https://doi.org/10.31891/csit-2026-2-1

Olha ATAMANIUK
Assistant of the Computer Engineering and Information Technologies Department, Khmelnytskyi National University
https://orcid.org/0000-0001-9802-864X
e-mail: olhaatamaniuk12@gmail.com
Volodymyr DUDNYK
Master student of Information Systems and Technologies,
Khmelnytskyi National University
https://orcid.org/0009-0006-8525-1926
e-mail: dudnvova@gmail.com
Nadiia LYSENKO
Student of Information Systems and Technologies, Khmelnytskyi National University
e-mail: nadialysenko07@gmail.com

Received: 11/04/2026
Accepted:04/05/2026
Published:31/05/2026

UDC 004.9

METHOD FOR COMPUTER NETWORK TRAFFIC ANALYSIS BASED ON ENTROPY CHARACTERISTICS AND MULTIVARIATE MATHEMATICAL STATISTICS

Modern computer networks generate traffic whose behaviour changes over time not only in volume but also in internal structure. Because of this, anomaly detection cannot be reduced to fixed thresholds on separate metrics; it must account for changes in address, port, and protocol distributions together with the joint variation of interrelated traffic descriptors.

This paper presents a method for computer network traffic analysis based on entropy characteristics and multivariate mathematical statistics. The method transforms packet or flow observations collected within a time window into a state vector that combines entropy measures of categorical traffic attributes with volumetric, dispersion, and flow descriptors.

The proposed approach includes formalization of the traffic analysis process, construction of an informative feature system, a multivariate model of normal traffic states, and a structural model of the detection procedure. Algorithmic implementation is organized as a sequence of window formation, empirical distribution estimation, entropy computation, standardization, principal component transformation, multivariate statistical control, and interpretation of feature contributions.

The paper also outlines a methodology for evaluating the developed method in terms of detection quality, robustness to parameter settings, sensitivity to structural changes, and interpretability of monitoring decisions. The resulting framework is intended for traffic monitoring tasks in which payload-independent analysis and adaptation to non-stationary network behaviour are required.

Keywords: computer networks, network traffic, traffic analysis, entropy characteristics, multivariate mathematical statistics, anomaly detection, PCA, Hotelling criterion, network monitoring.

Introduction

The growth of network services, distributed infrastructures, cloud platforms, and cyber-physical systems has made traffic behaviour more variable, heterogeneous, and context-dependent. In such conditions, timely analysis of network traffic is required not only for detecting obvious overloads but also for identifying early deviations that affect reliability, security, and quality of service. Traffic monitoring therefore remains a key component of network supervision and cybersecurity support in modern infrastructures [49–52].

Classical threshold-based monitoring remains useful for simple overload situations, yet it reacts poorly to dynamic baselines and often generates false alarms under ordinary workload fluctuations. Signature-based detection is effective for known malicious patterns, but it is less useful when deviations arise from previously unseen attacks, complex behavioural changes, or operational faults whose manifestations are distributed across several traffic indicators [43, 49, 50].

The aim of this paper is to present a coherent method of computer network traffic analysis based on entropy characteristics and multivariate mathematical statistics, and to describe its algorithmic implementation in a form suitable for further experimental validation. The proposed approach combines entropy-based representation of traffic structure with multivariate statistical decision-making, which creates a stronger basis for analysing abnormal states in non-stationary traffic environments [50, 52, 55].

Related works

Existing approaches to traffic anomaly detection can be grouped into signature-based, threshold-based, forecasting, entropy-based, and multivariate statistical methods. Signature detectors are precise when the traffic matches known patterns, but they cannot generalize to previously unseen threats [43]. Threshold and one-dimensional statistical schemes are easy to implement, although they react poorly to ordinary workload fluctuations and often miss low-intensity structural deviations [49, 50]. Forecasting methods are useful when the traffic exhibits a stable temporal pattern, yet their reliability depends strongly on model assumptions and on the availability of representative historical data [50, 53].

Entropy-based approaches are attractive because they operate on the structure of traffic rather than on aggregate volume alone. Changes in source and destination addresses, ports, or protocol distributions can reveal anomalies that remain weakly visible in pure volume metrics. At the same time, entropy taken in isolation is not always sufficient, because some abnormal states appear as coordinated shifts across several descriptors rather than as a large deviation of one distribution [49, 55].

Multivariate statistical methods, including covariance-based control, principal component analysis, and Hotelling-type criteria, address this limitation by considering traffic as a vector of interdependent descriptors. Their main advantage lies in the ability to detect coordinated deviations and to reduce the dependence of the final decision on a single metric. In practice, this logic is also consistent with the evolution of modern monitoring platforms and traffic analysis tools that combine flow inspection, behavioural analysis, and anomaly-oriented correlation mechanisms [29, 31, 33, 40, 42–44, 50–52].

