BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en

Vol. 31 No. 1. 2026

UDC 004.5
DOI: 10.62660/bcstu/1.2026.34

ISSN: 2306-4412
E-ISSN: 2708-6070

Method for detecting computer attacks based on adaptability and balancing of training samples

Bogdan Semenyuk*

Postgraduate Student
Khmelnytskyi National University
29016, 11 Instytutska Str., Khmelnytskyi, Ukraine
https://orcid.org/0009-0001-8831-8835

Bogdan Savenko

Doctor of Philosophy, Associate Professor
Khmelnytskyi National University
29016, 11 Institutska Str., Khmelnytskyi, Ukraine
https://orcid.org/0000-0001-5647-9979

Abstract. The relevance of this work is determined by the fact that modern cyber defence systems still face the problem of data imbalance in detecting computer attacks. Uneven class representation reduces the ability of machine learning models to recognise rare but critically important types of intrusions. Existing methods of synthetic sample augmentation often distort the data structure and lead to the loss of characteristic attack signatures, which negatively affects the reliability of classification. The aim of the study was to improve the accuracy and completeness of network attack detection by developing an adaptive method for balancing the training sample while preserving the statistical and signature properties of anomalies. The research methodology involved the creation of a deterministic algorithm that calculated the supplementation coefficients separately for each type of attack, taking into account its frequency and the minimum required sample size. A statistical approach based on local medians and extreme values of neighbouring samples was used to generate new examples, which ensured the reproduction of the typical structure of anomalous patterns without random interpolation. The developed method was integrated into a data processing sequence, within which network traffic parameters were converted into sound signals, and spectrograms were formed on their basis for further analysis by a two-dimensional convolutional neural network. The main results of the study showed an increase in the completeness of detection of rare types of attacks by an average of 10-12% compared to basic approaches, while maintaining a stable level of overall accuracy. Preserving signatures in synthetic samples ensured improved recognition of rare attacks and increased classification reliability. The practical value of the work lies in the possibility of applying the developed method to form balanced samples in intrusion detection systems and its integration with existing deep learning models in order to improve the reliability of corporate network cyber protection

Keywords: data imbalance; signature modelling of anomalies; synthetic sample augmentation; network traffic analysis; convolutional neural networks; acoustic data representation; classification robustness

INTRODUCTION

The rapid growth of the Internet of Things (IoT) ecosystem has led to an unprecedented increase in the number of connected devices and the amount of data generated in real time. The massive introduction of sensor networks, household smart devices, and industrial IoT solutions has created a complex and dynamic environment in which the speed of information exchange is constantly increasing. However, this scale

Article's History: Received: 05.11.2025; Revised: 13.02.2026; Accepted: 16.03.2026; Published: 08.04.2026

Suggested Citation:
Semenyuk, B., & Savenko, B. (2026). Method for detecting computer attacks based on adaptability and balancing of training samples. *Bulletin of Cherkasy State Technological University*, 31(1), 34-43. doi: 10.62660/bcstu/1.2026.34.

*Corresponding author

of connection is accompanied by an increase in the number of vulnerabilities.

T. Alladi *et al.* (2020) showed that IoT devices are often targeted by attacks due to limited resources, weak authentication protocols, and insufficient isolation of network segments. In these circumstances, intrusion detection systems that can track abnormal behaviour and respond quickly to threats become a critical component of protection. As noted by A. Hussain *et al.* (2023), the advanced Intrusion Detection System (IDS) for IoT environments serves as an adaptive monitoring mechanism that analyses network traffic, detects deviations from typical patterns, and complements other cyber defence tools, ensuring infrastructure resilience to new and complex attacks. The development of intrusion detection systems is largely associated with the introduction of machine and deep learning methods that provide the ability to automatically analyse complex network traffic patterns and adapt to new types of attacks. Comparative study by I.F. Kilincer *et al.* (2021) showed that machine learning algorithms demonstrate high efficiency in classifying network events due to their ability to consider multidimensional traffic features and detect anomalies that are not detected by conventional methods. Further improvement of accuracy is possible due to the use of deep models: G. Li & J.J. Jung (2023) emphasised that convolutional and other deep learning (DL) architectures are capable of modelling complex nonlinear dependencies that occur in multicomponent time series, providing higher noise resistance and efficient detection of previously unknown types of anomalies. The success of using Machine Learning (ML) or DL in IDS systems largely depends on the structure and balance of training data sets, which form the basis for correct training of models and prevent degradation of results when classifying rare attacks.

Ukrainian researchers also made a significant contribution to the development of methods for analysing anomalies. In particular, D. Denysiuk *et al.* (2023) confirmed the effectiveness of using local statistical features to detect hidden modifications in steganographic structures. The study by A. Kashtalian *et al.* (2025) was dedicated to the architecture of centralised multi-computer systems with built-in traps and baits, which increases the efficiency of detecting abnormal activity in distributed environments. A separate group of studies focused on the impact of sample imbalance on the quality of models. In particular, J. Dalou' *et al.* (2020) described an entropy-oriented approach for detecting Distributed Denial of Service (DDoS) attacks in software-defined networks (SDN), which allows adaptive response to changes in traffic intensity. The study by J.E. Joseph *et al.* (2025) proved that the accuracy of classification directly depends on the balance of the training sample, and depth models are not able to compensate for the lack of representation of rare attacks.

