INSTITUTE OF INFORMATION AND COMMUNICATION TECHNOLOGIES
BULGARIAN ACADEMY OF SCIENCES
CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **26**, No 1
Sofia • 2026 Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.2478/cait-2026-0007

ChSp-ACN: Channelized Spatial Attention Enabled Adam-Optimized Convolutional Neural Network for Intrusion Detection

Nishit Patil, Shubhlaxmi Joshi

School of Computer Science, Dr. Vishwanath Karad MIT-WPU, Pune, Maharashtra, India
E-mails: nsht.patil@gmail.com shubhalaxmi.joshi@mitwpu.edu.in

***Abstract*:** *Intrusion detection is a major concern in network systems where numerous smart devices are interconnected to handle sensitive information. Such interactions expose networks to threats like weak authentication, eavesdropping, malicious payloads, and a high false alarm rate. To address these challenges, a Channelized Spatial Attention enabled Adam optimized Convolutional Network* (*ChSp-ACN*) *is developed to identify malicious activities. The proposed ChSp-ACN effectively manages data imbalance and refines input features using KNN imputation and normalization. Its spatial attention mechanism enhances detection accuracy by focusing on relevant attack features, while the Adam optimizer fine-tunes model parameters to minimize false positives. Experimental evaluation demonstrates that ChSp-ACN achieves superior results compared to existing methods, attaining an accuracy of 97.22%, specificity of 97.24%, sensitivity of 97.20%, and a False positive rate of 0.03, which attains maximal performance effectiveness under the BoT-IoT dataset.*

***Keywords*:** *Channelized spatial attention, Cyber-attacks, Deep learning, Intrusion detection, Internet of Things* (*IoT*).

1. Introduction

In recent years, the emergence of a digital network environment with the utilization of the Internet of Things (IoT) has brought huge concern globally. Furthermore, the IoT plays a major role in various industrial domains such as energy, health care, and the manufacturing industry. Thus, the increased usage of IoT technology made significant lifestyle changes and provided several advances in technology [1].

Conceptually, IoT has several components such as sensors, communication channels, advanced security systems, control systems, Integration interfaces, smart house products, and vehicle network systems. The above-mentioned IoT things are connected to the internet and provide the desired performance on its basis [2, 3]. Thus, IoT facilitates effective communication between smart objects and the internet. Although IoT has various objectives, it shares certain common features [4]. Generally, IoT facilitates three different operations, including the collection phase, which is responsible for the collection of crucial physical data. The gathered data is transmitted during the transmission phase via several transmission mediums such as Digital Subscriber Line (DSL), Wi-Fi, and so on. The transmitted data is gathered by the respective physical medium in the utilization phase [5].

IoT serves as a significant paradigm in the Information and Communication Technology (ICT) industry and is extensively adopted in certain applications, including home automation applications, industrial procedures, health care, and environmental monitoring. Moreover, the internal attacks affect the host monitoring that relies on states, signatures, and system performance. Due to these obstacles, the data transmission process through the network causes security threats [4]. Nevertheless, the privacy and security of private information through IoT is still considered a difficult task. Moreover, the IoT has suffered from several cyberattacks and vulnerabilities. Hence, IDS is employed to act as a better protection aid for conventional networks and information systems by sending alert messages when the security is compromised [6]. Therefore, various approaches are undertaken to identify the intrusion behavior with suspicious activities in the network. These techniques effectively validate the network and enhance the security concern accordingly [7].

While evaluating traditional intrusion detection systems possessed certain drawbacks were identified, such as increased false alarm rates, imbalance issues, low detection rates, generalizability, and interpretability issues. Because of these limitations, the traditional methods reduced the reliability and performance of the model. So, researchers utilized Machine Learning (ML) approaches for detecting intrusion in the network. Some traditional ML techniques that are used in detecting network intrusion are Support Vector Machine (SVM), Random Forest (RF) [8], K-Nearest Neighbor (KNN), Decision Tree (DT) [9], Naïve Bayes [10], and other algorithmic techniques. However, these techniques incorporate minimal training accuracy with poor performance [11]. When compared with ML techniques, DL-based techniques provide superior detection accuracy with a lower false positive rate. Various Deep Learning (DL)-based techniques as Convolutional Neural Networks (CNN), Multi-Layer Perceptron (MLP) [12], Recurrent Neural Networks (RNN) [13], and other DL-based algorithms. It provides accurate detection accuracy but it consumes a long time for training. To address this limitation, develop a new model that generates accurate detection outcomes with a minimal false alarm rate and lower processing time [14].

The research is purposely developed for detecting the intrusion in Networks. The developed model incorporates various stages for detecting the internal attacks that affect the network efficacy. In the proposed Channelized Spatial attention enabled Adam optimized Convolution Network (ChSp-CAN), the unorganized input data sources are effectively organized by the data imputation process and proceed for further detection operation. The main aspects of the article are expressed as follows.

Adam optimizer. The Adam optimizer effectively measures the adaptive rate of every parameter for enhancing the generalization ability with effective computation. Thus, the developed optimization achieved a better convergence rate with minimal operating time.

ChSp-ACN: The ChSp-ACN effectively detects internal attacks by evaluating the correlation of the input data and target data variables based on the non-linearity principle. This Adam optimization-based ChSp-ACN model enhanced the training process and attained better detection results.

The remaining research is organized as follows: Section 2 explains the existing intrusion detection approaches related to this study, Section 3 discusses the developed model scheme, and the evaluation of the experiment is illustrated in Section 4, along with a comparative assessment. Convincingly, the conclusion of the proposed ChSp-ACN is detailed with possible trends for future reference in Section 5.

2. Literature review

The recent conventional intrusion detection techniques are briefly explained in the section.

Ullah et al. [15] developed an Intrusion Detection System based on Transformer Neural Network (TNN-IDS) to determine malicious attacks in the networks. In TNN-IDS, the detection efficacy of malicious activity in the IoT network was optimum. However, TNN-IDS was highly affected by imbalanced data issues and high false positive rates.

Jullian et al. [16] introduced a Distributed Attack Detection Framework (Dist-ADF), which eliminated diverse sources of vulnerabilities in the network effectively. The method utilized the BoT-IoT and NSL-KDD datasets, from which it effectively detected the cyber-attacks in IoT environments. Conceptually, the method deployed with better-correlated features to provide efficient intrusion detection outcomes. However, the model faced minimal performance efficiency with misclassification and over-fitting problems.

Sana et al. [17] employed a detection system with the inclusion of Vision Transformers (ViT). The method improved the security system of the network by preventing diverse attacks, which affected the network. Meanwhile, the reliability and generalization ability of the model were increased effectively. However, the developed ViT model faced certain limitations, including increased computational cost problems, minimal efficacy, and data imbalance issues.

Mohy-Eddine et al. [4] established an Isolation Forest-enabled Pearson's Correlation Coefficient (IF-PCC) model to detect cyberattacks. The model effectively reduced the time complexity problem with a better convergence rate and training time. Based on these advances, the ability of the model was enhanced significantly. But the IF-PCC model suffered from misclassification issues.

Chalichalamala, Govindan and Kasarapu [18] presented a Logistic Regression-based Ensemble Classifier (LR-EC) model, which effectively eliminated the imbalance issue and provided better detection outcomes. In the method, the significant features were extracted to stimulate the determination performance. However, the model suffered from increased minority samples that resulted in decreased performance efficacy with overfitting and generalization ability problems.