For a structured comparison of these approaches by advantages, disadvantages, and relevance to the present study, their characteristics are summarized in Table 1.

Table 1.

Comparison of network traffic anomaly detection methods

Method	Advantages	Disadvantages	Relevance to this study
Signature-based methods	High accuracy for known attacks	Do not detect unknown threats	Useful only as a baseline
Threshold-based methods	Simple and fast	Poor adaptability, many false alarms	Limited applicability
Entropy-based methods	Detect structural changes in traffic well	Weak as a standalone method for complex anomalies	One of the core components
Multivariate statistical methods	Account for correlations between features, reduce false positives	More complex to configure and compute	One of the core components
Behavioral / ML methods	Adaptive, can detect unknown anomalies	Harder to interpret, require more data	Useful for comparison
Proposed integrated method	Combines structural sensitivity and multidimensional analysis	Requires calibration and experimental validation	Main method of the study

As follows from Table 1, the most promising direction is the integration of entropy-based traffic representation with multivariate statistical analysis, since such a combination makes it possible to preserve structural sensitivity while improving the reliability of anomaly detection in multidimensional feature space.

Method

The proposed method treats network traffic as a sequence of packet or flow observations collected at a given observation point and aggregated over successive time windows. For a window W_t, the raw records are transformed into a feature vector $x_t = \varphi(W_t)$, where φ(·) maps the observed communications to a compact numerical representation of network state. The basic notation used in the formal description is summarized in Table2.

The method is intentionally payload-independent. It uses only the information that can be derived from packet headers or exported flow records, which makes it suitable for realistic monitoring environments. This choice is consistent with practical monitoring tools that rely on flow metadata and protocol-level observations rather than on deep payload inspection [42–44]. It also assumes that traffic is non-stationary; therefore, the model of normal behaviour must allow for ordinary variation without losing sensitivity to genuine structural deviations.

The notation introduced in Table 1 is used consistently in the subsequent stages of the method. It links the observation window, entropy descriptors, the traffic state vector, and the multivariate decision rule into a single formal description.

Entropy-based descriptors are selected because they reflect the degree of concentration or dispersion in traffic distributions. For monitoring purposes, the most informative attributes are usually source and destination addresses, source and destination ports, and protocol identifiers. Together with statistical characteristics of intensity and packet behaviour, they form the informative state vector summarized in Table 3 [49, 55].

In practical terms, entropy does not replace volume indicators; it complements them. For example, DDoS behaviour can combine concentration on a small set of destinations with a sharp rise in packet rate, whereas scanning can produce strong dispersion in destination ports even without a large increase in total traffic. For this reason, the method combines entropy measures and auxiliary statistical descriptors within one feature space instead of relying on a single indicator [49, 50, 55].

Table 2.

Notation used in the method

Notation	Meaning
W_t	Set of traffic records observed in time window t (packets or flows).
x_t	Vector of informative features describing the traffic state in window t.
A	Categorical traffic attribute, for example srcIP, dstIP, srcPort, dstPort, or protocol.
$\{c_i\}$	Counts of occurrences of attribute values in the current window.
$\{p_i\}$	Estimated empirical probabilities $p_i = \frac{c_i}{\Sigma_j c_j}$.
H(A)	Entropy characteristic of the empirical distribution of attribute A.
g(·)	Decision rule used for state classification or anomaly detection.
T^2	Hotelling statistic used for multivariate monitoring and deviation control.

Table 3.

Recommended informative features for the traffic state vector

Feature group	Examples of concrete features	Data type
Entropy-based	H(srcIP), H(dstIP), H(srcPort), H(dstPort), H(proto)	Numeric
Generalized entropies (optional)	H_α (*Rényi*), $S_q(Tsallis)$	Numeric
Diversity descriptors	$N_{distinct}(A)$ for selected attributes	Integer
Volumetric descriptors	Packets per window, bytes per window, mean packet size, packet rate	Numeric
Flow descriptors	Number of active flows, mean or quantile flow durations	Numeric
Divergence / change descriptors (optional)	KL or JSD distance for selected attributes	Numeric

The state vector formed from these descriptors allows the traffic window to be interpreted as a point in a multidimensional feature space. This representation makes it possible to analyse structural and volumetric deviations jointly rather than evaluating each metric in isolation.

