The object of this study is the process of recognizing intrusions in computer networks. Network intrusion detection systems (NIDS) have become an important area of research as they are used to protect computer systems from hacker attacks. Deep learning is becoming increasingly popular for detecting and classifying malicious network traffic, including for building NIDS.

In this paper, we propose a network intrusion detection model CNN-BiGRU-Attention based on a time-based approach to deep learning using the attention mechanism. The main goal of the study is to build an effective combined deep learning model that can detect various network cyber threats.

A 1D convolutional neural network is implemented to extract high-level representations of intrusion information features. A bidirectional gated recurrent unit (BiGRU) with an attention mechanism for traffic data classification has been designed. The attention mechanism plays a key role in the constructed model as it allows the system to focus only on important aspects of network traffic and allows the model to adapt to new types of threats.

The results of the study show that using a combination of CNN and BiGRU with the attention mechanism speeds up and improves the process of classifying network attacks. On the NSL-KDD and UNSW-NB15 training datasets, the model shows an accuracy of 99.81 % and 97.80 %. On the NSL-KDD and UNSW-NB15 test datasets, the model demonstrates 82.16 % and 97.72 % accuracy.

The proposed NIDS model will be considered for implementation in a real-time corporate network security system.

In general, the results of the study provide a new perspective on improving the performance of NIDS and are quite relevant in terms of using attention mechanisms to classify network traffic

Keywords: deep learning, combined model, network intrusion detection systems, attention mechanism

UDC 004
DOI: 10.15587/1729-4061.2024.305685

CONSTRUCTION OF A NETWORK INTRUSION DETECTION SYSTEM BASED ON A CONVOLUTIONAL NEURAL NETWORK AND A BIDIRECTIONAL GATED RECURRENT UNIT WITH ATTENTION MECHANISM

Andrii Nikitenko*
Corresponding author
E-mail: andrii.nikitenko@donntu.edu.ua
Yevhen Bashkov
Doctor of Technical Sciences, Professor*
*Department of Applied Mathematics and Informatics
Donetsk National Technical University
Potebni str., 56, Lutsk, Ukraine, 43003

Received date 05.03.2024
Accepted date 16.05.2024
Published date 03.06.2024

How to Cite: *Nikitenko, A., Bashkov, Y. (2024). Construction of a network intrusion detection system based on a convolutional neural network and a bidirectional gated recurrent unit with attention mechanism. Eastern-European Journal of Enterprise Technologies, 3 (9 (129)), 6–15. https://doi.org/10.15587/1729-4061.2024.305685*

1. Introduction

The world is experiencing a great revolution in the field of information technology; everyone is constantly exchanging information through the network. This involves creating new prevention and detection tools and mechanisms, as well as strengthening existing ones, such as NIDS, to strengthen security and protect the network from intrusions. The function of NIDS is to observe, evaluate, and classify the traffic passing through the network; it is based on pre-established methods and techniques to distinguish between normal and suspicious traffic. In addition, attackers are attracted to information and knowledge passing through the network, and in order to use it and make a profit, they are forced to overcome obstacles and security barriers, creating new attacks, and developing existing ones. Although modern NIDS are not evolutionary, their identification algorithms do not evolve to automatically detect new threats. These prompts thinking about advanced and intelligent detection methods that can identify new attacks and accompany the development of existing ones [1].

Deep Neural Networks (DNN) have seen great progress in many fields, including natural language processing, computer vision, and many others. With their incredible ability to detect complex dependences in data and automate the process of pattern recognition, DNNs are quite promising for improving NIDS.

The attention mechanism plays an equally important aspect in modern research. It allows the system to focus on specific aspects of network traffic that indicate possible attacks or unusual activity. Using the attention mechanism in NIDS makes it possible to detect even small changes or anomalies in traffic that may indicate hacking attacks, data leaks or other security threats.

In addition, with the emergence of new technologies such as the Internet of Things (IoT) or distributed computing, networks are becoming even more complex, making the process of detecting attacks more difficult. The use of the attention mechanism makes it possible to adapt the system to these changed conditions and increase its efficiency under the conditions of increasing complexity of the network environment.

This paper discusses key issues related to the use of DNNs in network intrusion detection, such as adapting to changing conditions, improving detection speed and accuracy. The implementation of deep learning neural networks in this context can expand the capabilities of network intrusion detection systems and make them more effective in the face of constantly changing cyber threats.

The results of scientific research can lead to the development of new, effective, and reliable intrusion detection systems that are able to detect and avert cyber threats in time. This can improve the overall security of networks and computer systems, which is important for corporate entities, government organi-

zations, and individual users. Also, the results demonstrate the effectiveness of using the attention mechanism in the classification of malicious network traffic. In addition, research results can stimulate the development of new deep learning technologies and methods. Thus, they can be applied not only in cyber security but also in other areas that require the analysis of large volumes of data and the detection of complex patterns.

Therefore, research into the detection of malicious traffic using network intrusion detection systems built on the basis of deep learning are quite relevant.

2. Literature review and problem statement

There are many studies that consider the issue of NIDS development; new solutions continue to appear every day. Work [2] reports the results of the analysis, which revealed that about 80 % of the proposed solutions are based on DL approaches, and AE and DNN architectures are most often used. Although DL schemes have much higher performance than ML-based methods, in terms of their self-learning ability of features and stronger model fitting ability. But these schemes are quite complex and require significant computing resources in terms of computing power and data storage capabilities. Paper [3] provides an analysis of popular data sets for creating NIDS. The KDD99 dataset was found to be the most popular, followed by the NSL-KDD and UNSW-NB15 datasets. But the problem with the KDD99 dataset is that it is a very old dataset, and it doesn't look like a modern traffic data stream. In [4], it is stated that DL is mostly used to study features in intrusion detection approaches. It is also suggested to review several areas using deep learning approaches instead of surface ML, as well as to conduct further comparative studies and explore hybrid architectures. It should be noted that the following neural network architectures are often used in the analyzed sources: DNN, recurrent neural networks (RNN), autoencoders (AE), convolutional neural networks (CNN). In addition, modern works propose combinations of the given architectures to improve the effectiveness of intrusion detection.