Alhenawi et al. [19] presented a hybrid feature selection framework for detecting anomalies via an Improved Intelligent Water Drop methodology. This method assists in decreasing the number of features in all the datasets. However, certain downsides, including an imbalance in the data and overfitting issues, affected the performance of the presented research.

Solimn, Oudah and Aljuhani [6] propounded a DL-based Intelligent Intrusion Detection System (DL-IIDS) to detect cyber-attacks in the industrial environment. Here, Singular Value Decomposition (SVD) was utilized to select only the appropriate features; the inclusion of the Synthetic Minority Over-sampling Technique (SMOTE) aided in the better prevention of the over-fitting problems that arise in the network. Yet, the presented framework failed to detect the unknown attacks that occur in the system.

Wang, Xu and Liu [20] suggested a Residual Transformer-enabled Bidirectional Long Short-Term Memory (Res-TranBiLSTM) model to facilitate intrusion detection. The SMOTE-Edited Nearest Neighbor (SMOTE-ENN) approach utilized in this framework alleviated the data imbalance problem, followed by the extraction of significant spatiotemporal features, which were further subjected to Res-TranBilSTM to determine and categorize the intrusions in the IOT network. Conversely, the instability nature of the collected data affected the accuracy of intrusion detection.

Zhang [21] demonstrated an ensemble learning model that incorporated CNN-GRU, CNN-LSTM, and DNN models for intrusion detection. These lightweight models attained high performance and reduced inference time, making them applicable for real-time IoT networks. Although the ensemble model has high robustness and predictive accuracy compared to single models, resource constraints affect generalizability on unseen data and introduce computational complexity.

Zhou and Li [31] presented a bio-inspired Resampled Spiking Neural Network (SNN) for intrusion detection, which utilized non-leaky neurons for effective extraction of spatio-temporal features. This model was developed to overcome the existing limitations, such as poor efficiency and latency issues faced in SNNs. However, it struggled with high computational cost.

There are some challenges.

- The TNN-IDS model was affected by potential and flexibility challenges concerning parameter overhead that, in turn, led to an imbalanced data problem occurring in the network, which minimized the intrusion detection performance of the model [14].
- In the Dist-ADF method, diverse cyber-attacks in distinct fields of the network were not detected properly. In addition, the model suffered from over-fitting and misclassification problems [16].
- While evaluating traffic anomalies, the ViT model did not capture the temporal dependencies of the model, which affected the ability and efficiency with increased computational complexities [17].

3. Channelized spatial attention enabled Adam optimized convolution network for detecting intrusion in the networks

In this work, a ChSp-ACN is designed to detect intrusion in IoT networks accurately. For this research, [24, 25] are used to provide the input required for intrusion detection, which is incorporated into the pre-processing phase. Both KNN imputation and normalization methods are used, where the data normalization technique effectively reduces the repetitive data in the input, whereas the KNN imputation finds the missing data instance and reorganizes the input data into quality form. The quality-improved data are subjected to the ChSp-ACN model, which incorporates the DCNN model to detect the intrusions. Further, the Adam optimizer is imposed in the model to achieve a better convergence rate with effective computation. Moreover, the ChSp attention mechanism is employed to speed up the intrusion detection phenomena by concentrating only on the intrusion-related features. Additionally, the optimizer tunes the bias and weights of the ChSp-ACN to reduce the optimization issues and enhance the security quality effectively. From this perspective, the ChSp-ACN effectively detects the intrusion in the IoT network with better reliability. The basic schematic illustration of the ChSp-ACN is mentioned in Fig. 1.

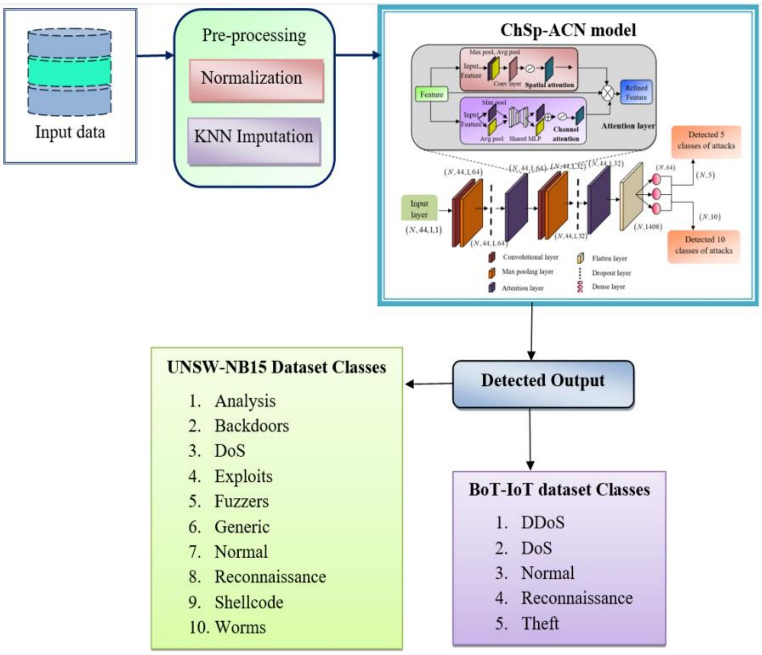


Fig. 1. Basic schematic illustration of the proposed intrusion detection method

3.1. Data input

The research captures the input from [24] and [25], which includes both botnet traffic and normal networks. The dataset source incorporates PCAP files, Argus files, and CSV files, which are categorized as attack and normal categories for achieving a proper labeling process. The obtained data input for the research is denoted as

$$F=\{k_1,k_2,...,k_N\}, \quad (1)$$

where F indicates the dataset with input data attributes $\{k_1,k_2,...,k_N\}$. The input dimensions for the UNSW-NB15 dataset are $(N, 45)$. Similarly, the input data dimensions for the BoT-IoT dataset are $(N, 20)$.

3.2. Pre-processing

In general, the data inputs suffer from various problems as irrelevant data, inaccurate data, missing data, and repetition of data. To enhance the quality of the input data, the above-mentioned problems should be eliminated effectively. In the research, the irrelevant and missing data are effectively reduced via the data normalization and KNN imputation process, which are elaborated as follows.

The basic problem that occurs in data attribute inputs is imbalance and missing data values, which do not capture significant information about the data, which causes an impact on detection performance. These limitations are resolved by the most familiar imputation technique called KNN imputation, which effectively provides a solution for many missing value imputation problems. Conceptually, the KNN imputation technique uses the Euclidean distance matrix, which identifies the nearest neighbor and imputes data for missing values [26]. In this context, both continuous and discrete data values are effectively predicted by the KNN technique. Initially, the method analyzes the current attributes based on distance evaluation to find whether the attributes have any missing values. These observed values are determined by the Euclidean distance matrix, which is mathematically notated as,

$$C = \sqrt{\mathrm{wei} \times \left(g_{\mathrm{dis}}\right)^2}. \tag{2}$$

Here g_{dis} denotes the distance of the present attributes that belong to F, and wei defines the weight of the Euclidean matrix, which is mathematically expressed as

$$\mathrm{wei} = \frac{h}{g}, \tag{3}$$

where h represents the total number of data attributes, g mentions the available data attributes.

Based on these evaluations, the KNN method determines the shortest distance attributes, which are used in the imputation process, and analyzes the shortest attributes with smaller distance values [27]. The outcome of KNN-imputation is mentioned Q.