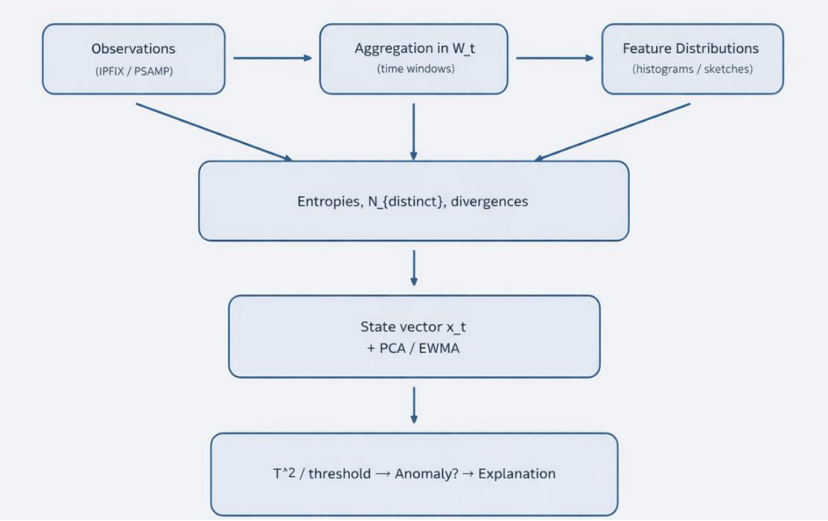

Fig. 1. Structural model of the method

After the feature system is defined, the normal state of the network is modelled statistically. Let μ denote the mean vector of the baseline traffic and Σ the corresponding covariance structure. Deviation is then interpreted not as a separate threshold violation of one coordinate but as a multivariate displacement of the current observation from the learned region of normal behaviour. Such an interpretation corresponds to the general logic of multivariate abnormal-state detection used in recent network-oriented studies [50–53].

The structural model of the method can be organized into six successive functional blocks: acquisition of traffic observations, time aggregation, construction of empirical distributions, computation of entropy and auxiliary

descriptors, formation of the multidimensional state vector with multivariate normalization, and statistical decision-making. This organization is shown in Figure 1 and reflects the implementation logic of the proposed method.

At the level of time aggregation, the method supports both non-overlapping and sliding windows. The first option is simpler and easier to interpret; the second reduces boundary effects and usually provides smoother temporal trajectories of entropy and multivariate scores. The main choices of window width and shift, together with their practical influence on detection behaviour, are summarized in Table 4.

Table 4.

Recommended windowing parameters and their influence

Parameter	Variant	What improves	Typical risk or trade-off
Window width	Smaller	Higher temporal resolution and better sensitivity to short incidents.	Unstable distribution and entropy estimates because of smaller samples.
Window width	Larger	More stable estimates and lower variance of the monitored statistics.	Short anomalies are diluted and detection may be delayed.
Step δ	δ = Δ (non-overlapping)	Simpler implementation and minimum repetition of records.	Pronounced boundary effects and fragmented temporal curves.
Step δ	δ < Δ (sliding)	Smoother time series and fewer discontinuities between adjacent windows.	Higher computation cost and stronger correlation between neighbouring windows.

For each window, empirical distributions are built for the selected categorical attributes. The basic entropy estimate is then obtained from the observed frequencies, optionally normalized to make values comparable across windows with different numbers of unique states. If the method is extended to generalized entropy measures, the same stage becomes the place where sensitivity to dominant or rare events can be adjusted [55].

Before multivariate decision-making, the feature space is standardized so that descriptors with different scales do not dominate the control statistic. Principal component analysis may be used to reduce dimensional redundancy and to separate the coordinated variation of normal traffic from residual deviations. The resulting sequence of operations that leads to an anomaly decision is presented in Figure 2 and corresponds to the class of multivariate statistical control procedures described in related studies [50–53].