Analysis of publications in 2020-2025 showed that the problem of qualitative balancing of training samples in the context of anomaly detection remains open. Most of the available approaches do not consider the internal structure of network traffic, which leads to a loss of signature characteristics of attacks. In addition, many studies lack mechanisms for adaptive selection of resampling coefficients for individual classes, as a result of which rare types of attacks remain underrepresented. Despite significant progress in the development of data balancing methods for computer attack detection tasks, existing approaches remain vulnerable to a number of technical limitations. Classical synthetic complement algorithms, although they partially compensated for the lack of rare attacks, often led to distortion of local structural properties of data, loss of signatures, and a decrease in the ability of models to correctly reproduce class boundaries. Additional difficulties arose due to the dependence of such methods on global parameters, which did not consider the heterogeneity of attacks and the variability of their internal subtypes. Under these conditions, there is a need for an approach that can overcome these shortcomings, preserve the characteristic features of rare samples, and ensure correct integration into advanced traffic processing pipelines, which include data conversion into spectrograms and subsequent classification by convolutional neural networks. It was certain limitations that became the basis for the development of the proposed balancing method aimed at overcoming these structural and methodological problems.

The purpose of the study was to develop and experimentally test the adaptive signature of a method for synthetic balancing of the training sample, aimed at improving the completeness of detection of rare attacks without reducing the overall classification accuracy. It was assumed that the use of this method will ensure the reproducibility of the new generation intrusion detection system, improve its resistance to class imbalance, and allow forming synthetic samples that correctly transmit local features of network traffic without losing signature properties.

MATERIALS AND METHODS

The purpose of the study was to develop and experimentally verify the adaptive signature of a method for synthetic balancing of training samples for computer attack detection tasks. The methodology included a combination of classical approaches to working with tabular features of network traffic, advanced techniques for preprocessing and converting data into two-dimensional structures, and the use of a compact convolutional neural network. All experiments were performed under the same conditions, with fixed initial values of random number generators, which ensured complete reproducibility of the simulation results. The paper used tabular parameters of network traffic, which contained numerical and categorical features. The imbalance of such samples is a common problem, and the lack of representation of rare attacks in training data leads to a systematic deterioration in the ability of models to

correctly recognise complex types of threats. Therefore, pretreatment provided for the need to correct the imbalance. Categorical attributes (protocol_type, service, flag) were encoded in whole indexes and then processed using the Synthetic Minority Oversampling Technique for Nominal and Continuous features (SMOTENC), which allowed their nature to be correctly taken into consideration. Numeric features were scaled to a range of -1 to 1, which stabilised calculations in the feature space and improved the behaviour of balancing algorithms.

A special place in the methodology was occupied by the stage of converting tabular data into a two-dimensional form using sonification. Each network traffic record was converted to a Pulse-Code Modulation (PCM) signal with a sampling rate of 2,000 Hz and duration of 0.4 seconds. The signals were pretreated by converting them into spectrograms generated using a short-time Fourier Transform (STFT) with parameters nperseg = 256 and noverlap = 128. The resulting spectrograms were converted to a logarithmic scale, after which they were interpreted as single-channel images suitable for further processing by a convolutional neural network. The imbalance was eliminated exclusively for the training part of the data, until the final scaling and coding. The modified version of SMOTENC was applied selectively – only to those subtypes of attacks that were characterised by a low detection completeness index or had an insufficient number of examples. Complement volumes were controlled using the SMOTE_MAX_MULTIPLIER and SMOTE_MIN_COUNT parameters, the limits for which were determined empirically during stratified five-part cross-validation. Balancing efficiency was evaluated using a macro-average completeness indicator (macro-recall), and monitoring the level of false-positive classifications.

A compact two-dimensional convolutional neural network was used for classification. At the first stage, the model used a Convolutional 2D convolutional layer with 16 3x3 filters with Rectified Linear unit activation, which helped to distinguish local patterns of spectral representations. The features were then aggregated using the maximum pooling sub-sample pooling layer, which reduced the dimension of the spectrogram and increased resistance to local fluctuations. Similar architectures are widely used in modern network anomaly detection systems.

The training was performed using the Adam optimiser at a rate of 5×10^{-4}, parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$, and the binary cross-entropy loss function. The package size was 128 objects. To prevent retraining, an early stopping mechanism was used with the patience = 3 parameter, monitoring the val_loss indicator, and checking the stability of the model behaviour. The effectiveness of the attack detection model was evaluated based on metrics that are sensitive to class imbalances, in particular, precision, recall, and F_1-measure, which reflects their harmonious average. The use of such metrics was critical for tasks where rare classes are crucial for the quality of IDS work, and an uneven number of examples can lead to a significant shift in results.

All experimental studies were conducted using the Network Security Laboratory – Knowledge Discovery Dataset (NSL-KDD), which contains 125,973 training examples and 22,544 testing examples and is an improved and duplicated version of the original KDD'99. NSL-KDD covered 41 network attributes, including both numeric and categorical parameters, and attacks were grouped into four key categories: DoS, Scanning/Probing attacks, User-to-Root, and Remote-to-Local. As part of this study, three comparable systems were formed: a basic model without balancing, a variant with classical SMOTENC, and a system with a signature – saving adaptive approach. All experiments were performed with the same learning parameters and the same model architecture, which ensured the validity and correctness of comparing the results.