A data normalization process is required for diverse data features that are generated in the CSV file source. In this context, the unstructured and redundant data are eliminated effectively from the data source and arranged in an organized manner for further detection processes. Conceptually, the normalization is achieved by both adaptive and sliding window approaches. In the adaptive method, the normalization process takes place in a sequence by considering the minimal and maximal values of the sequence. Consequently, the sliding window approach performs the normalization process by exploring the global maximum and minimum values of disjoint sliding windows [28]. In spite of these techniques, the unwanted data are effectually removed and organized in a structured manner, which further allows to perform the detection process. The outcome of the normalization process from the input Q is depicted as S, with attained dimensions (N, 19) for the BoT-IoT dataset; comparably, the dimension of the UNSW-NB15 dataset is (N, 44).

3.3. Channelized Spatial attention enabled Adam optimized Convolution Network model

The attained organized data are then allowed into the ChSp- ACN model for detecting the intrusion in networks. The achieved large amount of pre-processed data is effectually transferred through a series of convolutional, max-pooling, dense layer, flatten, and dropout layers, which detect the true label of data for capturing the intrusion attacks that affect the network performance. In this context, the developed ChSp-ACN model utilized 128 and 216 filter units for the convolutional layer, whereas the dense network utilizes 256 filters. Based on these filters, the output phase attains a 0.1 dropout rate. Further, in the ChSp-ACN model, the convolutional layer employs activation of Rectified Linear Units (ReLU), and at the output phase, the model utilizes the Softmax activation layer. Based on these activation layers, the ChSp-ACN model effectively detects the internal attacks of the networks. Meanwhile, the non-linearity and ability of the ChSp-ACN model are also improved by evaluating the correlation between the input and target variables. In order to improve the detection performance, a Channelized Spatial attention module plays a pivot role by concentrating only on the significant features. However, the model has numerous hidden layers, which consume high computational costs. In order to reduce the cost complexity issues, Adam optimizers are effectively enabled in the ChSp-ACN model, which enhances the training process with the imputation of faster leverage of the Graphics Processing Unit (GPU). Contextually, the GPU enhances the operating speed with effective time acceleration conditions. In addition, the weight regularization in the Adam optimizer limits the overfitting problem and reduces the loss function significantly [29]. Based on these executions, the generalization ability and robustness are improved while detecting the intrusion in IoT-based networks.

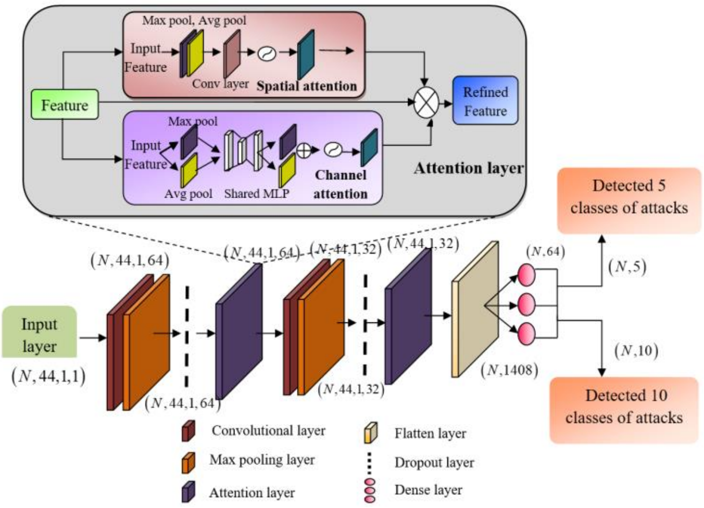


Fig. 2. Architecture of the ChSp-ACN model

Fig. 2 intercepts the architecture of the ChSp-ACN, which detects intrusion attacks in IoT-based networks. In this research, the quality-enhanced data obtained from the pre-processed phase with dimensions (*N*, 44) and (*N*, 19) is reshaped in (*N*, 44, 1, 1), (*N*, 19, 1, 1) which is assigned as the input to the input layer of the developed ChSp-ACN model. The resultant input layer output is modeled as $\mathbb{C}_{\mathrm{rs}}$.

Next is the convolutional layer, which receives $\mathbb{C}_{\mathrm{rs}}$ as its input and ensures the informative attributes of the data by performing a convolution operation via the effective utilization of the required kernel to produce the feature map as its output, and is expressed as $\mathbb{C}_{\mathrm{cn}}$ with the dimensions (*N*, 44, 1, 64) which are formulated in

$$\mathbb{C}_{\mathrm{cn}} = \vartheta\left(\mathbb{C}_{\mathrm{rs}} \times \nu\right) + \beta, \tag{4}$$

where, ϑ indicates the ReLU activation function, ν denotes the Adam optimized weight value, β and signifies the learning parameter. The Adam optimizer is the most widely accepted DL Algorithm utilized for the effective measurement of the adaptive rate of every parameter in the proposed ChSp-ACN because of its ability to overcome vanishing gradients and stimulate optimal convergence.

The feature map $\left(\mathbb{C}_{\mathrm{cn}}\right)$ thus obtained as output in the convolution layer $\left(\mathbb{C}_{\mathrm{cn}}\right)$, in turn, is further allowed into the max-pooling layer, which reduces the over-fitting problem through the exploration of the down-sampling process. The obtained outcome becomes $\mathbb{C}_{\mathrm{mp}}$ with dimensions (*N*, 44, 1, 64). The attained feature map is then projected into a dropout layer that effectively reduces the overlapping challenges and obtains better detection results modeled as $\mathbb{C}_{\mathrm{drp}}$ with dimension (*N*, 44, 1, 64). The feature map $\mathbb{C}_{\mathrm{drp}}$ thus obtained is injected into the upcoming ChSp attention module. The ChSp attention module [28] signifies the combination of both the channel and spatial attention approaches. In the Channel attention module, primarily the spatial dimension is squeezed to determine only the meaningful feature required for intrusion determination via the channel attention map generation. After that, the squeezed spatial details $\left(\mathbb{C}_{\mathrm{cn}}\right)$ are aggregated to generate two distinct spatial descriptors defining the average-pooled and max-pooled features, and are modeled as $\Im_{\mathrm{arg}}$ and $\Im_{\mathrm{max}}$.

The resultant spatial descriptors are modeled as $\Im_{\mathrm{arg}}$ and $\Im_{\mathrm{max}}$ are then passed into the shared network comprising the Multi-Layer Perceptron (MLP) to leverage the channel attention map as its output. The mathematical notation for the MLP output $\left(\mathbb{N}_{\mathrm{mlp}}\right)$ obtained after the application of MLP on$+\Im_{\mathrm{arg}}$ and $\Im_{\mathrm{max}}$ the via element-wise summation operation is expressed in

$$\mathbb{N}_{\mathrm{mlp}}\left(\mathbb{C}_{\mathrm{cn}}\right) = \varsigma\left(\mathbb{N}_{\mathrm{mlp}}\left(\Im_{\mathrm{arg}}\left(\mathbb{C}_{\mathrm{cn}}\right)\right) + \mathbb{N}_{\mathrm{mlp}}\left(\Im_{\mathrm{max}}\left(\mathbb{C}_{\mathrm{cn}}\right)\right)\right). \tag{5}$$