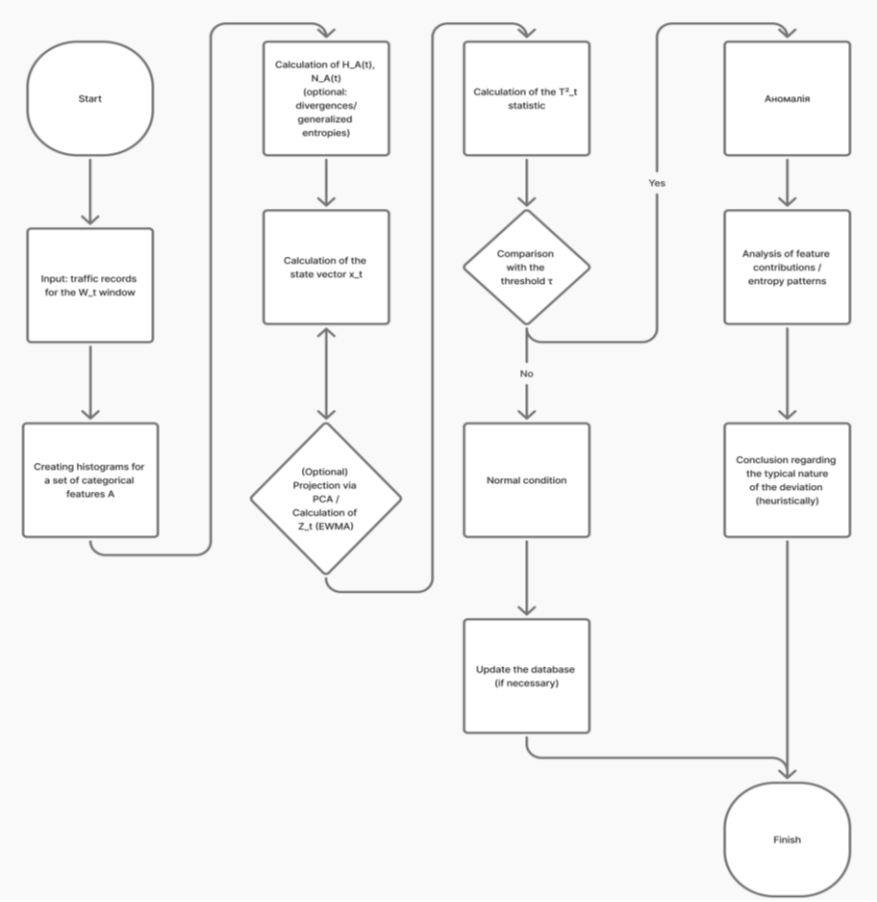

Fig. 2. Flowchart of anomaly detection

Experiments

The experimental stage verifies the proposed detector at the level of the complete decision chain described in Section 3. The goal is not only to confirm the fact of anomaly detection, but also to show how the selected feature composition, the windowing scheme, the PCA model, and the final T^2–Q decision rule behave on traffic that contains both normal operation and controlled disturbances. For this reason, the evaluation protocol is linked directly to the methodological components already introduced in the paper: the state vector follows Table 2, the temporal aggregation follows Table 3, the processing structure corresponds to Figure 1, and the final decision logic corresponds to Figure 2 [48–50, 53].

The traffic trace is divided into a training subset and a testing subset. The training subset contains only normal windows and is used to estimate the mean vector, standard deviations, covariance structure, and principal components of the baseline profile. The testing subset contains both normal traffic and controlled anomaly blocks. Such a separation is necessary because the inclusion of anomalous windows in the baseline sample would partially absorb abnormal behaviour into the normal subspace and would therefore weaken the contrast between normal and abnormal states. The main parameters of the computational experiment are summarized in Table 4 [21, 22, 48, 50].

In the article version of the experiment, traffic is aggregated into non-overlapping windows of $\Delta t = 10$ s, and the same value is used as the shift. The training sample contains 900 normal windows, while the testing sample contains 780 windows, including 500 normal and 280 anomalous windows. Each window is represented by the ten-dimensional feature vector described in Table 2: five normalized entropy descriptors and five statistical descriptors that characterize packet size, flow activity, and traffic intensity. After z-standardization, principal component analysis retains eight principal components, which explain 95.71% of the variance of the normal traffic profile. This provides a compact but still informative representation of coordinated traffic behaviour and creates the basis for subsequent control by Hotelling's T^2 statistic and the residual Q-statistic.

To test the method against qualitatively different disturbances, four anomaly scenarios are injected into the test trace: DDoS, port scan, flash crowd, and link failure. The DDoS block produces concentration of traffic on one destination together with an increase in intensity. The port-scan block keeps the aggregate rate close to the background level, but sharply changes the diversity of destination ports and addresses. The flash-crowd block imitates a legitimate demand surge with an increase in active flows and a shift of the service profile. The link-failure block causes a coordinated decrease in traffic intensity and a simplification of the active communication structure. The temporal response of one of the key structural descriptors, the normalized entropy of destination IP addresses, is shown in Figure 3 [23–25, 49, 55].

For comparison, the proposed detector is evaluated together with two baseline schemes. The first baseline is a fixed-threshold detector that marks a window as anomalous when the packet rate deviates from its normal mean by more than three standard deviations. The second baseline is an entropy-only detector that triggers when at least two entropy coordinates leave their normal ranges. Against these baselines, the proposed method uses the full chain of state-vector formation, PCA projection, computation of T^2 and Q, and final classification by multivariate control limits. The aggregate quality indicators of the compared methods are presented in Table 5.

Table 6 shows that the combined entropy-statistical detector provides the best balance between sensitivity and stability. The fixed-threshold baseline reacts mainly to large amplitude changes and therefore misses anomalies whose main manifestation is structural rather than volumetric. The entropy-only detector is much stronger on such anomalies, but it still evaluates the traffic state through separate descriptors and does not fully use the correlation structure of the feature space. The proposed method reaches the highest recall and F1-score while maintaining a very low false positive rate, which confirms the benefit of combining entropy descriptors with multivariate statistical control.

