as anomalous, but are not in fact so; True Negative (TN) – the number of events that were not classified as anomalous and are not in fact so (i.e., are normal); False Negative (FN) – the number of events that were not identified as anomalous, but are in fact anomalous [23, 24]. FPs are often called first-order errors, and FNs are called second-order errors.

After setting up the test environment, we ran the generated dataset as input for traffic classification with and without Snort. The quality metrics are shown in Table 1.

Table 1 illustrates the results of testing the three systems – Snort with hybrid library, original Snort, and Suricata – on the same artificially generated set of 5468657 records (of which 307672 are anomalous and 5160985 are normal). For Snort with the hybrid library, the number of correctly detected anomalies (TP) was 5141567, a missed anomaly (FN) was 19418, a false positive for normal traffic (FP) was 28739, and correctly classified normal events (TN) was 278933. In the case of the original Snort, these figures were 5027861 (TP), 133124 (FN), 46235 (FP), and 261437 (TN), and in Suricata, 5064128 (TP), 96857 (FN), 38138 (FP), and 269534 (TN). The comparison demonstrates that the hybrid method significantly reduces the number of missed anomalies and false positives compared to both traditional solutions.

Table 2 shows the key quality indicators of the systems in the laboratory, calculated on the basis of the data from Table 1. The Accuracy of the hybrid module reached 99.12 %, Precision was 99.44 %, Specificity was 90.66 %, Recall was 99.62 %, and F1 measure was 99.53 %. In comparison, the original Snort had an Accuracy of 96.72 %, Precision of 99.09%, Specificity of 84.97 %, Recall of 97.42 %, and F1 of 98.25 %, while Suricata showed 97.53%, 99.25%, 87.60%, 98.12% and 98.68%, respectively. These values demonstrate the superiority of the hybrid approach in all five metrics.

CONCLUSIONS

This paper proposes a hybrid method for detecting anomalous traffic that combines signature analysis, self-similarity, and fuzzy logic. The method is based on formalized models of user and intruder behavior, as well as on a multicomponent packet signature. The implementation of the three-stage classification allowed to increase the accuracy of threat detection and reduce the amount of uncertain traffic. Experimental results have shown the superiority of the method over traditional systems (Snort, Suricata) in terms of accuracy, completeness, and F1-measure. The novelty of the work is the integration of deterministic and fuzzy models with a focus on outbound traffic; practical value is the possibility of implementing into existing cyber defense architectures.

Table 1. **Testing with the dataset**

Dataset	Number of abnormal records	Number of records allowed	Total records	TP	TN	FP	FN
Snort with its own module	307672	5160985	5468657	5141567	278933	28739	19418
Snort				5027861	261437	46235	133124
Suricata				5064128	269534	38138	96857

Source: **compiled by the authors**

Table 2. **Characteristics of reliability assessment using the dataset**

Type of traffic	metric / system	Accuracy	Precision	Specificity	Recall	F1-score
Dataset	Snort with its own module	99.12 %	99.44 %	90.66 %	99.62 %	99.53 %
	Snort	96.72 %	99.09 %	84.97 %	97.42 %	98.25 %
	Suricata	97.53 %	99.25 %	87.60 %	98.12 %	98.68 %

Source: **compiled by the authors**

REFERENCES

1. Awan, M. J, Farooq, U, Babar, H. M. A., Yasin, A., Nobanee, H, Hussain, M., Hakeem, O. & Zain, A. M. "Real-time DDoS attack detection system using big data approach". *Sustainability*. 2021; 13 (19): 10743, https://www.scopus.com/record/display.uri?eid=2-s2.0-85116006755&origin=resultslist. DOI: https://doi.org/10.3390/su131910743.

2. Shafiq, M., Tian, Z., Bashir, A. K., Du, X. & Guizani, M. "CorrAUC: A malicious bot-IoT traffic detection method in iot network using machine-learning techniques". In *IEEE Internet of Things Journal*. 2021; 8 (5): 3242–3254, https://www.scopus.com/record/display.uri?eid=2-s2.0-85101711373& origin=resultslist. DOI: https://doi.org/10.1109/JIOT.2020.3002255.