The first thing to note is the researchers' use of DNNs. Work [5] shows the results of using DNN to build a NIDS capable of detecting intrusions in real time. It is shown that the model demonstrated the following results: 99.48 %, 99.52 %, 99.34 %, and 99.42 % for accuracy, precision, actual observations predicted correctly (recall), and f1-measure (f1-score), respectively. Disadvantages are low results on the accuracy and recall metrics for the NSL-KDD test data set and the lack of comparison with other well-known architectures. It should also be noted that today it is impossible to be completely sure that only the use of DNN is able to detect intrusions as efficiently as possible. That is why work [6] presents a methodology based on one-dimensional CNN (1D-CNN) for detecting normal and four different types of malicious network traffic. It is shown that the model is simple and less computationally expensive. The following results were also achieved: accuracy, 98.96 %; recall, 99.5 %; precision, 99.2 %; and f1-score, 99.34, based on the CIC-IDS-2017 dataset. The disadvantages are the comparison with other works based only on the accuracy metric, which does not allow us to fully understand how effectively the model works. Also, the comparison was performed only on the training data set; this is not enough for a competent assessment of the model.

In addition to using CNN, Long-Short Term Memory (LSTM) or Gated Recurrent unit (GRU) architectures, it is possible to use combinations of architectures to improve the effectiveness of the intrusion detection model. In [7], the results of using a unified model combining a multi-scale convolutional neural network with long short-term memory (MSCNN-LSTM) are reported. It is shown that the use of the UNSW-NB15 data set gives the following results: accuracy, 95.6 %; false alarm rate, 9.8 %; false negative rate, 1.6 %. This work requires additional analysis, which will include a comparison with the results of similar studies.

Paper [8] shows the results of using a high-precision IDS model using a combined model of optimized CNN (OCNN) and hierarchical multi-scale LSTM (HMLSTM) for effective selection and learning of spatial-temporal features. It is shown that the testing was carried out on the NSL-KDD, ISCX-IDS, and UNSW-NB15 datasets. Results for the NSL-KDD dataset for accuracy – 90.67 %, precision –86.71 %, recall – 95.19 %, and f1-score – 91.46 %. For the UNSW-NB15 dataset, accuracy – 96.33 %, precision – 100 %, recall – 95.87 %, and f1-score – 98.13 %. But in addition to the precision assessment, it is still possible to improve the performance of the model on the accuracy assessments. An option to increase the estimates is to modify the model, that is, to experiment with a different number of layers or to add regularization.

Combinations not only with CNN but also with AE are being studied. Paper [9] shows the results of using a two-stage deep learning (TSDL) model based on a stack AE with a softmax classifier for effective detection of network intrusions. Experiments conducted by the authors of [9] show accuracy results of 99.996 % and 89.134 % for KDD99 and UNSW-NB15 datasets, respectively. But the paper lacks additional model evaluation metrics, and it is not clear on which data set, training or testing, the conclusions were drawn.

Work [10] reports the results of using the machine learning algorithm (1D CAE OCSVM), which demonstrate the ability of the model to generalize invisible attacks and confirms its competitiveness compared to the latest modern works. It is shown that the experiments on the NSL-KDD dataset demonstrated the following results: accuracy, 91.58 %; f1-score, 92.87 %; false negatives, 2.43 %; and recall, 97.11 %. For the UNSW-NB15 data set, accuracy, 94.28 %; f1-score, 95.06 %; false alarm rate, 5.51 %; and recall, 96.49 %. Similarly, 1D CNN in combination with bidirectional long short-term memory (Bidirectional LSTM – BiLSTM) (DCNNBiLSTM) is presented in [11]. The paper shows that the model is optimal for system administrators and network companies to improve intrusion detection. It is shown that the model was trained on the CIC-IDS2018 and Edge_IIoT datasets. The effectiveness of the model was investigated using multi-class classification and achieved 100 % and 99.64 % accuracy, respectively, when training and testing with data sets. However, although the improved DCNNBiLSTM shows good results when using new training and testing datasets, there are still some areas for improvement. The disadvantages of the work are the lack of evaluation metrics. Also in the work, when compared with modern studies, the author showed a comparison with models that were tested with other data sets, which does not allow us to make an adequate comparison by evaluation metrics.

In [12], a combination of CNN and BiLSTM is presented to take into account the spatial and temporal features of the data. It is shown that the publicly available datasets NSL-KDD and UNSW-NB15 are used for training and testing. Experimental results for the NSL-KDD dataset are as follows: accuracy, 99.22 %; false positive rate, 0.43 %; and

recall, 98.882 %. For the UNSW-NB15 data set: accuracy, 82.08 %; false positive rate, 6.092 %; and recall, 92.506 %. The disadvantage of the work is the insufficient number of studies for a comparative analysis of the effectiveness of the model.

Consequently, most reviewed papers focus on combinations of neural network architectures to improve network intrusion detection. But taking into account the analysis of studies [5, 8, 10], we can say that after building the model and conducting experiments, the use of NIDS models in real time was not demonstrated in further works. It also shows the lack of accuracy in detecting network threats that were not known until today. Also, in works [6, 7, 9, 11, 12] insufficient evaluation metrics were presented in order to fully assess the significance of the developed models and insufficient number of works for comparative analysis. It provokes thought about other ways to solve these problems and predetermines the further study of new solutions to this task.

3. The aim and objectives of the study

The purpose of our study is to determine the potential capabilities of a combined model based on deep learning and an attention mechanism in the classification of malicious network traffic. This will make it possible to effectively and quickly detect network intrusions and further consider a model for real-time intrusion detection.

To achieve this goal, the following tasks were set:

– to propose a combined model for creating NIDS, which consists of CNN, BiGRU, and attention mechanism;

– to investigate existing models implemented in the TensorFlow framework, train the models on NSL-KDD and UNSW-NB15 data sets and compare them with the developed model;

– to compare the performance of the constructed model with the results of modern research.

4. The study materials and methods

4. 1. The object and hypothesis of the study

The object of our research is the process of recognizing intrusions in computer networks. The main hypothesis of the study assumes that the combined CNN-BiGRU model combined with the attention mechanism can accurately, efficiently, and quickly categorize network intrusions. When comparing the model with the results of modern research, only 4 evaluation metrics are taken into account. This is because some studies used the necessary evaluation metrics, while others did not.