The dynamics of Hotelling's T^2 statistic for the test trace are shown in Figure 4. In the normal part of the trace, the statistic fluctuates below the control threshold, whereas inside the anomalous blocks it rises sharply and remains above the limit until the disturbance disappears. This behaviour is important from the practical point of view because it demonstrates not only the fact of detection, but also the temporal stability of the alarm. When Figures 3 and 4 are interpreted jointly, it becomes clear that the detector reacts both to structural redistribution of traffic and to coordinated changes in intensity, which is precisely the effect expected from the method developed in Sections 2 and 3 [22–25, 48–50].

Table 5.

Parameters of the computational experiment

Parameter	Value
Window length Δt	10 s
Window shift δ	10 s (non-overlapping windows)
Training sample	900 normal windows
Test sample	780 windows: 500 normal + 280 anomalous
Feature vector dimension	10
Entropy descriptors	hsrcIP, hdstIP, hsrcPort, hdstPort, hproto
Retained principal components	8
Explained variance of the PCA model	95.71%
Decision rule	$T^2 > T^2_{0.999}$ or $Q > Q_{0.999}$

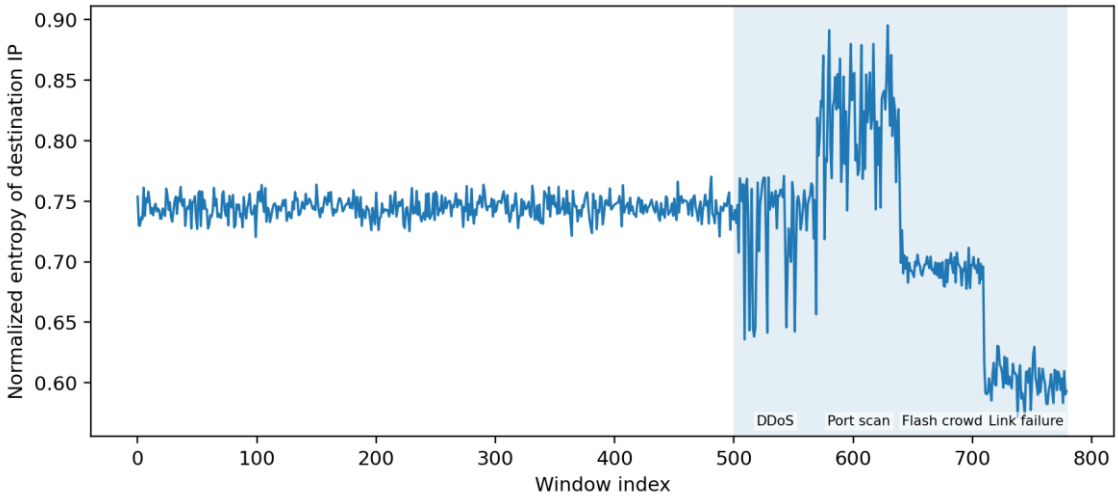

Fig. 3. Dynamics of the normalized entropy of destination IP addresses

Table 6.

Detection quality of compared methods

Method	Accuracy, %	Precision, %	Recall, %	F1-score, %	False positive rate, %
Fixed threshold	76.54	100.00	34.64	51.46	0.00
Entropy only	97.31	100.00	92.50	96.10	0.00
Proposed combined	99.74	99.29	100.00	99.64	0.40