3. Klots, Y., Titova, V., Petliak, N., Cheshun, V. & Salem, A.-B.M. "Research of the neural network module for detecting anomalies in network traffic". *CEUR Workshop Proceedings*. 2022; 3156: 378–389, https://www.scopus.com/record/display.uri?eid=2-s2.0-85133586586&origin=resultslist.

4. Kiflay, A., Tsokanos, A., Fazlali, M. & Kirner, R. "Network intrusion detection leveraging multimodal features". *Array*. 2024; 22: 100349, https://www.scopus.com/record/display.uri?eid=2-s2.0-85194046213&origin=resultslist. DOI: https://doi.org/10.1016/j.array.2024.100349.

5. Gu, X., Howells, G. & Yuan, H. "A soft prototype-based autonomous fuzzy inference system for network intrusion detection". *Information Sciences*. 2024; 677: 120964, https://www.scopus.com/record/display.uri?eid=2-s2.0-85195662655&origin=resultslist. DOI: https://doi.org/10.1016/j.ins.2024.120964.

6. Kushal, S., Shanmugam, B., Sundaram, J. et al. "Self-healing hybrid intrusion detection system: an ensemble machine learning approach". *Discover Artificial Intelligence 4*. 2024; 28, https://www.scopus.com/record/display.uri?eid=2-s2.0-85190784502&origin=resultslist. DOI: https://doi.org/10.1007/s44163-024-00120-9.

7. Mbona, I. & Eloff, J. H. P. "Detecting zero-day intrusion attacks using semi-supervised machine learning approaches". *IEEE Access*. 2022; 10: 69822–69838, https://www.scopus.com/record/display.uri?eid=2-s2.0-85133746314&origin=resultslist. DOI: https://doi.org/10.1109/ACCESS.2022.3187116.

8. Le-Khac, M. S. E., Sayed, N.-A., Azer, M. A. & Jurcut, A. D. "A flow-based anomaly detection approach with feature selection method against DDoS attacks in SDNs". *IEEE Transactions on Cognitive Communications and Networking*. 2022; 8 (4): 1862–1880. DOI: https://doi.org/10.1109/TCCN.2022.3186331.

9. Shaikh, A. & Gupta, P. "Advanced signature-based intrusion detection system". In *Rajakumar, G., Du, KL., Vuppalapati, C., Beligiannis, G. N. (eds) "Intelligent Communication Technologies and Virtual Mobile Networks". Lecture Notes on Data Engineering and Communications Technologies*. *Springer, Singapore*. 2023; 131. DOI: https://doi.org/10.1007/978-981-19-1844-5_24.

10. Sivasankari, N. & Kamalakkannan. S. "Detection and prevention of man-in-the-middle attack in iot network using regression modeling". *Advances in Engineering Software*. 2022; 169: 103126. DOI: https://doi.org/10.1016/j.advengsoft.2022.103126.

11. Ishaque, M., Johar, Md G. Md, Khatibi, A. & Yamin, M. "A novel hybrid technique using fuzzy logic, neural networks and genetic algorithm for intrusion detection system". *Measurement: Sensors*. 2023; 30: 100933. DOI: https://doi.org/10.1016/j.measen.2023.100933.

12. Radivilova, T., Kirichenko, L., Tawalbeh, M., & Ilkov, A. "Detection of anomalies in the telecommunications traffic by statistical methods". *Electronic Professional Scientific Journal «Cybersecurity: Education, Science, Technique»*. 2021; 3 (11): 183–194. DOI: https://doi.org/10.28925/2663-4023.2021.11.183194.

13. Novaes, M. P., Carvalho, L. F., Lloret, J. & Proença. M. L. "Long short-term memory and fuzzy logic for anomaly detection and mitigation in software-defined network environment". *Access.* 2020; 8: 83765-83781. DOI: https://doi.org/10.1109/ACCESS.2020.2992044.

14. Canavese, D., Regano, L., Basile, C., Ciravegna, G. & Lioy, A. "Encryption-agnostic classifiers of traffic originators and their application to anomaly detection." *Computers & Electrical Engineering*. 2022; 97: 107621. DOI: https://doi.org/10.1016/j.compeleceng.2021.107621.