4. 2. Evaluation metrics

The proposed CNN-BiGRU-Attention model is evaluated using accuracy, precision, recall, and f1-score metrics [13]. Accuracy is the proportion of correct recognitions in the total amount of data. Precision is a measure of the accuracy of the positive predictions made by the model. It is calculated as the ratio of the number of correctly predicted positive observations to the total number of predicted positive results. Recall is a measure of the model's ability to capture all positive outcomes. It calculates the ratio of correctly predicted positive observations to the actual positive observations. F1-score gives the harmonic mean of precision and recall, which is a balanced measure that accounts for both false positives and false negatives.

4. 3. Experimental environment

As part of the research, Google Colab Pro cloud environment was used for experiments. This environment provides access to powerful computing resources, making it ideal for resource-intensive tasks.

Python version 3.10.12, TensorFlow version 2.15.0, and Keras version 2.15.0 were used for all experiments. The Optuna library was chosen to optimize hyperparameters. This open-source library, written in Python, uses Bayesian optimization algorithms to find the best sets of hyperparameters for machine learning.

5. Results of the CNN-BiGRU-Attention model performance in the classification of malicious network traffic

5. 1. Designing a combined CNN-BiGRU-Attention model for creating NIDS

To overcome the current problems of current NIDS methods, such as feature selection and finding a balance between model performance and computational cost, a hybrid intrusion detection model is proposed. It is based on a time-dependent deep learning approach using the attention mechanism. First, a 1D convolutional neural network is built to extract high-level features. This ensures preferential preservation of information about the original data and speeds up the subsequent classification process. Second, a two-layer BiGRU with a soft attention mechanism is added as a classifier, as shown in Fig. 1.

The application of 1D CNN is reflected in the model's ability to detect local patterns in sequential data and reduce their dimensionality. By using convolutional and pooling layers, the number of parameters and computational complexity of the model is reduced. The use of BiGRU is a key element of the framework, as it allows efficient modeling of dependences in time series data. BiGRU consists of two recurrent gradient modules that process information in both forward and reverse directions. This configuration allows the model to take into account both past and future context when processing the current item in the time sequence. Adding an attention mechanism to the model structure helps focus on the most important or relevant parts of the input data, reducing the amount of computation required to classify intrusions. After the weight mechanism, a fully connected layer is used to combine data features, and a dropout layer reduces overtraining and improves the generalization ability of the model.

That is why it is assumed that the presented complex model can provide effective detection of intrusions under conditions of significant computational load.

Convolutional neural network. CNN uses several layers of arrays for data processing [11]. The task of the convolutional layer is to extract features while preserving consistent information. CNN exploits the spatial correlations present in the input data. Each subsequent layer of the neural network consists of connections of input neurons. This particular area is known as the local receptive field [14]. The local receptive field is concentrated on invisible neurons. Hidden neurons analyze the input data inside the specified field, without knowing about the changes that occur outside it. 1D CNNs are capable of automatically extracting important features from network data such as data packets, event logs, etc. This enables the detection of hidden patterns and anomalies that may indicate network attacks. Figure 1 shows a model that uses a 1D CNN.