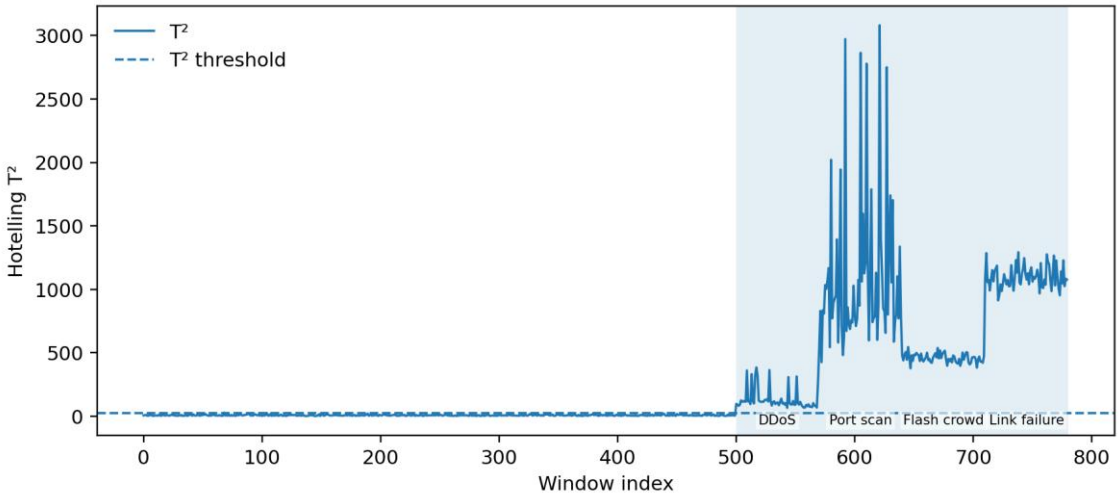

Fig. 4. Dynamics of Hotelling's T^2 statistic on the test trace

Conclusions

The paper consolidates the methodological content of the developed traffic analysis approach into an article format while preserving the core logic of the original work. The proposed method models network traffic through time-windowed state vectors that combine entropy characteristics of categorical attributes with volumetric, diversity, and flow descriptors. This makes it possible to describe anomalies not only through separate metric violations but through coordinated changes in traffic structure and behaviour.

The algorithmic implementation of the method is organized as a consistent chain of window formation, empirical distribution construction, entropy computation, feature standardization, multivariate statistical control, and contribution-based interpretation. The article preserves the key implementation components of the developed approach, including the notation of the method, the feature composition, the structural model, the main parameterization choices, and the anomaly detection flow.

The developed framework is suitable for further experimental validation on labelled traffic traces and for subsequent adaptation to practical monitoring systems. Its main value lies in combining payload-independent structural descriptors with statistically interpretable multivariate decision rules, which creates a stronger basis for analysing abnormal network states in modern, non-stationary traffic environments.

ADDITIONAL INFORMATION

AUTHOR CONTRIBUTIONS

Conceptualization, O.A.; methodology, V.D.; validation, N.L.; formal analysis, V.D.; investigation, N.L.; data curation, V.D..; writing-original draft preparation, N.L.; writing-review and editing, O.A.; visualization, N.L.; supervision, O.A.; project administration, O.A. All authors have read and agreed to the published version of the manuscript.

DECLARATION ON THE USE OF GENERATIVE ARTIFICIAL INTELLIGENCE TOOLS

In preparing this work, the author used DeepL Translate and Grammarly for: grammar and spelling checks, paraphrasing, and rephrasing. After using these tools/services, the author reviewed and edited the content and takes full responsibility for the content of this publication.

REFERENCES

1. Worch E., Hydrochemistry: basic concepts and exercises, De Gruyter, Berlin/ Kashtalian Antonina, Serhii Lysenko, et al. "Control and Decision-Making in Deceptive Multi-Computer Systems Based on Previous Experience for Cybersecurity of Critical Infrastructure." *Applied Sciences* Vol 15 P. 12286.

2. Denysiuk Dmytro, Oleg Savenko, Sergii Lysenko, Bohdan Savenko, Andrii Nicheporuk. "Detecting software implants using system decoys." (2024).

3. Bokhonko Oleksandr, Sergii Lysenko, Piotr Gaj. "Development of the social engineering attack models." (2024).

4. Yang, Xiaodong. "Efficient and security-enhanced certificateless aggregate signature-based authentication scheme with conditional privacy preservation for VANETs." *IEEE Transactions on Intelligent Transportation Systems* 25.9. 2024 P. 12256-12268.

5. LaRocque Anthony. Effective ransomware detection using autonomous patternbased signature extraction. 2024. P. 1-12

6. Rehman, Fazalur, Farhan Mushtaq, Hafsah Zaman. "A Host-based Intrusion Detection: Using Signature-based and AI-driven Anomaly Detection for Enhanced Cybersecurity." *2024 4th International Conference on Digital Futures and Transformative Technologies (ICoDT2)*. IEEE, 2024. P. 1-7.

7. Loaiza, Carlos. Dynamic temporal signature analysis for ransomware detection using sequential entropy monitoring. *Authorea Preprints.* 2024. P. 30-45.

8. Kalla, Dinesh. Investigating the Impact of Heuristic Algorithms on Cyberthreat Detection. In: *2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT)*. IEEE, 2024. p. 450-455.

9. Taylor, Theodore. Dynamic anomaly-driven detection for ransomware identification: An innovative approach based on heuristic analysis. *Authorea Preprints*, 2024. P. 1-5.

10. Yunmar, Rajif Agung. Hybrid Android malware detection: A Review of heuristic-based approach. *IEEe Access*, 2024, 12. P. 41255-41286.

11. Novak, Pavel, OUJEZSKY, Vaclav. Heuristic malware detection method based on structured cti data: A research study and proposal. In: *2024 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. IEEE, 2024. p. 1-6.

12. Hu, Wengui. A deep analysis of nature-inspired and meta-heuristic algorithms for designing intrusion detection systems in cloud/edge and IoT: state-of-the-art techniques, challenges, and future directions. *Cluster Computing*, 2024, 27.7. P. 8789-8815.