In this expression, $\mathbb{N}_{\mathrm{mlp}}\left(\mathbb{C}_{\mathrm{cn}}\right)$ highlights the MLP operation on $\mathbb{C}_{\mathrm{cn}}$, ς unveils the sigmoid activation function, $\Im_{\mathrm{arg}}\left(\mathbb{C}_{\mathrm{cn}}\right)$ and $\Im_{\mathrm{max}}\left(\mathbb{C}_{\mathrm{cn}}\right)$ defines the MLP on $\Im_{\mathrm{arg}}$ and $\Im_{\mathrm{max}}$. The final feature map obtained as output using Channel attention becomes A_{ch} and is formulated in

$$\mathrm{A}_{\mathrm{ch}} = \varsigma\left(\nu\left(\Im_{\mathrm{arg}}\left(\mathbb{C}_{\mathrm{cn}}\right)\right) + \nu\left(\Im_{\mathrm{max}}\left(\mathbb{C}_{\mathrm{cn}}\right)\right)\right). \tag{6}$$

Simultaneously, the inter-spatial relationships of the feature maps present $\mathbb{C}_{\mathrm{cn}}$ are utilized to stimulate the generation of spatial-attention maps to facilitate efficient detection. Conceptually, only the relevant features are extracted by using spatial attention [5] via the application of both the max-pooling and average-pooling operations on $\mathbb{C}_{\mathrm{cn}}$ which are then concatenated to offer only meaningful features $(\mathbb{N}_{\mathrm{cctn}})$ and are represented in the following equation. After that, the concatenated feature $(\mathbb{N}_{\mathrm{cctn}})$ is fed into the convolution layer to produce the spatial attention map as its output, which is denoted as A_{sp},

$$(7) \qquad \mathbb{N}_{\mathrm{cctn}} = \left\{ \left\| \Im_{\mathrm{arg}}\left(\mathbb{C}_{\mathrm{cn}}\right) \right\| \left\| \Im_{\mathrm{max}}\left(\mathbb{C}_{\mathrm{cn}}\right) \right\| \right\},$$

$$(8) \qquad \mathrm{A}_{\mathrm{sp}} = \varsigma\left(\gamma\left(\left\| \Im_{\mathrm{arg}}\left(\mathbb{C}_{\mathrm{cn}}\right) \right\| \left\| \Im_{\mathrm{max}}\left(\mathbb{C}_{\mathrm{cn}}\right) \right\| \right)\right),$$

where γ depicts the convolution operator with a fixed filter size. Furthermore, the features thus obtained in the Equation (4), (6), and (8) then undergo an element-wise multiplication operation $(\otimes)$ to produce the critical features required for intrusion detection, and are represented in

$$(9) \qquad \mathbb{C}_{\mathrm{chsp}} = \mathbb{C}_{\mathrm{cn}} \otimes \mathrm{A}_{\mathrm{ch}} \otimes \mathrm{A}_{\mathrm{sp}}.$$

The ChSp attention module output is expressed as $\mathbb{C}_{\mathrm{chsp}}$ and has a dimension of. Thus, the critical features obtained are then passed through several convolutions and attention modules to produce feature vectors with dimension (*N*, 44, 1, 32). The features thus attained with dimensions are allowed into the flattened layer, which effectively generalizes the model's ability. Based on this generalization ability, the flattened layer output (*N*, 1408) becomes $\mathbb{C}_{\mathrm{fl}}$. Meanwhile, the resized dimensions are then subjected to the dense layer that achieves better interpretability information of data variables and generates $\mathbb{C}_{\mathrm{dns}}$ as output with dimensions (*N*, 64), which in turn effectually reduces the loss function and achieves precise outcomes of intrusion detection with dimensions (*N*, 5), and (*N*, 10) for the BoT-IoT dataset and UNSW-NB15 datasets, respectively. In this, the intrusion has the following class of attacks, such as DDoS, DOS, normal, reconnaissance, and theft in the BoT-IoT dataset, whereas the attacks of Analysis, Backdoors, DoS, Exploits, Fuzzers, Generic, Normal, Reconnaissance, Shellcode, and Worms in UNSW-NB15 datasets.

4. Results

The section depicts the entire performance of the ChSp-ACN during the evaluation of intrusion detection. Along with this, the performance of the developed ChSp-ACN is compared with other conventional methods, and a discussion about the detection accuracy is significant.

4.1. System specification

The research is conducted on a Windows 11 system configuration, with a memory of 128 GB ROM and 16 GB RAM. The implementation of the proposed ChSp-ACN model utilized a Python application with Version 3.7 at an available clock speed of 3 GHz. In this context, the PyCharm software version 2024.2.1 is utilized for implementing the code for detecting the intrusion in the IoT network. The initial parameters involve a batch size of 32, a Learning rate of 0.001, several epochs of 500, a Dropout rate of 0.2, an Activation function of "ReLU", a Loss function of "Categorical cross-entropy", with default Optimizer Adam.

4.2. Dataset description

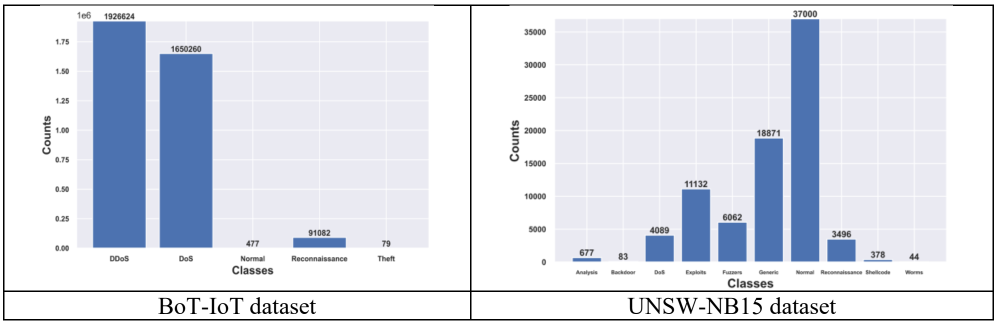
In the research, the input data is captured from the BoT-IoT and UNSW-NB15 datasets to perform the detection of attacks in the IoT-based networks, which are briefly elaborated as follows,

BoT-IoT dataset [24] contains more than 72,000,000 records, which are in the size of 69.3 GB. Mainly, the BoT-IoT dataset is assembled by botnet traffic and normal network environments, which contains original PCAP files, generated CSV files, and generated Argus files. These source files of the dataset are detached by attack and subcategory for performing the evaluation process. The dataset includes DDoS, DoS, Normal, Reconnaissance, and Theft attacks. By considering this dataset, the research effectively performs the intrusion detection process.

UNSWNB15 dataset [25] is generated in a cyber range lab for estimating the activities of normal and the behaviors of contemporary attacks. The dataset contains nine labeled class features with the inclusion of nine attacks, such as Backdoors, Generic, Shellcode, Fuzzers, DoS, Analysis, Exploits, worms, normal, and Reconnaissance. Thus, these attack-insured dataset is obtained in the form of CSV files, which are utilized in the research of intrusion detection.

4.3. Class distribution

Fig. 3 displays the class distribution information of the BoT-IoT dataset and UNSW-NB15 dataset utilized in the proposed ChSp-CAN model. About 37,000 counts are categorized as normal in the UNSW-NB15 dataset.