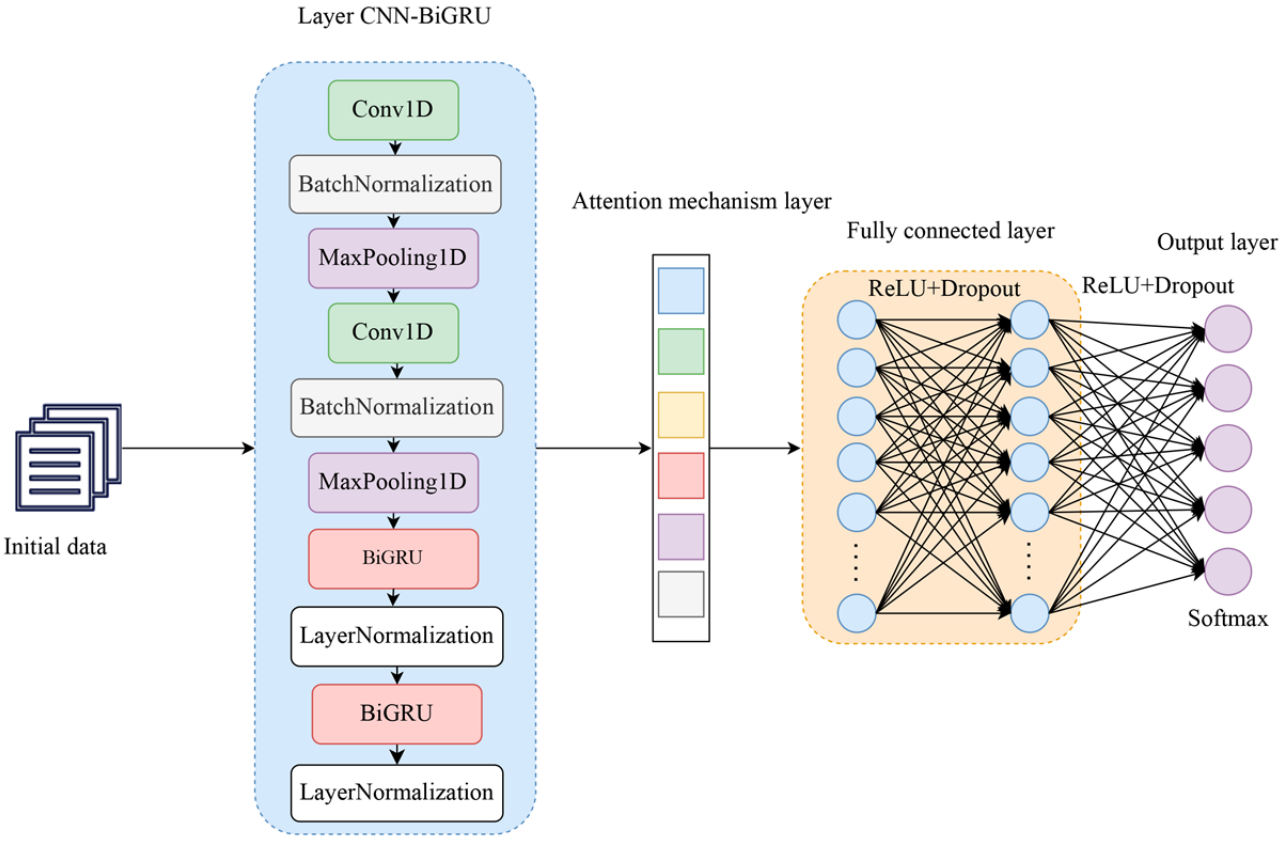

Fig. 1. CNN-BiGRU-Attention model for network intrusion detection

Bidirectional gated recurrent unit. RNNs are prone to gradient fading and gradient explosion, so a neural network like LSTM [15] or GRU [15] is usually chosen to prevent this. GRU and LSTM belong to the network architecture of RNN, and both use control mechanisms to control the flow of information. The difference is that LSTM has three nodes, including an input node, a forget node, and an output node, while GRU is a simplified version of LSTM and contains only a reset node and a refinement node. Compared with LSTM, GRU has a simpler model structure with fewer parameters, as shown in experimental results [16]; the functional diagram of GRU is shown in Fig. 2.

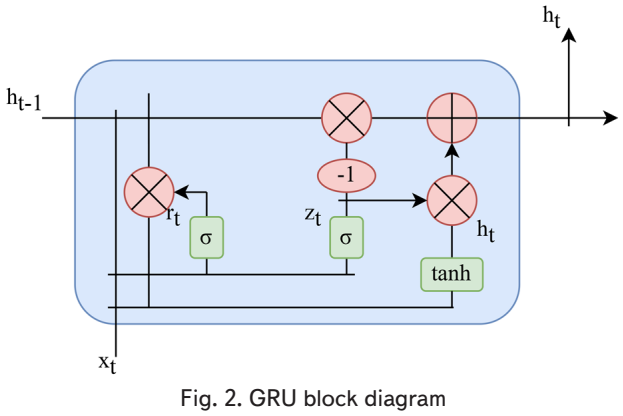

Fig. 2. GRU block diagram

It should be noted that the current input is x_t, and the new state h_t after time t consists of two parts: the candidate state $\tilde{h}_t$ and the past state h_{t-1}:

$$h_t = (1 - z_t) h_{t-1} + z_t \tilde{h}_t. \quad (1)$$

The reset node r_t is working in the process of obtaining the candidate state. The method of obtaining a candidate state is similar to a traditional RNN, except for the mechanism of nodes:

$$\tilde{h}_t = \tanh\left(x_t W_{xh} + W_{hh}\left(r_i \circ h_{t-1}\right) + b_h\right), \quad (2)$$

where ∘ is the Hadamard product,

W is the weight matrix,

b – displacement.

Here, r_t helps control how much information from the past state can be added to the candidate state. The reset node r_t is updated as follows:

$$r_t = \sigma\left(x_t W_{xr} + h_{t-1} W_{hr} + b_r\right). \quad (3)$$

According to equation (3), to obtain the new state h_t, the refinement node z_t plays the role of balancing the previous state h_{t-1} and the current candidate state $\tilde{h}_t$. z_t can be regarded as a node for sharing past information and new information. Updating z_t is similar to updating r_t:

$$z_t = \sigma\left(x_t W_{xz} + h_{t-1} W_{hz} + b_z\right). \quad (4)$$

BiGRU consists of forward and reverse GRU, so it can not only use past information in forward order but also use future information in reverse order [17]:

$$\vec{h}_t = \overrightarrow{GRU}\left(x_t\right), t \in [1, T],$$

$$\overleftarrow{h_t} = \overleftarrow{GRU}(x_t), t \in [1,T]. \quad (5)$$

Experiments [13] showed that BiGRU and BiLSTM work more efficiently than GRU and LSTM in creating network intrusion detection systems, so BiGRU was chosen for further research.

Mechanism of soft attention. Traffic detection environments are typically deployed on firewalls. The hardware platform on which the firewall is hosted is usually limited in computing and storage resources. More than the nominal traffic capacity causes the firewall to become a bottleneck of the network transmission channel, which does not facilitate network transmission. Especially in the case of limited computing resources, more traffic must pass through the firewall in real time. Therefore, the traffic detector must use reasonable computer resources. And the attention mechanism can accurately solve a complex problem, the attention mechanism is a resource allocation scheme, which is the main means of solving the problem of information overload [18]. The rational and efficient use of computing resources allows the detection model to focus on the recognition of malicious traffic feature maps. The mechanism of attention is divided into soft attention, hard attention, and self-attention [19]. The mechanism of soft attention is used in this work. First of all, the model has an attention weight matrix that can be trained, after the function is activated, the value is passed to the softmax function to obtain the weight value, and the K-dimensional weight vector of the sum of the values is 1. Finally, the attention vector can be obtained by weighted computation of the hidden state:

$$u_t = \tanh(W_w h_t + b_w), \quad (6)$$

$$a_t = \frac{\exp(u_t^T u_w)}{\sum_t \exp(u_t^T u_w)}, \quad (7)$$

$$V = \sum_t a_t h_t, \quad (8)$$

where h_t is a hidden state, W_w is a matrix of attention weights, b_w is an attention bias, a_t is a matrix of weight coefficients, and V is an attention vector weighted by the attention mechanism.

5. 2. Training of existing models on data sets and their comparison

The effectiveness of NIDS depends largely on the quality and quantity of data used to train and evaluate the system. With the development of ML and neural networks, datasets have become a key component in building accurate and reliable intrusion detection systems. The selected data set is used not only for training the NIDS model but also for evaluating the effectiveness of the proposed NIDS model [20]. Because of the difficulty of directly collecting real-time attack data, researchers can use publicly available standard datasets based on network traffic. NSL-KDD and UNSW-NB15 are selected for further research. These two datasets are widely used in current models, making experimental results better comparable to state-of-the-art models.