RESULTS AND DISCUSSION

Below is a mathematical formalisation of a method for detecting computer attacks based on adaptability and balancing of the training sample. Formalisation describes the mechanisms for determining the target number of synthetic examples, local statistics, bias directions, and principles for preserving the signature properties of data that distinguish the proposed approach from classical interpolation algorithms. Parameter $c > 0$ determined the sensitivity of the deterministic deviation coefficient $\beta(x)$, given by equation (7), to the local distance $d(x)$, calculated according to (6). Lower values of c led to a more drastic change of $\beta(x)$ relative to $d(x)$, while larger – contributed to smoothing and reducing the variability of synthesis. Parameter $\gamma_j \in [0,1]$ was used to determine the degree of shift of synthetic coordinates towards local extremes, which helped to preserve the characteristic signature properties of the sample without distorting its statistical structure. The selection of parameters c and γ_j was performed empirically using a five-component stratified cross-validation method with the macro-recall target metric and control of the false positive rate. A grid search was performed for each parameter $c \in \{0,05; 0,1; 0,2; 0,5; 1,0\}$, $\gamma \in \{0;0,05;0,1;0,2;0,3\}$, after normalising the data to the range [-1,1]. Additionally, a limit on the median absolute deviation (MAD) was introduced:

$$|\gamma_j (e_j - x_j^{syn})| \leqslant 1.5 \cdot MAD_j, \qquad (1)$$

where $\gamma_j \in [061]$ – extreme gain for j-th coordinate, $e_j \equiv e_j(x)$ – local extremum (minimum or maximum) j-th numerical feature in the sample area x, x_j^{syn} – j-th coordinate of the synthetic sample, MAD_j – median absolute deviation for j-th coordinate within the attack class t, $|\cdot|$ – module.

This prevented the development of synthetic anomalies. The optimal values were set as $c = 0.2$ and $\gamma_j = 0.1$, since

they provided a stable trade-off between stability and preservation of local signatures with controlled synthesis variability in the selected CNN architecture.

A set of training examples for the attack class t:

$$S_t = \{x^{(i)}\}_{i=1}^{n_t}, \quad (2)$$

where t – index of the attack type (class), n_t – number of training examples of this class, $x^{(i)}$ – i-th feature vector divided into a numeric part $x_{num}^{(i)} \in R^d$ (with d numeric attributes) and the categorical part $x_{cat}^{(i)}$. The target number of examples after resampling is denoted as N_t^{target}. Enter parameters k – number of nearest neighbours used to estimate the local structure of the class, M – upper limit of relative class expansion, N_{min} – minimum allowed number of examples for rare classes after balancing.

The regular goal is calculated using the equation:

$$\tilde{N}_t = min(\lfloor n_t \cdot M \rfloor, n_{max}), \quad (3)$$

where $\tilde{N}_t$ – target number of examples before threshold correction, n_{max} – maximum number of examples among all abnormal (attack) classes, $\lfloor \cdot \rfloor$ – rounding down to an integer operation:

$$N_t^{target} = max(\tilde{N}_t, N_{min}). \quad (4)$$

That is N_t^{target} – final target number of class examples t after accounting as a relative constraint M, and the absolute threshold N_{min}.

Local statistics are defined for each sample. The set $N_k(x)$ includes k nearest neighbours within the class S_t based on Euclidean or mixed distance. Local median of numeric coordinates:

$$m_{loc}(x) = median(\{y_{num} : y \in N_k(x) \cup \{x\}\}), \quad (5)$$

where $m_{loc}(x) \in R^d$ – element-wise local median of numeric coordinates in the vicinity x, y_{num} – numeric part of a neighbouring vector y.

Percentiles are defined: $p_{10}(x)$, $p_{90}(x)$. Distance to the local median:

$$d(x) = \| x_{num} - m_{loc}(x) \|_2, \quad (6)$$

where $\|\cdot\|_2$ – Euclidean norm, $d(x) \geq 0$ – local distance of the sample x to its local median $m_{loc}(x)$.

It is used to calculate the deterministic displacement coefficient:

$$\beta(x) = clamp\left(\frac{d(x)}{d(x)+c}, 0\beta 1\right), \quad (7)$$

where $\beta(x) \in [0,1]$ – deterministic bias coefficient for the sample x, $c > 0$ – parameter that adjusts the sensitivity of $\beta(x)$ to the distance $d(x)$, $clamp(z, 0, 1) = min(max(z, 0), 1)$ – value constraint operator z with interval [0,1]. This is a deterministic function (no randomness).

In the standard Synthetic Minority Oversampling Technique (SMOTE), categorical values are selected as a mode among neighbours:

$$x_{cat,j}^{syn} = arg \max_v \#\{y \in \mathcal{N}_k(x) \cup \{x\} : y_{cat,j} = v\}, \quad (8)$$

where $x_{(cat,j)}^{syn}$ – value of j-th categorical feature in a synthetic sample, $y_{(cat,j)}$ – value of the same attribute in a neighbouring sample y, v – possible categorical value, $\#\{\cdot\}$ – power of the set (number of elements). Thus, for categorical features, the mode is selected in the neighbourhood x.

In the proposed SPAS method, the amplitude of local variation is defined as:

$$\Delta(x) := p_{90}(x) - p_{10}(x) \in R^d, \quad (9)$$

where $p_{10}(x)$, $p_{90}(x) \in R^d$ – element-wise 10th and 90th percentiles of numerical coordinates in the local neighbourhood x, $\Delta(x)$ – vector of the local amplitude of variation for each numerical feature.