BoT-IoT dataset

UNSW-NB15 dataset

Fig. 3. Class distribution of datasets

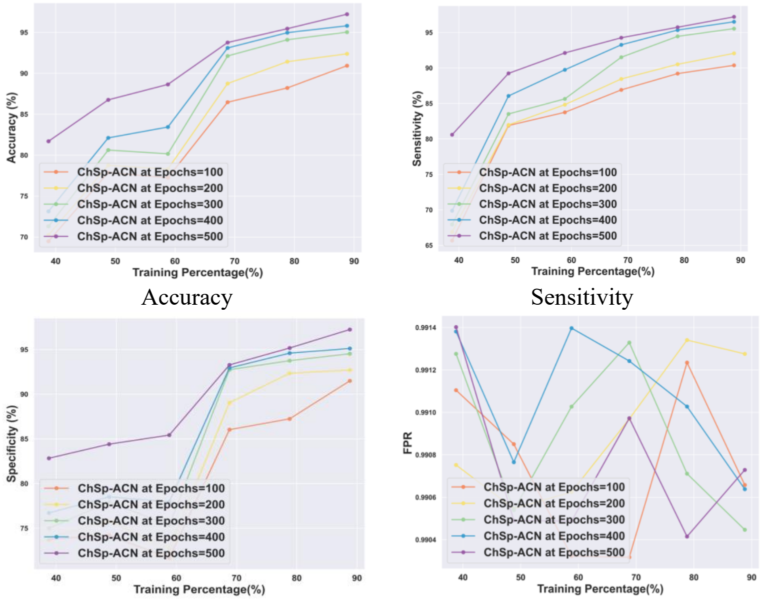
Further, the rest of the 677 counts fall under the analysis category, 83 in backdoor, 4089 in DoS, 11,132 in exploits, 6062 in fuzzers, 18,871 in generic, 3496 in reconnaissance, 378 in shellcode, and 44 in worms. Meanwhile, the BoT-IoT dataset comprises 1,926,624 counts under DDoS, 1,650,260 in DOS class, 477 in normal class, 91,082 in reconnaissance, and 79 under theft category.

4.5. Performance analysis

In this sub-section, the performance of the proposed ChSp-ACN is analyzed by varying the epoch values and for various percentages of training data in terms of certain metrics, including accuracy, specificity, sensitivity, and False Positive Rate (FPR), which are detailed as follows.

4.5.1. Performance analysis of the proposed ChSp-CAN using the BoT-IoT dataset

The performance measure of the proposed ChSp-ACN using the BoT-IoT dataset is elucidated in Fig. 4, and it is clear that the proposed ChSp-ACN gets an accuracy of 97.22% for 90% of the training data with the epoch value of 500. Similarly, the sensitivity and specificity values achieved by the proposed ChSp-ACN for the same 90% of training data with the epoch value of 500 lie in the range of 97.20% and 97.24% respectively. Hence, it is evident that the inclusion of the ChSp attention mechanism and the Adam optimizer in the proposed ChSp-ACN facilitated accurate detection of the intrusion in the IoT network by selecting only the meaningful features by downsizing the over-fitting and local convergence problem that arises during the intrusion determination procedure.



Accuracy

Sensitivity

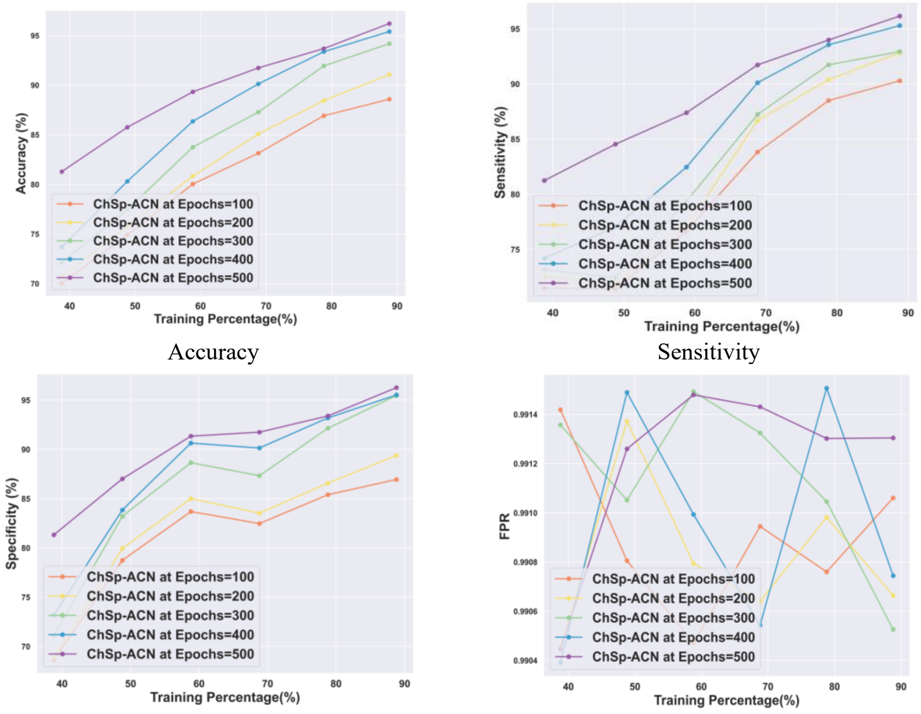
Specificity

FPR

Fig. 4. Performance Analysis of the ChSp-ACN model using the BoT-IoT dataset

4.5.2. Performance analysis of proposed ChSp-ACN using the UNSW-NB15 dataset

Fig. 5 displays the effectiveness of the proposed ChSp-ACN using the UNSW-NB15 dataset. Generally, for varying percentages of training data and with varying epoch sizes, the performance metric values also change. Accordingly, the accuracy, specificity, sensitivity, and FPR measures obtained by the proposed ChSp-ACN for 90% of training data and with the epoch value of 500 are at the order of 96.21%, 96.27%, 96.16%, and 0.99, respectively. Thus, this achievement in the proposed ChSp-ACN is due to the accomplishment of optimal convergence via the utilization of the Adam optimizer and the ChSp attention mechanism responsible for the selection of relevant features, thereby reducing the computational issues that arise in the attack detection framework.



Accuracy

Sensitivity

Specificity

FPR

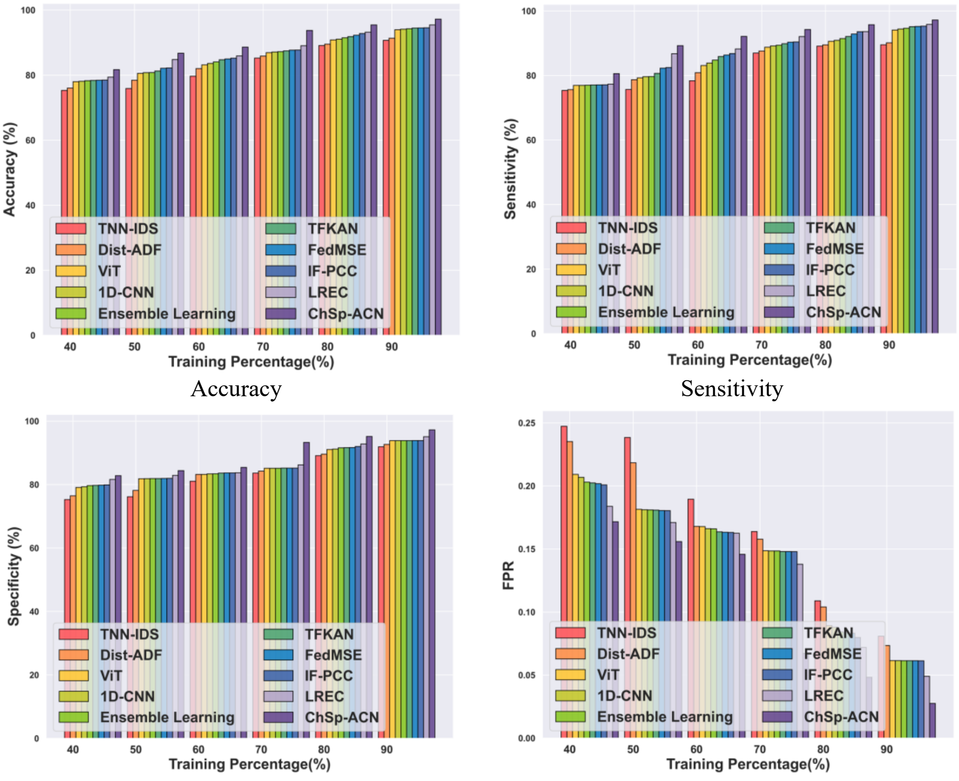
Fig. 5. Performance Analysis of the ChSp-ACN model using the UNSW-NB15 dataset