13. More, Sanjana A., KACHAVIMATH, Amit V. SDN Intrusion Detection using Meta-Heuristic Optimization and K-Nearest Neighbors Classifier. *Procedia Computer Science*, 2025, 260. P. 1137-1144.

14. Fuller, Richard. A novel hybrid machine learning approach for real-time ransomware detection using behavior-driven heuristic features. 2024. P. 120-125

15. Okoli, Ugochukwu Ikechukwu. Machine learning in cybersecurity: A review of threat detection and defense mechanisms. *World Journal of Advanced Research and Reviews*, 2024, 21.1. P. 2286-2295.

16. Kumar, Busireddy Hemanth. Big Data in Cybersecurity: Enhancing Threat Detection with AI and ML. *Metallurgical and Materials Engineering*, 2025, 31.3. P. 12-20.

17. Mohammed, Anwar. Transforming SOC Operations: Harnessing the Power of AI and ML for Enhanced Threat Detection. *INTERNATIONAL JOURNAL OF RESEARCH CULTURE SOCIETY Monthly Peer-Reviewed, Refereed, Indexed*, 2024. P. 8.

18. Katiyar, Nirvikar. AI and Cyber-Security: Enhancing threat detection and response with machine learning. *Educational Administration: Theory and Practice*, 2024, 30.4. P. 6273-6282.

19. Marimuthu, Oviya, RAVI, Priyadharshini, JANARTHANAN, Senthil. Application of AI and ML in Threat Detection. *Protecting and Mitigating Against Cyber Threats: Deploying Artificial Intelligence and Machine Learning*, 2025. P. 29.

20. Alzaabi, Fatima Rashed, MEHMOOD, Abid. A review of recent advances, challenges, and

opportunities in malicious insider threat detection using machine learning methods. *IEEE Access*, 2024, 12. P. 30907-30927.

21. Mizanur, Mohammad. Machine Learning-Based Anomaly Detection for Cyber Threat Prevention. *Journal of Primeasia*. 2025, 6.1. P. 1-8.

22. Subrahmanyam, Satya. Behavioral Analysis for Threat Detection. In: *Handbook of AI-Driven Threat Detection and Prevention*. CRC Press, 2025. P. 95-115.

23. Ozturk, Mehmet. Dynamic behavioural analysis of privacy-breaching and data theft ransomware. 2024.

24. Blowing, Adam. Performing ransomware detection through predictive behavioral mapping to autonomous threat identification. 2024. P. 7.

25. Shanks, Gene. Innovative framework for ransomware detection using adaptive cryptographic behavior analysis. 2024. P. 10-14.

26. Microsoft Defender for Endpoint. URL: https://www.microsoft.com/en-us/security/business/endpoint-security/microsoft-defender-endpoint?msockid=382eb02bfecb63240938a310ffd9621e .

27. CrowdStrike Falcon. URL: https://www.crowdstrike.com/en-us/platform/ .

28. Cortex XDR. URL: https://docs-cortex.paloaltonetworks.com/r/Cortex-XDR/Cortex-XDR-5.x-Documentation/What-is-Cortex-XDR .

29. Darktrace Platform: https://www.darktrace.com/ .

30. Elastic Security Platform. URL: https://www.elastic.co/security .

31. IBM Qradar. URL: https://www.ibm.com/products/qradar .

32. Google Chronicle. URL: https://cloud.google.com/blog/products/identity-security/introducing-the-unified-chronicle-security-operations-platform/ .

33. Splunk Enterprise Security. URL: https://www.splunk.com/en_us/products/enterprise-security.html .

34. SentinelOne Singularity. URL: https://www.sentinelone.com/platform/singularity-complete/ .

35. Cisco SecureX. URL: Introducing SecureX - Cisco Blogs .

36. Fortinet FortiEDR. URL: https://www.fortinet.com/products/endpoint-security/fortiedr .

37. Sophos Intercept X. URL: https://www.sophos.com/en-us/products/mobile-control/intercept-x .

38. Trend Micro Apex One. URL: Apex One and Apex Central are now available .

39. Trellix MVISION. URL : https://techdirectarchive.com/2025/08/12/how-to-install-trellix-mvison-endpoint/.

40. Rapid7 InsightIDR. URL: https://www.rapid7.com/products/siem/ .

41. Sumo Logic Cloud SIEM. URL: https://www.sumologic.com/solutions/cloud-siem .

42. Zeek Overview. URL: https://zeek.org/ .

43. Suricata Homepage. URL: https://suricata.io/ .

44. Wazuh Platform. URL: https://wazuh.com/ .

45. MISP Platform. URL: https://www.misp-project.org/ .

46. OpenCTI. URL: https://docs.opencti.io/latest/ .

47. YARA Rules. URL: https://github.com/Yara-Rules/rules .

48. Bokhonko Oleksandr, Atamaniuk Olha. Method for synthesis of a scalable architecture of a distributed cs, resistant to social engineering attacks. Computer Systems and Information Technologies. Vol. (4). P. 60–76. 2025

49. Chornobuk Maksym, Dubrovin Valeriy, Deineha Larysa. Cybersecurity: Research on Methods for Detecting DDoS Attacks. Computer Systems and Information Technologies. Vol. (4). P. 6–9. 2023.