Direction $s(x)$ is defined deterministically as the sign of the average deviation:

$$s_j(x) = sign\left(\frac{1}{k}\sum_{y \in \mathcal{N}_k(x)} (x_j - y_j)\right) \in \{-1,0,1\}, \quad (10)$$

where $s_j(x)$ – sign of average deviation of j-th coordinates of the sample x from its neighbours, x_j, y_j – value of j-th numerical attribute, respectively, for samples x and y, $sign(\cdot)$ – standard sign function that returns -1, 0, or 1.

A synthetic numerical vector is obtained by the equation:

$$x_{num}^{syn} = m_{loc}(x) + \beta(x)(\Delta(x) \circ s(x)), \quad (11)$$

where $x_{num}^{syn} \in \mathbb{R}^d$ – numerical part of the synthetic sample, $s(x) \in \{-1, 0, 1\}^d$ – direction vector for each coordinate, "$\circ$" – element-wise (Hadamard) multiplication of vectors. If for a certain coordinate , the offset direction is additionally set by the nearest neighbour.

Saving extremes (signature-preserving shift). For coordinates where extreme values exist in neighbours (outside), an additional offset is introduced:

$$x_j^{syn} \leftarrow x_j^{syn} + \gamma_j(e_j(x) - x_j^{syn}), \quad (12)$$

where x_j^{syn} – current value of j-th coordinates of the synthetic sample, $e_j(x)$ – local extremum (minimum or maximum) of j-th coordinates among the set $N_k(x) \cup \{x\}$, $\gamma_j \in [0,1]$ – coefficient that sets the fraction of "pull-up" to the extreme value, operator "$\leftarrow$" means updating the coordinate.

Categorical features in classical methods, the choice of categories is random, and in the proposed method:

$$x_{num}^{syn} = m_{loc}(x) + \beta(x)(\Delta(x) \circ s(x)), \quad (13)$$

in other words, the mode is selected. If several categories have the same frequency, the nearest neighbour

category is used. Index j numbers categorical attributes, and $x_{(cat,j)}^{syn}$ – value of j-th categorical feature in a synthetic sample.

The built formal device allows generating signature-resistant synthetic samples that preserve the structure of rare attacks. After mathematical specification, an experimental verification of the method was carried out on the NSL-KDD benchmark set, within which a comparison was made between the basic model without balancing, the variant with classic SMOTENC, and the model using the proposed SPAS approach In the basic model, which was trained on initially unbalanced data, the following results were obtained: the average error in the test sample was 0.6538, the classification accuracy was 0.7654, and the indicator of completeness of attack detection was only 0.62. This indicates the model's tendency to overestimate normal traffic and its limited ability to recognise rare types of attacks.

The analysis of the results obtained is consistent with the conclusions of other researchers. According to M. Farooq & F. Ahmad (2024), even well-optimised multilevel models show a significant drop in completeness for rare attacks due to the predominance of normal traffic, where convolutional architectures have been shown to retain sensitivity to class unevenness even after hyperparameter optimisation. This relationship can also be traced in the obtained experiments: the model overestimates the normal class and loses a significant part of the rare attacks, which corresponds to the general patterns described in contemporary studies. As shown in Table 1, the majority of errors in the baseline model occurred when attacks were classified as normal activity (3,484 false-negative triggers), while the number of false-positive classifications was relatively small (734 cases). This distribution confirms an imbalance in the representation of classes and a lack of ability of the model to detect rare events.

Table 1. Interface matrix

Actual \ Predicted	Pred = 0 (Normal)	Pred = 1 (Attack)
Actual = 0 (Normal)	8,977	734
Actual = 1 (Attack)	3,484	9,349

Source: compiled by the authors

Using the classic SMOTENC improved the situation: the test error decreased to 0.4573, the accuracy increased to 0.7900, and the completeness increased to 0.70. Partial indicators (Table 2) showed a significant increase in recall for the mscan (0.82) and neptune (0.99) attack types, although the values remained low for the warezmaster (0.26) and guess_passwd (0.30) subtypes, indicating a loss of signatures during interpolation.

Table 2. Recall indicators for individual types of attacks (clippings)

Attack	Support	TP	Recall
neptune	4,657	4,641	0.9966
guess_passwd	1,231	371	0.3014
mscan	996	819	0.8223
warezmaster	944	241	0.2553

Source: compiled by the authors

The results obtained on the use of classical SMOTENC are also consistent with the conclusions presented in the paper by P. Sathaporn *et al.* (2025), who demonstrated that interpolation methods can increase completeness for certain types of attacks, but simultaneously create the risk of distorting local statistical characteristics of data. The researchers emphasised that when modelling network anomalies, random or over-smoothed interpolation leads to the loss of signature patterns that are crucial for classifying complex or rare threats. This behaviour is particularly noticeable in conditions of significant imbalance, when small classes are represented by a limited number of samples, and therefore require a more accurate and structurally sound approach to synthetic complement. In the conducted experiment, this problem is manifested due to the low recall rates for warezmaster and guess_passwd, which confirms the system flaw of the standard SMOTE/SMOTENC. Figure 1 shows the value of attack detection completeness for the 20 most common classes when using regular SMOTE. The image shows a strong unevenness of the results: individual classes, in particular neptune, smurf, and portsweep, appear almost perfectly, while guess_passwd, mailbomb, and multihop remain virtually undetected.