Datasets. The UNSW-NB15 data set [21] was built in the UNSW Cyber Polygon Laboratory to create a hybrid of real modern routine activity and synthetic modern behavior. Each UNSW-NB15 record has 49 features, which are divided into flow features, basic features, content features, time features, additionally generated and labeled signs, as given in Table 1. The main categories of records are normal records and attack records; attack records are divided into nine families according to the nature of the attack. According to [22], the data set has the following types of attacks: normal, fuzzers, analysis, backdoors, DoS, exploits, generic, reconnaissance, shellcode, worms.

Table 1

Features of the dataset UNSW-NB15

No.	Feature ID
1	scrip
2	sport
3	dstip
4	dsport
5	proto
6	state
7	dur
8	sbytes
9	dbytes
10	sttl
11	dttl
12	sloss
13	dloss
14	service
15	sload
16	dload
17	spkts
18	dpkts
19	swin
20	dwin
21	stcpb
22	dtcpb
23	smeansz
24	dmeansz
25	trans-depth
26	res_bdy_len
27	sjit
28	djit
29	stime
30	ltime
31	sintpkt
32	dintpkt
33	tcprtt
34	synack
35	ackdat
36	is_sm_ipsports
37	ct_state_ttl
38	ct_flw_http_mthd
39	is_ftp_login
40	ct_ftp_cmd
41	ct_stv_src
42	ct_stv_dst
43	ct_dst_ltm
44	ct_src_ltm
45	ct_src_dport_ltm
46	ct_dst_sport_ltm
47	ct_dst_src_ltm
48	attack_cat
49	label

The NSL-KDD dataset [23] is widely used in intrusion detection experiments. Most researchers use the NSL-KDD dataset as a reference dataset. Dataset categories are made more balanced by intelligently customizing them. Each record has 41 features that reveal different features of the stream, as well as a label assigned to each of them as either an attack type or a normal one. Features are divided into basic features, content features, time features, and host based. It should be noted that the 42nd attribute contains data on different 5 classes of network connection vectors, which are divided into one normal class and four attack classes. These 4 classes of attacks are divided into DoS, Probe, R2L and U2R [22]. Table 2 gives all the features of the data set.

Table 2

Features of the dataset NSL-KDD

No.	Feature ID
1	duration
2	protocol_type
3	Service
4	Flag
5	Src_bytes
6	Dst_bytes
7	Land
8	Wrong_fragment
9	Urgent
10	Hot
11	Num_failed_logins
12	Logged_in
13	Num_compromised
14	Root_shell
15	Su_attempted
16	Num_root
17	Num_file_creations
18	Num_shells
19	Num_access_files
20	Num_outbound_cmds
21	Is_host_login
22	Is_guest_login
23	Count
24	Srv_count
25	Serror_rate
26	Srv_serror_rate
27	Rerror_rate
28	Srv_rerror_rate
29	Save_srv_rate
30	Diff_srv_rate
31	Srv_diff_host_rate
32	Dst_host_count
33	Dst_host_srv_count
34	Dst_host_same_srv_rate
35	Dst_host_diff_srv_rate
36	Dst_host_same_src_port_rate
37	Dst_host_srv_diff_host_rate
38	Dst_host_serror_rate
39	Dst_host_srv_serror_rate
40	Dst_host_rerror_rate
41	Dst_host_srv_rerror_rate

The Optuna library was used to optimize the hyperparameters of the developed model. Other main parameters of the model are given below: dropout 0.5, number of epochs 50, batch size 192; Adam is chosen as the optimizer for parameter training. The model uses the loss function categorical_crossentropy for multiclass classification and binary_crossentropy for binary classification. Minimizing loss functions allows the model to predict results that are closer to real labels.

Comparison of existing models with new NIDS models becomes necessary to evaluate their performance and determine the optimal solution for specific application conditions. The following models are implemented using the TensorFlow framework: DNN, AE-DNN, BiLSTM, LSTM, GRU, and CNN-BiLSTM. The same basic parameters used for training the CNN-BiGRU-Attention model were applied to train the models. Below, Tables 3–6 give the results of multiclass classification.

Table 3

Accuracy of models on the NSL-KDD data set (multiclass classification)

Model ID	Training	Validation	Testing
DNN	99.70	99.71	77.92
AE-DNN	99.50	99.51	76.22
LSTM	91.53	91.71	71.57
GRU	98.70	98.68	74.56
BiLSTM	97.26	97.27	74.03
CNN-BiLSTM	99.65	99.66	72.60
CNN-BiGRU-Attention	99.81	99.81	82.16

Table 4

Comparison of models by precision, recall, and f1-score metrics based on the NSL-KDD data set (multi-class classification)

Model ID	Training			Testing		
	Precision	Recall	F1-score	Precision	Recall	F1-score
DNN	99.75	99.74	99.74	82.02	77.92	79.91
AE-DNN	99.51	99.50	99.50	77.38	76.22	76.79
LSTM	91.83	91.52	91.67	77.14	71.57	74.25
GRU	98.61	98.69	98.65	72.40	74.56	73.46
BiLSTM	96.49	97.26	96.87	68.97	74.02	71.40
CNN-BiLSTM	99.65	99.65	99.65	77.49	72.60	74.96
CNN-BiGRU-Attention	99.81	99.81	99.81	83.75	82.16	82.95

Table 5

Accuracy of models on the UNSW-NB15 data set (multi-class classification)

Model ID	Training	Validation	Testing
DNN	97.71	97.69	97.67
AE-DNN	97.67	97.58	97.63
LSTM	97.08	97.01	97.06
GRU	97.75	97.67	97.72
BiLSTM	97.77	97.69	97.77
CNN-BiLSTM	97.72	97.65	97.70
CNN-BiGRU-Attention	97.80	97.81	97.72

Table 6

Comparison of models by precision, recall, and f1-score metrics based on the UNSW-NB15 data set (multi-class classification)

Model ID	Training			Testing		
	Precision	Recall	F1-score	Precision	Recall	F1-score
DNN	97.60	97.71	97.65	97.47	97.71	97.58
AE-DNN	97.67	97.45	97.55	97.25	97.63	97.43
LSTM	96.98	97.09	97,03	97.00	97.06	97.02
GRU	97.46	97,80	97.62	97.41	97.71	97.55
BiLSTM	97.53	97.77	97.64	97.52	97.74	97.62
CNN-BiLSTM	97.70	97.71	97.70	97.62	97.70	97.65
CNN-BiGRU-Attention	97.61	97.81	97.71	97.50	97.72	97,61

As a result, the CNN-BiGRU-Attention model demonstrated the best performance on the NSL-KDD dataset, while BiLSTM, LSTM, and GRU performed worse. But if you look at the UNSW-NB15 data set, according to the precision evaluation metric, the CNN-BiLSTM model has higher accuracy for multiclass classification and identical results for recall and f1-score for binary classification. Therefore, there is still room for improvement of the developed CNN-BiGRU-Attention model.