50. Savenko Bohdan, Kashtalian Antonina. A Method for Determining the Effectiveness of a Distributed System for Detecting Abnormal Manifestations. Computer Systems and Information Technologies. Vol. (2). P. 14–22. 2022.

51. Ramskyi Ihor, Drozd Andriy, Lyhun Oleksii, Ponochovna Olena. System for Cybersecurity Evaluation of Corporate Networks. Computer Systems and Information Technologies. Vol. (2). P. 123–131. 2025.

52. Andrieiev Dmytro, Lyhun Oleksii, Drozd Andriy, Ponochovna Olena. Monitoring System for Critical Infrastructure Objects Based on Digital Twins. Computer Systems and Information Technologies. Vol. (2). P. 27–35. 2025.

53. Ivanets Ivan, Ovsyak Volodymyr, Ovsyak Oleksandr. Edge-Native Cable Access Network with UDP Termination. Computer Systems and Information Technologies. Vol. (2). P. 45–54. 2025.

54. Shelest Mykhailo, Pidlisnyi Yurii, Kapustian Mariia. Methods of Hiding Data in Computer Networks: From Classics to IoT and AI. Computer Systems and Information Technologies. Vol. (4). P. 27–34. 2025.

55. Kashtan Vita, Hnatushenko Volodymyr. Entropy-Cryptographic Approach for Transmission of Satellite Data in Telecommunication Networks. Computer Systems and Information Technologies. Vol. (4). P. 53–59. 2025.

Ольга АТАМАНЮК, Володимир ДУДНИК, Надія ЛИСЕНКО
Хмельницький національний університет

МЕТОД АНАЛІЗУ ТРАФІКУ КОМП'ЮТЕРНИХ МЕРЕЖ НА ОСНОВІ ЕНТРОПІЙНИХ ХАРАКТЕРИСТИК ТА БАГАТОВИМІРНОЇ МАТЕМАТИЧНОЇ СТАТИСТИКИ

Сучасні комп'ютерні мережі формують трафік, поведінка якого змінюється не лише за обсягом, а й за внутрішньою структурою. Тому виявлення аномалій не може зводитися до фіксованих порогів для окремих метрик; воно має враховувати зміни у розподілах адрес, портів і протоколів разом зі спільною варіацією взаємопов'язаних характеристик трафіку.

У статті запропоновано метод аналізу трафіку комп'ютерних мереж, що ґрунтується на використанні ентропійних характеристик і засобів багатовимірної математичної статистики. Актуальність роботи зумовлена тим, що сучасний мережевий трафік є нестаціонарним: його поведінка з часом змінюється не лише за обсягом, а й за внутрішньою структурою. У зв'язку з цим виявлення аномалій не може базуватися виключно на фіксованих порогах окремих показників, оскільки потребує врахування змін у розподілах адрес, портів, протоколів, а також спільної варіації взаємопов'язаних дескрипторів трафіку.

Розроблений метод передбачає перетворення спостережень за пакетами або потоками, зібраними в межах заданого часового вікна, у вектор стану мережевого трафіку. Такий вектор поєднує ентропійні міри категоріальних атрибутів із об'ємними, дисперсійними та потоковими характеристиками. Запропонований підхід охоплює формалізацію процесу аналізу трафіку, побудову інформативної системи ознак, створення багатовимірної моделі нормальних станів трафіку та структурної моделі процедури виявлення відхилень.

Алгоритмічна реалізація методу організована як послідовність етапів: формування часових вікон, оцінювання емпіричних розподілів, обчислення ентропії, стандартизація ознак, перетворення методом головних компонент, багатовимірний статистичний контроль і подальша інтерпретація внеску окремих ознак у виявлені зміни. Така організація забезпечує можливість не лише фіксувати аномальні стани, а й пояснювати причини їх появи.

Окремо визначено методику оцінювання запропонованого методу за показниками якості виявлення, стійкості до вибору параметрів, чутливості до структурних змін і рівня інтерпретованості результатів моніторингу. Запропонований підхід орієнтований на задачі моніторингу мережевого трафіку, у яких необхідний аналіз без урахування вмісту корисного навантаження та адаптація до змінної поведінки комп'ютерних мереж.

Ключові слова: комп'ютерні мережі, мережевий трафік, аналіз трафіку, ентропійні характеристики, багатовимірна математична статистика, виявлення аномалій, PCA, критерій Хотеллінга, моніторинг мереж.