Applying the proposed method, which takes signature properties into consideration when creating new examples, provided further improvements. In this case, the test error was 0.3985, the classification accuracy was 0.8129, and the completeness of attack detection was 0.73. Thus, the increase in completeness was 11% compared to the basic model and 3% compared to the classic SMOTENC. Results of measuring the time spent on training a convolutional neural network under various balancing strategies (Table 3). The basic model without balancing required 343 seconds for five training epochs, which corresponds to the average time of one epoch of 68.6 seconds. In the configuration with the proposed SPAS approach, which includes an additional stage of synthetic sample

addition based on modified SMOTENC, the total training time was 287 seconds, and the average epoch time was 57.4 seconds. Despite the increase in the size of the training sample (from 113,375 to 125,627 records), the training time does not increase, but rather decreases due to the stabilisation of the model behaviour and faster convergence. Thus, the additional costs of sample synthesis do not lead to a significant increase in the overall processing time and are acceptable for practical IDS systems.

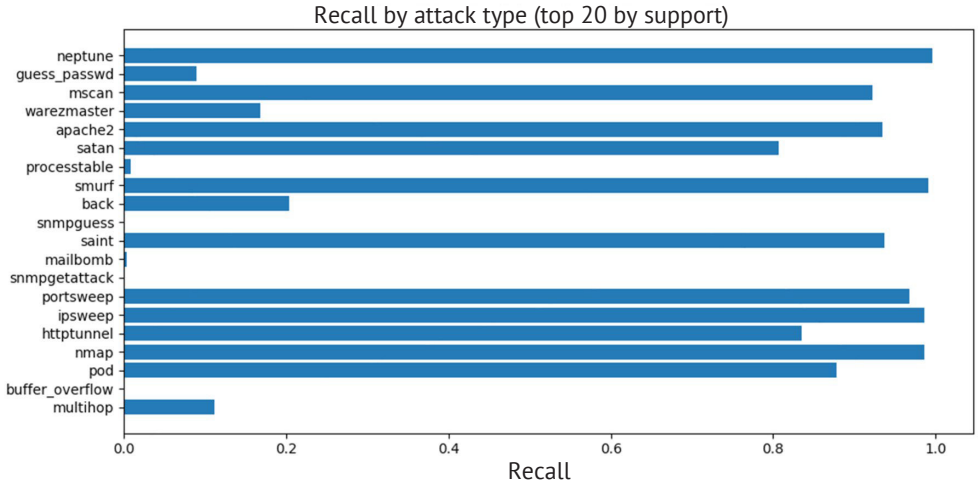

Figure 1. Recall by attack type conventional SMOTE

Source: compiled by the authors

Table 3. Time spent on different balancing configurations

Model configuration	Training sample size	Total training time, s	Average time of one epoch, s
No balancing (initial sampling)	125,627	343	68.6
Proposed SPAS (selective SMOTENC)	125,627	287	57.4

Source: compiled by the authors

Figure 2 shows a comparative diagram of detection completeness for the same types of attacks after applying the developed method. Guess_passwd and warezmaster scores improved significantly (an increase of 0.18 and 0.21, respectively), while the back and snmpguess classes approached the level of stable detection. The advantages of the signature-saving approach are consistent with a number of papers devoted to the use of local statistical features in anomaly detection problems, that the reproduction of local extremes and medians improves the resolution of models, which fully correlates with the result of increasing recall for problem attacks. That is why the SPAS approach provides not only higher averages, but also lower variability between attack types.

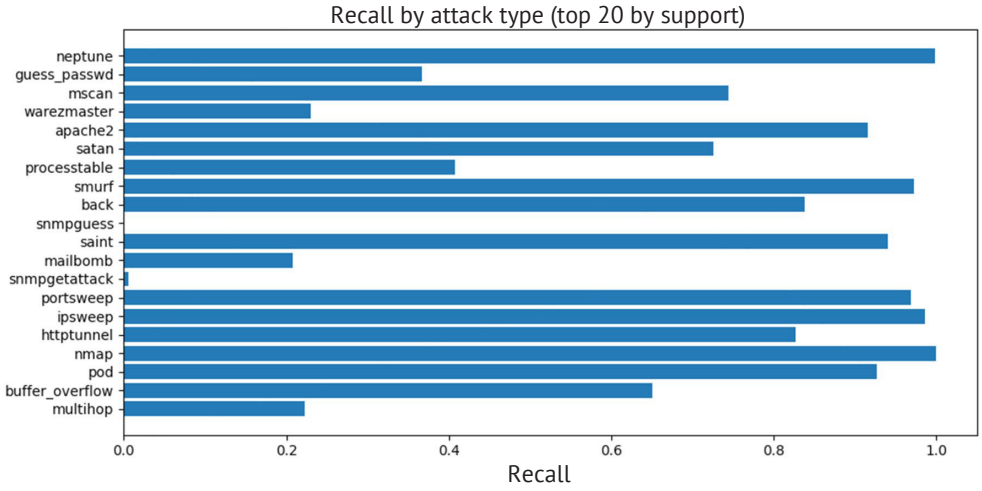

Figure 2. Recall by attack type offered by SMOTENC

Source: compiled by the authors

Along with an increase in the average values, there was a decrease in the spread between classes, which indicates a more stable behaviour of the model. Analysis of the error matrix confirms improvements in recognition of back, multihop, and rootkit attacks in the absence of an increase in false-positive positives. The difference does not exceed 18%, which makes the method suitable for use in practical IDS systems. Thus, the conducted experiments showed that the method of detecting computer attacks based on the adaptability and balancing of the training sample provides a sustainable improvement in the quality of the intrusion detection system due to the preservation of structural signatures of attacks and flexible adjustment of the volume of synthetic data supplementation without losing the overall classification accuracy.