A comparison of the effectiveness of the proposed model with well-known models on the NSL-KDD and UNSW-NB15 data sets for binary classification is given in Tables 7–10.

Table 7

Accuracy of models on the NSL-KDD data set (binary classification)

Model ID	Training	Validation	Testing
DNN	99.80	99.74	81.53
AE-DNN	99.38	99.36	78.77
LSTM	98.71	98.71	79.48
GRU	98.39	98.52	78.44
BiLSTM	97.64	97.52	74.80
CNN-BiLSTM	99.63	99.61	76.20
CNN-BiGRU-Attention	99.83	99.84	85.26

Table 8

Comparison of models by precision, recall, and f1-score metrics based on the NSL-KDD dataset (binary classification)

Model ID	Training			Testing		
	Precision	Recall	F1-score	Precision	Recall	F1-score
DNN	99.79	99.79	99.79	85.56	81.52	83.49
AE-DNN	99.37	99.37	99.37	82.28	78.77	80.48
LSTM	98.71	98.71	98.71	82.68	79.47	81.04
GRU	98.38	98.38	98.38	81.86	78.44	80.11
BiLSTM	97.65	97.64	97.64	82.12	74.79	78.12
CNN-BiLSTM	99.63	99.63	99.63	81.10	76.20	78.57
CNN-BiGRU-Attention	99.83	99.83	99.83	86.10	85.30	85.7

Table 9

Accuracy of models on the UNSW-NB15 data set (binary classification)

Model ID	Training	Validation	Testing
DNN	99.16	99.17	99.16
AE+DNN	99.02	98.96	98.97
LSTM	98.72	98.71	98.71
GRU	99.00	99.01	98.98
BiLSTM	99.16	99.15	99.13
CNN-BiLSTM	99.17	99.15	99.15
CNN-BiGRU-Attention	99.18	99.16	99.15

Table 10

Comparison of models by precision, recall, and f1-score metrics on the UNSW-NB15 data set (binary classification)

Model ID	Training			Testing		
	Precision	Recall	F1-score	Precision	Recall	F1-score
DNN	99.16	99.16	99.16	99.13	99.16	99.14
AE-DNN	99.00	98.97	98.98	98.99	98.96	98.98
LSTM	98.83	98.72	98.77	98.18	98.71	98.44
GRU	99.00	99.00	99.00	98.99	98.89	98.93
BiLSTM	99.16	99.16	99.16	99.13	99.13	99.13
CNN-BiLSTM	99.18	99.17	99.17	99.15	99.14	99.14
CNN-BiGRU-Attention	99.17	99.17	99.17	99.15	99.15	99.15

5. 3. Comparison of the proposed CNN-BiGRU-Attention model with modern developments

Tables 11, 12 give data on modern developments, which have new solutions for detecting network intrusions, justified by the need to devise more effective and reliable methods of protection against cyber-attacks. These developments focus on various aspects of cyber security, including the detection of new threats, vulnerability analysis, the development of ML algorithms to detect anomalous network activity, and the improvement of security incident response techniques. Comparison with these works allows us to understand how well the proposed model meets modern requirements and whether it is possible that they will bring new ideas or improvements to existing approaches.

Table 11

Comparison of the proposed model with modern developments (NSL-KDD)

Model ID	Accuracy	Precision	Recall	F1-score
DNN [5]	99.48	99.52	99.34	99.42
1D CAE OCSVM [10]	91.58	N/A	97.11	N/A
CNN-BiLSTM [12]	99.22	N/A	98.88	N/A
SRFCNN-BiGRU [13]	99.81	99.76	99.81	99.79
S-ResNet [24]	99.53	N/A	99.53	99.54
CNN-GRU [25]	99.69	99.65	99.69	99.70
Proposed model	99.81	99.81	99.81	99.81

Table 12

Comparison of the proposed model with modern developments (UNSW-NB15)

Model ID	Accuracy	Precision	Recall	F1-score
MSCNN [7]	95.60	N/A	N/A	N/A
TSDL (stacked auto-encoder) [9]	89.134	N/A	N/A	N/A
1D CAE OCSVM [10]	94.28	N/A	96.49	95.06
CANET [26]	89.39	N/A	N/A	N/A
Ensemble SVM [27]	N/A	90.5	97.21	93.72
Proposed model	97.80	97.61	97.81	97.71

It should be noted that Tables 11, 12 give values "N/A", which shows the absence of data on these metrics in the work. It should also be noted that the comparison was based on the results of the models on the training data sets, which is due to the lack of information on the test data in some works.

6. Discussion of results from comparing the proposed CNN-BiGRU-Attention model with well-known models and modern developments

Tables 3, 4 demonstrate that the proposed model has the highest indicators according to all evaluation metrics. The LSTM model has the lowest performance of 91.53 % and 71.57 % for the NSL-KDD training and testing datasets, respectively. This may be due to the gradient vanishing or dropout problem, especially for long sequences, resulting in poor adaptation to complex patterns in network traffic. At the same time, according to evaluation metrics, the closest model to the proposed one is DNN, with an accuracy of 99.70 % and 77.92 % for training and testing data, which can detect rapidly changing abnormal patterns. However, the proposed model combined with the attention mechanism better focuses on the important parts of the input data, so it has accuracy rates of 99.81 % and 82.16 % for training and testing data.