4.6. Comparative estimation

The effective performance of the ChSp-ACN for detecting the intrusion in the IoT-based models is significantly compared with other state-of-the-art models such as TNN-IDS [14], Dist-ADF [16], ViT [17], IF-PCC [4], LREC [18], 1D-CNN [23], Ensemble learning [21], TFKAN [30], and FedMSE [22]. The comparative analysis of these models is explained in the following sections.

4.5.1. Comparative estimation of the BoT-IoT dataset

In this context, the ChSp-ACN model is effectively evaluated by numerous performance measures as accuracy, sensitivity, specificity, and FPR. While detecting intrusion in IoT-based networks using the BoT-IoT dataset at a maximal training percentage of 90%, the obtained value of accuracy under the ChSp-ACN model is 97.22%, which shows a development of 2.8% over TFKAN and outperforms other existing models. Likewise, the attained value of sensitivity for the ChSp-ACN model is 97.2%, which shows an improvement of 2.12% over the TFKAN model. Similarly, the achieved value of specificity under the ChSp-ACN model is 97.24%, which is 3.48% higher than the TFKAN model. Meanwhile, the proposed model gained a minimum FPR of 0.03 and attained a difference of 0.02 and above over existing models. Thus, the overall comparative estimation of the ChSp-ACN model under the BoT-IoT dataset is schematically expressed in Fig. 6.



Accuracy

Sensitivity

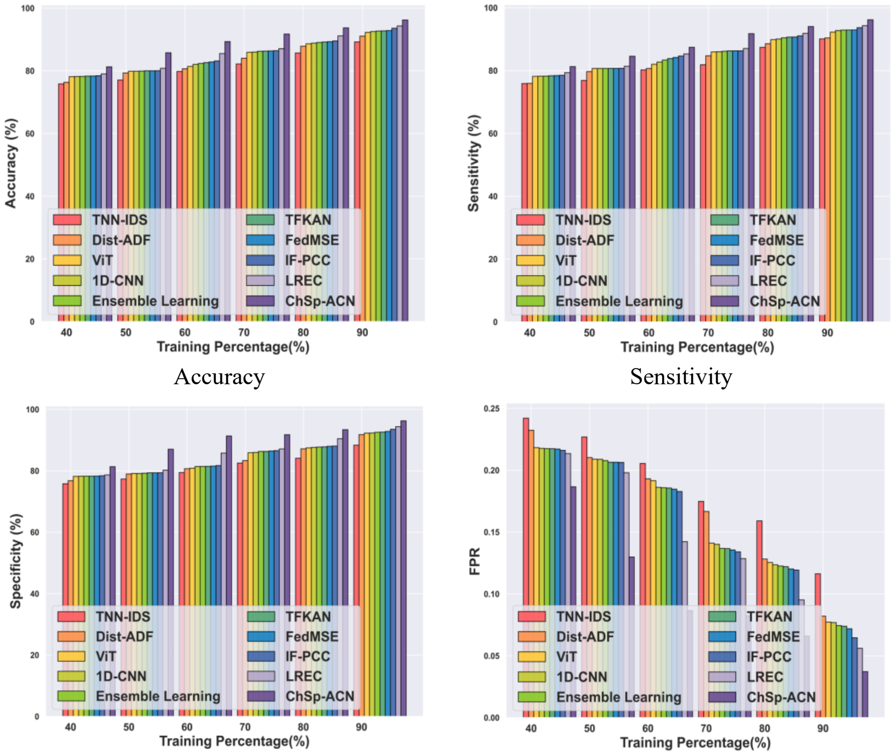
Specificity

FPR

Fig. 6. Comparative evaluation of the ChSp-ACN model under the BoT-IoT dataset

4.5.2. Comparative estimation of the UNSW-NB15 dataset

Fig. 7, intercepts the comparative estimation of the ChSp- ACN under the UNSW-NB15 dataset. During evaluation at 90% training, the achieved accuracy value of the proposed ChSp-ACN model is 96.22%, which is 3.61% greater than the TFKAN model. Similarly, the achieved value for sensitivity under the ChSp-ACN model is 96.16%, which shows an improvement of 3.4% for the TFKAN model. Consequently, the specificity shows an improvement over the TFKAN method is 3.81%, whereas the achieved value of specificity for the ChSp-ACN model is 96.28%. Moreover, the proposed model reached an FPR of 0.04, which is 0.04 less than the TFKAN model and outperforms the other models. Based on these estimation techniques, the performance of the proposed ChSp-ACN model exceeded the other existing models and is highlighted effectively.



Accuracy

Sensitivity

Specificity

FPR

Fig. 7. Comparative estimation of ChSp-ACN under the UNSW-NB15 dataset

4.6. Comparative discussion

In the intrusion detection research, the proposed ChSp-ACN model effectively detects intrusions or internal attacks that affect the IoT networking models. In addition, the supremacy of the various intrusion detection approaches is comparatively analyzed with state-of-the-art methods. The existing TNN-IDS [26] approach suffers from certain downsides, including the issues related to flexibility, imbalance problems that occur in the data, parameter overhead issues, and limitations in potential characteristics, which in turn are overcome by the inclusion of the data imputation phase in the proposed ChSp-ACN. Furthermore, the detection accuracy of the conventional Dist-ADF was highly affected by the misclassification problems that occur during the determination of the internal attacks on IoT-based networks [12]. Moreover, the computational complexity of the traditional ViT [3] was highly increased with the corresponding increment in the False Alarm Rate (FAR). Likewise, the obstacles faced by the state-of-the-art IF-PCC [14] decreased the attack determination ability. Meanwhile, LREC suffered from misclassification, generalization ability, and interpretability issues that affect the detection performance [7]. Further, the state-of-the-art models, such as 1D-CNN, Ensemble Learning, TFKAN, and FedMSE models, struggle with computational complexity and limited resources. To overcome these disadvantages, the ChSp-ACN model is developed, which offers robust performance and aids in the precise determination of the abnormalities via the selection of significant features using the modified attention mechanism. Furthermore, optimization of Adam tunes the model parameters for accurate intrusion detection with high performance. The benefits of these techniques overcome the vanishing gradient problem and lead to accurate intrusion detection. The discussion regarding the performance of the proposed ChSp-ACN with other existing models is tabulated in Table 1.