In the broader context of modern balancing methods, the results of the proposed approach confirmed trends that emphasise the feasibility of adaptive control of the degree of data addition for different classes. The researchers point out that universal synthesis coefficients are not effective enough for highly sparse or structurally complex classes. The observations available in this paper fully correspond to this opinion: for a number of complex attacks (for example, guess_passwd, warezmaster), the adaptive approach provided a significant improvement due to deterministic bias formation and the use of local statistics, while the classical SMOTENC showed uneven results. Research on methods for detecting network attacks over the past decade has been intensively developing in two areas: improving deep classification models and developing approaches to balancing unbalanced samples. Despite a significant amount of studies, the problems of correct reproduction of rare attacks, stability of models on small classes, and preservation of signatures remain open.

One of the reference data sets for the analysis of IDS methods is NSL-KDD, which is widely used in the study of classification algorithms and evaluation of their resistance to class imbalance. According to M.S. Sheikh & Y. Peng (2022), this dataset remains suitable for analysing basic attack modelling patterns and improving recognition accuracy, making it a useful tool for experimental research in the field of network security. Although NSL-KDD is often criticised for breaking away from the real complexity of traffic, it provides a structured framework for comparing algorithms and reproducibility of experiments. Unlike its predecessor, NSL-KDD significantly eliminates the problem of an excessive number of identical records, which contributes to a more objective assessment of model performance and reduces the risk of shifting results towards duplicate samples. This characteristic is particularly important in the context of modern approaches to constructing IDS, where excessive correlation between samples can lead to an overestimated generalising ability of models. In addition, as noted by A.-G. Mari *et al.* (2023), NSL-KDD retains its popularity in intrusion detection systems research due to its balanced task structure, the presence of clearly delineated training and test subsets, and transparent classification of attack types, making it a convenient benchmark for benchmarking.

Moreover, numerous studies emphasise that the accuracy of IDS models significantly depends on the specifics of the data on which training is carried out. Z. Ahmad *et al.* (2021) noted that an imbalance between normal and abnormal traffic causes a significant decrease in the ability of models to detect complex or rare attacks. An additional complication is the high level of false-negative classifications reported by J. Mijalkovic & A. Spognardi (2022), emphasising that it is False Negative errors that pose the greatest danger to cyber defence systems, since they allow attacks to go unnoticed. A similar relationship was described by Z. Cai *et al.* (2023), who showed that an imbalance in training samples leads to a deterioration in the sensitivity of even well-optimised deep learning models. As stated by M. Sheibani *et al.* (2024), a comprehensive analysis of the sensitivity and specificity of the classifier allows adequately assessing its behaviour in cases where standard accuracy does not reflect the actual ability of the model to detect harmful actions. In this study, F_1-metric was used as the main indicator of balanced performance, while precision and recall were additionally monitored to assess the accuracy and completeness of classification in conditions of a significant imbalance in training data.

A separate area of research is devoted to improving the sustainability of IDS in IoT environments, where the volume, speed, and heterogeneity of traffic pose additional challenges. The paper by M. Baich & N. Sael (2025) emphasised that the intensive growth in the number of connected devices leads to the emergence of complex and persistent cyber threats, such as botnets, which are difficult to detect using classical models. The researchers demonstrated that the use of machine learning methods in combination with feature selection and data balancing mechanisms can significantly improve the quality of detection, reducing processing time, and increasing classification accuracy. M.H. Shahriar *et al.* (2020) showed that the use of generative adversarial networks (GAN) in IDS systems can serve as an additional mechanism for synthetic sample expansion, although the quality of the obtained samples significantly depends on the stability of the GAN model training.

Due to the problem of uneven representation of classes, a significant part of studies focused on the investigation of different approaches to data balancing. Classical resampling methods, in particular the SMOTE method and its modifications, are known for their ability to improve the representation of rare classes in a sample. J.H. Joloudari *et al.* (2023) showed that the use of SMOTE in combination with artificial neural networks significantly improves the accuracy of attack detection in the Bot-IoT Suite. The results show that classical synthetic complement methods often do not consider the local signature properties of rare attacks,

which can lead to a shift in the statistical data structure and a deterioration in the ability of models to correctly reproduce the features of complex samples.

The problem of feature selection also occupies an important place in contemporary research – IDS. Feature Selection methods allow reducing the dimension of data, increasing the information content of features, and reducing the risk of retraining. The paper by P. Maniriho *et al.* (2020) used a hybrid approach that combines Information Gain and Gain Ratio to select optimal subsets of features that are later used in the Random Forest model. This helped to obtain high accuracy of DoS classification, scanning, and MITM attacks. However, the researchers also noted that even with efficient feature selection, the problem of class imbalance remains a key factor that limits the ability of models to detect rare attacks.