15. Berjab, N., Le H. H. & Yokota, H. "Recovering missing data via top-k repeated patterns for fuzzy-based abnormal node detection in sensor networks". *IEEE Access.* 2022; 10: 61046–61064. DOI: https://doi.org/10.1109/ACCESS.2022.3181742.

16. Chen, L., Gao, S. & Liu, B. “An improved density peaks clustering algorithm based on grid screening and mutual neighborhood degree for network anomaly detection”. *Scientific Reports*. 2022; 12: 1409. DOI: https://doi.org/10.1038/s41598-021-02038-z.

17. Zhang, J., Yuan, Y., Zhang, J., Yang, Y. & Xie, W. “Anomaly detection method based on penalty least squares algorithm and time window entropy for Cyber–Physical Systems”. *Journal of King Saud University – Computer and Information Sciences*. 2023; 35: 101860. DOI: https://doi.org/10.1016/j.jksuci.2023.101860.

18. Pelc, M., Galus, D., Zolubak, M., Ozana, S., Chlewicki, W., Cichon, K., Podpora, M. & Kawala-Sterniuk, A. “Behavioral approach to network anomaly detection for resource-constrained system – presentation of the novel solution – preliminary study”. *IFAC-PapersOnLine*. 2019; 52 (27): 121–126. DOI: https://doi.org/10.1016/j.ifacol.2019.12.743.

19. Shpinareva, I. M., Yakushina, A. A., Voloshchuk, L. A., & Rudnichenko, N. D. “Detection and classification of network attacks using the deep neural network cascade”. *Herald of Advanced Information Technology*. 2021; 4 (3): 244–254. DOI: https://doi.org/10.15276/hait.03.2021.4.

20. Saprykin, O. S. “Models and methods for diagnosing zero-day threats in cyberspace”. *Herald of Advanced Information Technology*. 2021; 4 (2): 155–167. DOI: https://doi.org/10.15276/hait.02.2021.5.

21. Shagin B., Nicheporuk A., & Kashtalian A. “Centralized distributed attack detection system in corporate computer networks based on multifractal analysis.” *Measuring and computing devices in technological processes.* 2021; 1: 50–55. DOI: https://doi.org/10.31891/2219-9365-2021-67-1-7.

22. Petliak N. “Hybrid method and system for detecting abnormal traffic in information and communication systems”. *Herald of Khmelnytskyi national university*. 2025; 2 (349): 561–569. DOI: https://doi.org/10.31891/2307-5732-2025-349-82.

23.Titova V., Klots Y., Petliak N., Cheshun V. & Salem A.-B. M. “Detection of network attacks in cyber-physical systems using a rule-based logical neural network”. *1st International Workshop on Intelligent and CyberPhysical Systems (ICyberPhyS)*. Khmelnytskyi, Ukraine. 2024; 3736: 255–268. – Available from: https://ceur-ws.org/Vol-3736/paper19.pdf.

24. Klots, Y., Petliak, N. & Titova, V. “Evaluation of the efficiency of the system for detecting malicious outgoing traffic in public networks”. *13th International Conference on Dependable Systems, Services and Technologies (DESSERT).* Athens, Greece. 2023. p. 1–5, https://www.scopus.com/record/display.uri?eid=2-s2.0-85185838366&origin=resultslist.
DOI: https://doi.org/10.1109/DESSERT61349.2023.10416502.

Conflicts of Interest: The authors declare that they have no conflict of interest regarding this study, including financial, personal, authorship or other, which could influence the research and its results presented in this article

Received 20.03.2025.
Received after revision 29.05.2025
Accepted 10.06.2025

DOI: https://doi.org/10.15276/aait.08.2025.13
УДК 004.056.5

Гібридний метод виявлення аномального трафіку в комп'ютерних мережах

Кльоц Юрій Павлович[1)]
ORCID: https://orcid.org/0000-0002-3914-0989; klots@khmnu.edu.ua. Scopus Author ID: 6504043018
Петляк Наталія Сергіївна[1)]
ORCID: https://orcid.org/0000-0001-5971-4428; npetlyak@khmnu.edu.ua. Scopus Author ID: 57786856200
Тітова Віра Юріївна[1)]
ORCID: https://orcid.org/0000-0001-8668-4834; titovav@khmnu.edu.ua. Scopus Author ID: 57786263500
[1)] Хмельницький національний університет, вул. Інститутська, 11. Хмельницький, 29016, Україна