But not everything is so clear if you look at Tables 5, 6. Based on Table 5, the proposed model performed better on the training and validation data, but the BiLSTM model has a better accuracy rate on the test data. In Table 6, part of the evaluation metrics is better in the proposed model, and the other part in the CNN-BiLSTM model, again the attention mechanism allowed us to focus on the important parts of the network traffic.

In Tables 7, 8, according to all evaluation metrics, the proposed model is better, the closest model in terms of results is DNN. In Table 9, according to all evaluation metrics, the proposed model is almost better, the f1-score indicator for the proposed model and CNN-BiLSTM turned out to be identical. In Table 10, as in Table 6, part of the evaluation metrics is better in the proposed model, and the other is better in CNN-BiLSTM, therefore, in future studies, attention should be paid not only to the combination of CNN-BiGRU but also to CNN-BiLSTM, considering the good performance on the UNSW-NB15 dataset.

Considering the results in Tables 11, 12, the proposed model has good performance in contrast to other models and current developments of NIDS based on deep learning. The high performance of the model became possible owing to the addition of an attention mechanism that focused on different parts of the input data, giving priority weight to important aspects, unlike the developments in Tables 11, 12. Also, hyperparameter optimization helped tune parameters such as learning rate, network size, etc., to achieve optimal performance.

The results of the proposed model for the NSL-KDD data set for all evaluation metrics are 99.81 %. For the UNSW-NB15 data set, the evaluation metrics are as follows: accuracy, 97.80 %; precision, 97.61 %; recall, 97.81 %; f1-score, 97.71 %. The results reflect a significant improvement in the efficiency and reliability of the developed NIDS, in particular, the proposed solution made it possible to increase the accuracy of detecting anomalous activity and improve the system's ability to respond to intrusions. Unlike [5], in which only DNN was used to build the model, the results of the proposed model show an improvement in the detection of network intrusions according to all evaluation metrics. This is made possible by considering not only DNNs but also other architectures and their combinations. Unlike [10], in which a combination of 1D convolutional AE architecture and SVM was used, the results of the proposed model are better in accuracy and recall metrics for the NSL-KDD and UNSW-NB15 datasets. This becomes possible owing to the combined architecture of 1D CNN and BiGRU, and it gives an understanding that the use of machine learning algorithms is not enough for effective detection of intrusions. In contrast to [25], in which a combination of CNN and BiGRU architectures were used for construction, the results of the proposed model on the NSL-KDD dataset show better results on all evaluation metrics. This is made possible not only by the combination of architectures but also by the addition of an attention mechanism. In contrast to [7], in which multi-scale CNN and LSTM were used, the results of the proposed model on the UNSW-NB15 data set show better results on the accuracy metric. This is made possible by using a 1D CNN rather than a multi-scale CNN to extract spatial features of the data. From the point of view of computational resources, multiscale CNN has higher computational complexity compared to 1D CNN. In contrast to [26], in which a hierarchical CNN-Attention network was used, our model improves the best results according to the accuracy metric. This is made possible by adding the BiGRU architecture to automatically determine the importance of features during training, taking into account the dependences between features and selecting the most informative ones for classification.

Although the current research has produced satisfactory results, it is still not perfect. A limitation in trying to apply the model in practice is problems with the interpretation of the results, that is, the model's solutions can be difficult for a person to understand and reproduce. The disadvantage of the study is the lack of use of other types of the attention mechanism and the lack of demonstration of the data pre-processing process. Therefore, additional evaluation metrics should be paid attention to in future studies, such as time spent on classification, time spent on model training. This is due to the fact that it is necessary to take into account the limited resources of critical enterprises to provide for the security of corporate networks.

Therefore, in future studies we plan to investigate in detail the impact of various attentional mechanisms on the performance of network intrusion classification using NIDS and add additional evaluation metrics to ensure completeness of studies.

7. Conclusions

1. A combination of two CNN and BiGRU architectures with an attention mechanism into a combined CNN-BiGRU-Attention model has been proposed. This combination of CNN-BiGRU architectures allows us to potentially improve feature extraction and classification of cyber threats in network attacks. On the other hand, the attention mechanism is able to promote the focus on malicious flows in segments of the data flow, which is an important aspect to ensure efficiency under conditions of limited computing resources.

2. Existing models were studied: DNN, AE-DNN, LSTM, GRU, BiLSTM, CNN-BiLSTM, which were trained and compared with each other, including the developed CNN-BiGRU-Attention model. As a result of the comparison, it was found that the developed model demonstrated the best indicators. Training and testing accuracy for NSL-KDD data set: 99.81 %, 82.16 %, for UNSW-NB15 data set: 97.80 %, 97.72 %, respectively.

3. CNN-BiGRU-Attention model was compared to modern studies, based on which it was found that the model has the highest performance for the UNSW-NB15 data set, where accuracy, precision, recall, f1-score were 97.80 %, 97.61 %, 97.81 %, 97.71 %. At the same time, for the NSL-KDD data set, the proposed model has higher indicators only for precision and f1-score. Compared to the SRFCNN-BiGRU model [13], the accuracy of 99.81 % and recall of 99.81 % are identical. Therefore, it becomes clear that there is room for further development and improvement.

Conflicts of interest

The authors declare that they have no conflicts of interest in relation to the current study, including financial, personal, authorship, or any other, that could affect the study and the results reported in this paper.

Funding

This work was supported by the European Commission [grant number ENI/2019/413-664 «EDUTIP»].