Summarising, it can be noted that contemporary research confirms the need for methods that can work correctly with uneven samples, preserve the structural and signature properties of rare attacks, and ensure the stability of models in complex network environments. Despite the presence of numerous approaches, the problem of matching balancing, feature selection, and deep modelling remains insufficiently solved, which opens up opportunities for developing more adaptive and signature-saving methods that combine these components.

CONCLUSIONS

As a result of the research, a method for detecting computer attacks based on adaptive balancing of the training sample with the preservation of signature features was developed, which increased the efficiency of intrusion detection systems in environments with data imbalance. The proposed approach combined statistical sample analysis with a deterministic synthetic complement mechanism based on local medians, percentiles, and modal values. This approach ensures the preservation of characteristic attack structures and avoids distortion of internal patterns of network traffic during model training. A comparative experiment on the NSL-KDD reference set confirmed the effectiveness of the method: the completeness of attack detection increased from 0.62 to 0.73, and the overall classification accuracy increased from 0.76 to 0.81, while the average test error value decreased to 0.3985. The results show that the preservation of signatures during synthetic data supplementation increases the sensitivity of the model to rare attacks without reducing its generalising ability. A critical advantage was a significant improvement in the recognition of rare, structurally complex attacks, for which the classic SMOTENC was ineffective. Thus, the Recall indicator for the guess_passwd attack increased by about 0.18, and for warezmaster – by 0.21. This proves that deterministic signature-preserving synthesis is a reliable basis for improving the stability of intrusion detection systems. An important advantage of the proposed method is the determinism of the process of constructing new examples, which guarantees stability and reproducibility of the results. Integration of the method into the data processing pipeline – from converting network traffic to signal format (PCM) and building spectrograms to analysis using convolutional neural networks – helped to increase the efficiency of deep learning through a qualitatively balanced sample. Further research should focus on automating the selection of complement coefficients of the training sample, implementing Bayesian methods to optimise model parameters, and expanding the experimental base to data sets such as CIC-IDS2017 and UNSW-NB15. This will confirm the universality and adaptability of the proposed approach in real network traffic conditions and ensure further development of intelligent cyber defence systems.

ACKNOWLEDGEMENTS

None.

FUNDING

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Ahmad, Z., Khan, A.S., Shiang, C.W., Abdullah, J., & Ahmad, F. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Emerging Telecommunications Technologies*, 32(1), article number e4150. doi: 10.1002/ett.4150.

[2] Alladi, T., Chamola, V., Sikdar, B., & Choo, K.-K.R. (2020). Consumer IoT: Security vulnerability case studies and solutions. *IEEE Consumer Electronics Magazine*, 9(2), 17-25. doi: 10.1109/MCE.2019.2953740.

[3] Baich, M., & Sael, N. (2025). Enhancing machine learning model prediction with feature selection for botnet intrusion detection. *Engineering Proceedings*, 112(1), article number 55. doi: 10.3390/engproc2025112055.

[4] Cai, Z., Du, H., Wang, H., Zhang, J., Si, Y., & Li, P. (2023). One-dimensional convolutional Wasserstein generative adversarial network based intrusion detection method for industrial control systems. *Electronics*, 12(22), article number 4653. doi: 10.3390/electronics12224653.

[5] Dalou', J., Al-Duwairi, B., & Al-Jarrah, M. (2020). Adaptive entropy-based detection and mitigation of DDoS attacks in SDN networks. *International Journal of Computing*, 19(3), 399-410. doi: 10.47839/ijc.19.3.1889.

[6] Denysiuk, D., Savenko, O., Lysenko, S., Savenko, B., & Kashtalian, A. (2023). Method for detecting steganographic changes in images using machine learning. In *Proceedings of the 13th international conference on dependable systems,services and technologies (DESSERT)* (pp.1-6).Athens: IEEE.doi:10.1109/DESSERT61349.2023.10416453.