АНОТАЦІЯ

Це дослідження присвячене проблемі виявлення аномалій у мережевому трафіку, спричинених зростаючими загрозами для інформаційно-комунікаційних систем. Традиційні системи виявлення вторгнень часто не можуть адаптуватися до нових загроз, особливо при аналізі вихідного трафіку, який може сигналізувати про внутрішню компрометацію. Для подолання цих

обмежень у дослідженні запропоновано гібридний метод виявлення, спрямований на підвищення точності ідентифікації аномалій. Метод складається з трьох компонентів. По-перше, трафік класифікується за допомогою сигнатурного підходу із заздалегідь визначеними наборами дозволених і заборонених сигнатур. По-друге, аналіз самоподібності за допомогою коефіцієнта Херста виявляє довгострокові шаблони трафіку. По-третє, нечітка логіка застосовується для інтерпретації невизначених характеристик трафіку, таких як номери портів, протоколи, інтенсивність та розміри пакетів, з використанням лінгвістичних змінних та нечітких правил. У дослідженні представлені формалізовані моделі поведінки як легітимних, так і зловмисних користувачів, а також комбінована модель сигнатур пакетів для всебічного аналізу трафіку. Такий підхід підвищує адаптивність та зменшує частку некласифікованого трафіку. Експериментальна перевірка з використанням реальних та синтетичних даних підтверджує покращену точність виявлення та менший відсоток хибних спрацьовувань порівняно з традиційними методами. Наукова новизна полягає в поєднанні детермінованої класифікації з нечіткою логікою в єдиному конвеєрі виявлення, з особливим акцентом на моніторинг вихідного трафіку. Практична цінність запропонованої системи полягає в її придатності для інтеграції в існуючі системи кібербезпеки, що сприяє більш ефективному виявленню загроз і зниженню операційних ризиків в мережевих середовищах, що розвиваються.

Ключові слова: виявлення аномалій, гібридний метод, мережевий трафік, сигнатура пакетів, нечітка логіка, самоподібність, поведінкова модель, вихідний трафік, виявлення вторгнень, кібербезпека

ABOUT THE AUTHORS

Yurii P. Klots – Candidate of Engineering Sciences, Associate Professor, head of Cybersecurity Department, Khmelnytskyi National University, 11, Instytuts'ka Str. Khmelnytskyi, 29016, Ukraine
ORCID: https://orcid.org/0000-0002-3914-0989; klots@khmnu.edu.ua. Scopus Author ID: 6504043018
***Research field*:** Computer networks; computer network security; traffic analysis for anomaly detection

Кльоц Юрій Павлович – кандидат технічних наук, завідувач кафедри Кібербезпеки
Хмельницький національний університет, вул. Інститутська, 11. Хмельницький, 29016, Україна

Nataliia S. Petliak – Senior lecturer, Cybersecurity Department. Khmelnytskyi National University, 11, Instytuts'ka Str. Khmelnytskyi, 29016, Ukraine
ORCID: https://orcid.org/0000-0001-5971-4428; npetlyak@khmnu.edu.ua. Scopus Author ID: 57786856200
***Research field*:** Security of computer networks; traffic analysis to detect anomalies; detection of vulnerabilities in information and communication systems

Петляк Наталія Сергіївна – асистент кафедри Кібербезпеки. Хмельницький національний університет, Інститутська, 11. Хмельницький, 29016, Україна

Vira Y. Titova – Candidate of Engineering Sciences, Associate Professor of the Cybersecurity Department. Khmelnytskyi National University, 11, Instytuts'ka Str. Khmelnytskyi, 29016, Ukraine
ORCID: https://orcid.org/0000-0001-8668-4834; titovav@khmnu.edu.ua. Scopus Author ID: 57786263500
***Research field*:** Decision support systems and artificial intelligence systems; risk assessment; cyber incidents

Тітова Віра Юріївна – кандидат технічних наук, доцент кафедри Кібербезпеки
Хмельницький національний університет, Інститутська. 11, Хмельницький, 29016, Україна