Data availability

The data will be provided upon reasonable request.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

References

1. Boukhalfa, A., Abdellaoui, A., Hmina, N., Chaoui, H. (2020). LSTM deep learning method for network intrusion detection system. International Journal of Electrical and Computer Engineering (IJECE), 10 (3), 3315. https://doi.org/10.11591/ijece.v10i3.pp3315-3322
2. Ahmad, Z., Shahid Khan, A., Wai Shiang, C., Abdullah, J., Ahmad, F. (2020). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. Transactions on Emerging Telecommunications Technologies, 32 (1). https://doi.org/10.1002/ett.4150
3. Kumar, S., Gupta, S., Arora, S. (2021). Research Trends in Network-Based Intrusion Detection Systems: A Review. IEEE Access, 9, 157761–157779. https://doi.org/10.1109/access.2021.3129775
4. Aldweesh, A., Derhab, A., Emam, A. Z. (2020). Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. Knowledge-Based Systems, 189, 105124. https://doi.org/10.1016/j.knosys.2019.105124
5. Thirimanne, S. P., Jayawardana, L., Yasakethu, L., Liyanaarachchi, P., Hewage, C. (2022). Deep Neural Network Based Real-Time Intrusion Detection System. SN Computer Science, 3 (2). https://doi.org/10.1007/s42979-022-01031-1
6. Qazi, E. U. H., Almorjan, A., Zia, T. (2022). A One-Dimensional Convolutional Neural Network (1D-CNN) Based Deep Learning System for Network Intrusion Detection. Applied Sciences, 12 (16), 7986. https://doi.org/10.3390/app12167986
7. Zhang, J., Ling, Y., Fu, X., Yang, X., Xiong, G., Zhang, R. (2020). Model of the intrusion detection system based on the integration of spatial-temporal features. Computers & Security, 89, 101681. https://doi.org/10.1016/j.cose.2019.101681
8. Rajesh Kanna, P., Santhi, P. (2021). Unified Deep Learning approach for Efficient Intrusion Detection System using Integrated Spatial–Temporal Features. Knowledge-Based Systems, 226, 107132. https://doi.org/10.1016/j.knosys.2021.107132
9. Khan, F. A., Gumaei, A., Derhab, A., Hussain, A. (2019). TSDL: A Two-Stage Deep Learning Model for Efficient Network Intrusion Detection. IEEE Access, 7, 30373–30385. https://doi.org/10.1109/access.2019.2899721
10. Binbusayyis, A., Vaiyapuri, T. (2021). Unsupervised deep learning approach for network intrusion detection combining convolutional autoencoder and one-class SVM. Applied Intelligence, 51 (10), 7094–7108. https://doi.org/10.1007/s10489-021-02205-9
11. Hnamte, V., Hussain, J. (2023). DCNNBiLSTM: An Efficient Hybrid Deep Learning-Based Intrusion Detection System. Telematics and Informatics Reports, 10, 100053. https://doi.org/10.1016/j.teler.2023.100053
12. Sinha, J., Manollas, M. (2020). Efficient Deep CNN-BiLSTM Model for Network Intrusion Detection. Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Pattern Recognition. https://doi.org/10.1145/3430199.3430224
13. Cao, B., Li, C., Song, Y., Fan, X. (2022). Network Intrusion Detection Technology Based on Convolutional Neural Network and BiGRU. Computational Intelligence and Neuroscience, 2022, 1–20. https://doi.org/10.1155/2022/1942847
14. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B. et al. (2018). Recent advances in convolutional neural networks. Pattern Recognition, 77, 354–377. https://doi.org/10.1016/j.patcog.2017.10.013
15. Song, Y., Luktarhan, N., Shi, Z., Wu, H. (2023). TGA: A Novel Network Intrusion Detection Method Based on TCN, BiGRU and Attention Mechanism. Electronics, 12 (13), 2849. https://doi.org/10.3390/electronics12132849

16. Li, X. (2023). CNN-GRU model based on attention mechanism for large-scale energy storage optimization in smart grid. Frontiers in Energy Research, 11. https://doi.org/10.3389/fenrg.2023.1228256
17. Yu, X., Li, T., Hu, A. (2020). Time-series Network Anomaly Detection Based on Behaviour Characteristics. 2020 IEEE 6th International Conference on Computer and Communications (ICCC). https://doi.org/10.1109/iccc51575.2020.9345249
18. Liu, X., Liu, J. (2021). Malicious traffic detection combined deep neural network with hierarchical attention mechanism. Scientific Reports, 11 (1). https://doi.org/10.1038/s41598-021-91805-z
19. Huang, Y., Chen, J., Zheng, S., Xue, Y., Hu, X. (2021). Hierarchical multi-attention networks for document classification. International Journal of Machine Learning and Cybernetics, 12 (6), 1639–1647. https://doi.org/10.1007/s13042-020-01260-x x
20. Nikitenko, A. (2023). Datasets for creating intrusion detection systems using neural networks. Proceedings of the XLI Scientific and Technical Conference of Young Scientists and Specialists of the Pukhov Institute for Modeling Problems in Energy of the National Academy of Sciences of Ukraine, 104–106. Available at: https://ipme.kiev.ua/wp-content/uploads/2023/05/%D0%9C%D0%B0%D1%82%D0%B5%D1%80%D1%96%D0%B0%D0%BB%D0%B8-%D0%BA%D0%BE%D0%BD%D1%84%D0%B5%D1%80%D0%B5%D0%BD%D1%86%D1%96%D1%97-2023.pdf
21. The UNSW-NB15 Dataset. Available at: https://research.unsw.edu.au/projects/unsw-nb15-dataset
22. Nikitenko, A. (2023). Network intrusion detection systems based on deep learning neural networks. Scientific papers of DonNTU. Series: "Informatics, Cybernetics and Computer Science", 2 (37), 15–21.
23. ISCX NSL-KDD dataset 2009. Available at: https://www.unb.ca/cic/datasets/nsl.html
24. Xiao, Y., Xiao, X. (2019). An Intrusion Detection System Based on a Simplified Residual Network. Information, 10 (11), 356. https://doi.org/10.3390/info10110356
25. Cao, B., Li, C., Song, Y., Qin, Y., Chen, C. (2022). Network Intrusion Detection Model Based on CNN and GRU. Applied Sciences, 12 (9), 4184. https://doi.org/10.3390/app12094184
26. Ren, K., Yuan, S., Zhang, C., Shi, Y., Huang, Z. (2023). CANET: A hierarchical CNN-Attention model for Network Intrusion Detection. Computer Communications, 205, 170–181. https://doi.org/10.1016/j.comcom.2023.04.018
27. Marir, N., Wang, H., Feng, G., Li, B., Jia, M. (2018). Distributed Abnormal Behavior Detection Approach Based on Deep Belief Network and Ensemble SVM Using Spark. IEEE Access, 6, 59657–59671. https://doi.org/10.1109/access.2018.2875045