[7] Farooq, M., & Ahmad, F. (2024). Improved intrusion detection in IoT using multi-layered neural architectures. *International Journal of Computing*, 23(2), 268-273. doi: 10.47839/ijc.23.2.3546.
[8] Hussain, A., Sharif, H., Rehman, F., Kirn, H., Sadiq, A., Khan, M.S., Riaz, A., Ali, C.N., & Chandio, A.H. (2023). A systematic review of intrusion detection systems in Internet of Things using ML and DL. In *2023 4th international conference on computing, mathematics and engineering technologies (iCoMET)*. Sukkur: IEEE. doi: 10.1109/iCoMET57998.2023.10099142.
[9] Joloudari, J.H., Marefat, A., Nematollahi, M.A., Oyelere, S.S., & Hussain, S. (2023). Effective class-imbalance learning based on SMOTE and convolutional neural networks. *Applied Sciences*, 13(6), article number 4006. doi: 10.3390/app13064006.
[10] Joseph, J.E., Aleke, N.T., & Onyeanisi, O.P. (2025). Deep learning based intrusion detection system for network security in IoT system. *International Journal of Education, Management, and Technology*, 3(1), 119-138. doi: 10.58578/ijemt.v3i1.4539.
[11] Kashtalian, A., Sergii, L., Sachenko, A., Savenko, B., Savenko, O., & Nicheporuk, A. (2025). Evaluation criteria of centralization options in the architecture of multicomputer systems with traps and baits. *Radioelectronic and Computer Systems*, 2025(1), 264-297. doi: 10.32620/reks.2025.1.18.
[12] Kilincer, I.F., Ertam, F., & Sengur, A. (2021). Machine learning methods for cyber security intrusion detection: Datasets and comparative study. *Computer Networks*, 188, article number 107840. doi: 10.1016/j.comnet.2021.107840.
[13] Li, G., & Jung, J.J. (2023). Deep learning for anomaly detection in multivariate time series: Approaches, applications, and challenges. *Information Fusion*, 91, 93-102. doi: 10.1016/j.inffus.2022.10.008.
[14] Maniriho, P., Niyigaba, E., Bizimana, Z., Twiringiyimana, V., Mahoro, L.J., & Ahmad, T. (2020). Anomaly-based intrusion detection approach for IoT networks using machine learning. In *2020 international conference on computer engineering, network, and intelligent multimedia (CENIM)* (pp. 303-308). Surabaya: IEEE. doi: 10.1109/CENIM51130.2020.9297958.
[15] Mari, A.-G., Zinca, D., & Dobrota, V. (2023). Development of a machine-learning intrusion detection system and testing of its performance using a generative adversarial network. *Sensors*, 23(3), article number 1315. doi: 10.3390/s23031315.
[16] Mijalkovic, J., & Spognardi, A. (2022). Reducing the false negative rate in deep learning based network intrusion detection systems. *Algorithms*, 15(8), article number 258. doi: 10.3390/a15080258.
[17] Sathaporn, P., Krungseanmuang, W., Chaowalittawin, V., Benjangkaprasert, C., & Purahong, B. (2025). DDoS detection using a hybrid CNN-RNN model enhanced with multi-head attention for cloud infrastructure. *Applied Sciences*, 15(21), article number 11567. doi: 10.3390/app152111567.
[18] Shahriar, M.H., Haque, N.I., Rahman, M.A., & Alonso, M. (2020). G-IDS: Generative adversarial networks assisted intrusion detection system. In *2020 IEEE 44th annual computers, software, and applications conference (COMPSAC)* (pp. 376-385). Madrid: IEEE. doi: 10.1109/compsac48688.2020.0-218.
[19] Sheibani, M., Konur, S., Awan, I., & Qureshi, A. (2024). A multi-layered defence strategy against DDoS attacks in SDN/NFV-based 5G mobile networks. *Electronics*, 13(8), article number 1515. doi: 10.3390/electronics13081515.
[20] Sheikh, M.S., & Peng, Y. (2022). Procedures, criteria, and machine learning techniques for network traffic classification: A survey. *IEEE Access*, 10, 64806-64829. doi: 10.1109/access.2022.3181135.

Метод виявлення комп'ютерних атак на основі адаптивності та балансування навчальної вибірки

Богдан Семенюк

Аспірант
Хмельницький національний університет
29016, вул. Інститутська, 11, м. Хмельницький, Україна
https://orcid.org/0009-0001-8831-8835

Богдан Савенко

Доктор філософії, доцент
Хмельницький національний університет
29016, вул. Інститутська, 11, м. Хмельницький, Україна
https://orcid.org/0000-0001-5647-9979

Анотація. Актуальність роботи визначається тим, що в сучасних системах кіберзахисту зберігається проблема дисбалансу даних у задачах виявлення комп'ютерних атак. Нерівномірне представлення класів призводить до зниження здатності моделей машинного навчання розпізнавати рідкісні, але критично важливі типи вторгнень. Наявні методи синтетичного доповнення вибірок часто спотворюють структуру даних і призводять до втрати характерних сигнатур атак, що негативно впливає на достовірність класифікації. Метою дослідження було підвищення точності та повноти виявлення мережевих атак шляхом розроблення адаптивного методу балансування навчальної вибірки зі збереженням статистичних і сигнатурних властивостей аномалій. Методологія дослідження передбачала створення детермінованого алгоритму, який обчислював коефіцієнти доповнення окремо для кожного типу атаки з урахуванням його частоти та мінімально необхідного обсягу вибірки. Для генерації нових прикладів застосовувався статистичний підхід, заснований на локальних медіанах і крайніх значеннях сусідніх зразків, що забезпечувало відтворення типової структури аномальних патернів без випадкової інтерполяції. Розроблений метод інтегрувався у послідовність обробки даних, у межах якої параметри мережевого трафіку перетворювалися у звукові сигнали, а на їх основі формувалися спектрограми для подальшого аналізу двовимірною згортковою нейронною мережею. Основні результати дослідження показали підвищення повноти виявлення малопоширених типів атак у середньому на 10-12 % порівняно з базовими підходами, водночас зберігаючи стабільний рівень загальної точності. Збереження сигнатур у синтетичних зразках забезпечувало покращене розпізнавання рідкісних атак і підвищувало достовірність класифікації. Практична цінність роботи полягає у можливості застосування розробленого методу для формування збалансованих вибірок у системах виявлення вторгнень та його інтеграції з існуючими моделями глибинного навчання з метою підвищення надійності кіберзахисту корпоративних мереж

Ключові слова: дисбаланс даних; сигнатурне моделювання аномалій; синтетичне доповнення вибірки; аналіз мережевого трафіку; згорткові нейронні мережі; акустичне представлення даних; стійкість класифікації