Table 1. Comparative discussion of the ChSp-ACN model

Method	Training percentage – 90							
	BoT-IoT dataset				UNSW-NB15 dataset			
	Accuracy (%)	Sensitivity (%)	Specificity (%)	FPR	Accuracy (%)	Sensitivity (%)	Specificity (%)	FPR
TNN-IDS [14]	90.70	89.53	91.91	0.08	89.22	90.08	88.37	0.12
Dist-ADF [16]	91.37	90.12	92.65	0.07	91.08	90.38	91.79	0.08
ViT [17]	93.97	94.09	93.86	0.06	92.25	92.22	92.28	0.08
1D-CNN [23]	94.14	94.42	93.86	0.06	92.52	92.74	92.31	0.08
Ensemble learning [21]	94.25	94.66	93.86	0.06	92.70	92.85	92.56	0.07
TFKAN [30]	94.50	95.14	93.86	0.06	92.75	92.89	92.61	0.07
FedMSE [22]	94.53	95.21	93.86	0.06	92.86	92.90	92.82	0.07
IF-PCC [4]	94.57	95.29	93.87	0.06	93.60	93.65	93.54	0.06
LREC [18]	95.47	95.84	95.09	0.05	94.35	94.29	94.40	0.06
ChSp-ACN	97.22	97.20	97.24	0.03	96.22	96.16	96.28	0.04

4.7. Convergence analysis

The convergence graph representation of the proposed ChSp-ACN is displayed in Fig. 8.

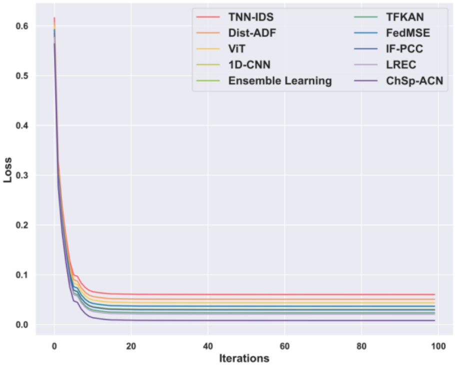
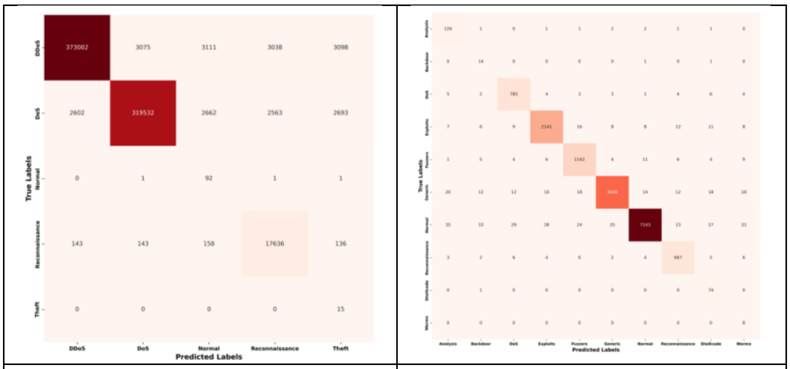


Fig. 8. Convergence graph representation of the proposed ChSp-ACN

From the analysis, it is evident that the proposed ChSp-ACN determines the intrusion in the IoT with a reduced loss of about 0.013 for 40 epochs, whereas the loss attained by the conventional TNS-IDS is 0.049, which is higher than the proposed model. As the epoch size varies, the loss attained by the proposed method also varies, but it is not higher than the existing models. At the 98th epoch, the proposed model attained a very low loss of 0.008 and outperforms other existing models. Hence, it is concluded that the utilization of the ChSp attention and the Adam optimizer in the proposed ChSp-ACN model enhanced the intrusion detection performance by obtaining optimal convergence.

4.8. Confusion matrix

Fig. 9 shows the confusion matrix for both BoT-IoT and UNSW-NB15 datasets utilized in the proposed ChSp-ACN. The information regarding the true label and the predictive labels is obtained using this confusion matrix. Hence, it is obvious that the proposed ChSp-ACN model accurately detected 723,385 as DDoS, 1,288,456 as DoS, 71,143 as Reconnaissance, 371 as normal, and 63 as Theft attacks using the BoT-IoT dataset. Subsequently, the proposed model correctly detects 14,730 as Generic, 28,789 as normal, 8653 as Exploits, 4720 as Fuzzers, 3170 as DoS, 2716 as Reconnaissance, 516 as Analysis, 295 as Shellcode, 64 as Backdoor, and 34 as Worms using the UNSW-NB15 dataset. However, the model predicts only a very few number of false positives, which will be negligible. Thus, it is revealed that the proposed approach stimulates effective detection due to the inclusion of the ChSp attention, Adam optimization, and thereby alleviates the vanishing gradient and overfitting issues.



(a)BoT-IoT dataset

(b)UNSW-NB15 dataset

Fig. 9. Confusion matrix

5. Conclusion

In this research, a new ChSp-ACN model is developed for detecting attacks in the network system. The malicious internal attacks affect the model performance and cause weak authentication problems, security threats, and eavesdropping. These limitations are effectively overcome by the ChSp-ACN model, which is specifically designed by the implementation of the ChSp Attention mechanism with a neural network and Adam optimizer that effectively measures the adaptive rate of every parameter in the model. The executed Adam optimizer possessed high generalizability to attain a better convergence rate with the reduction of overfitting issues. Initially, the model organizes the data attributes and further detects the network behavior as normal or an intrusion. In this intrusion detection evaluation, the model attained better performance with the acquisition of various evaluation metrics as FPR, sensitivity, specificity, and accuracy. While evaluating the BoT-IoT dataset, the ChSp-ACN possessed superior performance with accuracy of 97.22%, sensitivity of 97.20%, specificity of 97.24% and FPR of 0.03. Whereas, the UNSW-NB15 dataset attained an accuracy of 96.22%, sensitivity of 96.16%, specificity of 96.28% and FPR of 0.04, respectively. Future directions will explore the use of advanced deep learning (DL) techniques and explainability for intrusion detection, integrating meta-heuristic algorithms to achieve more effective performance.

References

1. Centenaro, M., C. E. Costa, F. Granelli, C. Sacchi, L. Vangelista. A Survey on Technologies, Standards, and Open Challenges in Satellite IoT. – IEEE Communications Surveys & Tutorials, Vol. **23**, 2021, No 3, pp. 1693-1720. DOI: 10.1109/COMST.2021.3078433.
2. Aversano, L., M. L. Bernardi, M. Cimitile, R. Pecori. A Systematic Review of Deep Learning Approaches for IoT Security. – Computer Science Review, 2021, No 40, 100389. DOI: 10.1016/j.cosrev.2021.100389.
3. Altunay, H. C., Z. Albayrak. A Hybrid CNN+ LSTM-Based Intrusion Detection System for Industrial IoT Networks. – Engineering Science and Technology, an International Journal, Vol. **38**, 2023, 101322. DOI: 10.1016/j.jestch.2022.101322.
4. Mohy-Eddine, M., A. Guezzaz, S. Benkirane, M. Azrour, Y. Farhaoui. An Ensemble Learning-Based Intrusion Detection Model for Industrial IoT Security. – Big Data Mining and Analytics, Vol. **6**, 2023, No 3, pp. 273-287. DOI: 10.26599/BDMA.2022.9020032.
5. Woo, S., J. Park, J. Y. Lee, I. S. Kweon. CBAM: Convolutional Block Attention Module. – In: Proc. of Eur. Conf. Comput. Vision, 2018, pp. 3-19. DOI: 10.48550/arXiv.1807.06521.
6. Soliman, S., W. Oudah, A. Aljuhani. Deep Learning-Based Intrusion Detection Approach for Securing Industrial Internet of Things. – Alexandria Eng. J., Vol. **81**, 2023, pp. 371-383. DOI: 10.1016/j.aej.2023.09.023.
7. Hnamte, V., J. Hussain. DCNNBiLSTM: An Efficient Hybrid Deep Learning-Based Intrusion Detection System. – Telematics and Informatics Reports, Vol. **10**, 2023, 100053. DOI: 10.1016/j.teler.2023.100053.
8. Sahu, S., B. M. Mehtre. Network Intrusion Detection System Using J48 Decision Tree. – In: 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI'15), 2015, pp. 2023-2026. DOI:10.1109/ICACCI.2015.7275914.
9. Chang, Y., W. Li, Z. Yang. Network Intrusion Detection Based on Random Forest and Support Vector Machine. – In: 2017 IEEE International Conference on Computational Science and Engineering (CSE'17) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC'17), Vol. **1**, 2017, pp. 635-638. DOI: 10.1109/CSE-EUC.2017.118.
10. Koc, L., T. A. Mazzuchi, S. Sarkani. A Network Intrusion Detection System Based on a Hidden Naïve Bayes Multiclass Classifier. – Expert Systems with Applications, Vol. **39**, 2012, No 18, pp. 13492-13500. DOI: 10.1016/j.eswa.2012.07.009.
11. Şen, S. Y, N. Özkurt. October Convolutional Neural Network Hyperparameter Tuning with Adam Optimizer for ECG Classification. – In: Innovations Intell. Syst. Appl. Conf. (ASYU'20), IEEE, 2020, pp. 1-6. DOI: 10.1109/ASYU50717.2020.9259896.
12. Rosay, A., F. Carlier, P. Leroux. MLP4NIDS: An Efficient MLP-Based Network Intrusion Detection for the CICIDS2017 Dataset. – In: Machine Learning for Networking: 2nd IFIP TC 6 International Conference, MLN 2019, Paris, France, 3-5 December 2019, Revised Selected Papers, Vol. **2**, 2020, pp. 240-254. DOI: 10.1007/978-3-030-45778-5_16.
13. Yue, C., L. Wang, D. Wang, R. Duo, X. Nie. An Ensemble Intrusion Detection Method for Train Ethernet Consist Network Based on CNN and RNN. – IEEE Access, Vol. **9**, 2021, pp. 59527-59539. DOI: 10.1109/ACCESS.2021.3073413.
14. Ullah, F., S. Ullah, G. Srivastava, J. C. W. Lin. IDS-INT: Intrusion Detection System Using Transformer-Based Transfer Learning for Imbalanced Network Traffic. – Digital Commun. Networks, Vol. **10**, 2024, No 1, pp. 190-204. DOI: 10.1016/j.dcan.2023.03.008.

15. Ullah, S., J. Ahmad, M. A. Khan, M. S. Alshehri, W. Boulila, A. Koubaa, S. U. Jan, M. M. I. Ch. TNN-IDS: Transformer Neural Network-Based Intrusion Detection System for MQTT-Enabled IoT Networks. – Comput. Networks, Vol. **237**, 2023, 110072. DOI: 10.1016/j.comnet.2023.110072.

16. Jullian, O., B. Otero, E. Rodríguez, N. Gutiérrez, H. Antona, R. Canal. Deep-Learning-Based Detection for Cyber-Attacks in IoT Networks: A Distributed Attack Detection Framework. – J. Network Syst. Manage, Vol. **31**, 2023, No 2, 33. DOI: 10.1007/s10922-023-09722-7.

17. Sana, L., M. M. Nazir, J. Yang, L. Hussain, Y. L. Chen, C. S. Ku, M. Alatiyyah, L. Y. Por. Securing the IoT Cyber Environment: Enhancing Intrusion Anomaly Detection with Vision Transformers. – IEEE Access, 2024. DOI: 10.1109/ACCESS.2024.3404778.

18. Chalichalamala, S., N. Govindan, R. Kasarapu. Logistic Regression Ensemble Classifier for Intrusion Detection System in the Internet of Things. – Sensors, Vol. **23**, 2023, No 23, 9583. DOI: 10.3390/s23239583.

19. Alhenawi, E. A., H. Alazzam, R. Al-Sayyed, O. Abu-Alghanam, O. Adwan. Hybrid Feature Selection Method for Intrusion Detection Systems Based on an Improved Intelligent Water Drop Algorithm. – Cybernetics and Information Technologies, Vol. **22**, 2022, No 4, pp. 73-90.

20. Wang, S., W. Xu, Y. Liu. Res-TranBiLSTM: An Intelligent Approach for Intrusion Detection in the Internet of Things. – Comput. Networks, Vol. **235**, 2023, 109982. DOI: 10.1016/j.comnet.2023.109982.

21. Zhang, H. Development of an Intelligent Intrusion Detection System for IoT Networks Using Deep Learning. – Discover Internet of Things, Vol. **5**, 2025, No 1, 74. DOI: 10.1007/s43926-025-00177-7.

22. Beuran, R. FEDMSE: Semi-Supervised Federated Learning Approach for IoT Network Intrusion Detection. – Computers & Security, Vol. **151**, 2025, 104337. DOI: 10.1016/j.cose.2025.104337.

23. Hossain, M. A. Deep Learning-Based Intrusion Detection for IoT Networks: A Scalable and Efficient Approach. – Eurasip Journal on Information Security, Vol. **1**, 2025, 28. DOI: 10.1186/s13635-025-00202-w.

24. The BoT-IoT Dataset (Accessed on October 2024). **https://www.kaggle.com/datasets/vigneshvenkateswaran/bot-iot**

25. The UNSW-NB15 dataset (Accessed on October 2024). **https://research.unsw.edu.au/projects/unsw-nb15-dataset**

26. Cascone, L., K. Alnowaiser, A. Alarfaj, E. A. Alabdulqader, M. Umer, B. Alankar. IoT-Based Smart Framework to Predict Air Quality in Congested Traffic Areas Using Sv-Cnn Ensemble and Knn Imputation Model. – Available at SSRN, 4737721. DOI: 10.1016/j.compeleceng.2024.109311.

27. Fadlil, A. K. Nearest Neighbor Imputation Performance on Missing Value Data for Graduate User Satisfaction. – Jurnal Rekayasa Sistem Informasi dan Teknologi, Vol. **6**, 2022, No 4, pp. 570-576. DOI: 10.59407/jrsit.v1i2.77.

28. Ali, P. J. M., R. H. Faraj, E. Koya, P. J. M. Ali, R. H. Faraj. Data Normalization and Standardization: A Technical Report. – Mach. Learn. Tech. Rep., Vol. **1**, 2014, No 1, pp. 1-6. DOI: 10.13140/RG.2.2.28948.04489.

29. Hnamte, V., J. Hussain. Dependable Intrusion Detection System Using Deep Convolutional Neural Network: A Novel Framework and Performance Evaluation Approach. – Telematics Inf. Rep., Vol **11**, 2023, 100077. DOI: 10.1016/j.teler.2023.100077.

30. Fares, I., M. Abd Elaziz, A. Aseeri, H. Zied, A. Abdellatif. TFKAN: Transformer Based on Kolmogorov-Arnold Networks for Intrusion Detection in IoT Environment. – Egyptian Informatics Journal, Vol. **30**, 2025, 100666.

31. Zhou, S., X. Li. Spiking Neural Networks with Single-Spike Temporal-Coded Neurons for Network Intrusion Detection. – In: 25th International Conference on Pattern Recognition (ICPR'20), 2021, pp. 8148-8155.

Fast-track. Received: 13.10.2025, Revised version: 01.12.2025, Accepted: 07.12.2025