International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

ATTENTION-AUGMENTED DEEP CONVOLUTIONAL FRAMEWORK FOR INTELLIGENT NETWORK TRAFFIC ANOMALY DETECTION

[1] Dr.N.Bhanupriya,[2] M Sowjanya
[1]Assistant Professor, [2]MCA Student
Department Of MCA
Sree Chaitanya College Of Engineering, Karimnagar

ABSTRACT
The exponential growth of internet-connected devices has led to increasingly complex and high-volume network traffic, making anomaly detection a crucial challenge in cybersecurity. This research proposes an attention-augmented deep convolutional framework for intelligent network traffic anomaly detection. The model integrates attention mechanisms with big-step convolutional neural networks (CNNs) to enhance feature extraction and focus on critical traffic attributes. By leveraging datasets such as CICIDS2017, the system learns to differentiate between normal and malicious patterns efficiently. Experimental results show that the proposed model significantly improves detection accuracy, precision, and recall compared to traditional machine learning and basic CNN-based approaches. This framework demonstrates scalability, robustness, and adaptability, offering a promising solution for next-generation intrusion detection systems in modern digital networks.

Received: 23-09-2025 Accepted: 28-10-2025 Published: 04-11-2025

I. INTRODUCTION

With the rapid expansion of digital communication and the proliferation of IoT and cloud-based systems, network infrastructures face growing exposure to cyber threats. Detecting abnormal or malicious network traffic has become a vital component of cybersecurity strategies. Conventional rule-based intrusion detection systems (IDS) are limited in handling the dynamic and evolving nature of network attacks. Therefore, intelligent anomaly detection frameworks that leverage deep learning are gaining increasing attention for their ability to automatically learn and generalize complex traffic behaviors.

Deep learning models, particularly Convolutional Neural Networks (CNNs), have demonstrated high effectiveness in analyzing structured network data and identifying abnormal patterns. However, traditional CNNs may struggle with capturing long-range dependencies and important contextual relationships across features. To address this limitation, attention mechanisms have been introduced to prioritize critical information and improve model interpretability. These mechanisms enable the model to focus on the most relevant aspects of traffic data, thereby enhancing detection performance and reducing false positives.

This research integrates attention modules with a big-step convolutional architecture to create a powerful anomaly detection model. The proposed framework processes high-dimensional network data efficiently and dynamically adapts to different attack types. The combination of attention layers and large convolutional kernels enhances spatial representation, making it particularly suitable for complex traffic environments. This approach aims to deliver a scalable, intelligent system capable of real-time detection in modern network infrastructures.

II. LITERATURE SURVEY

Kim et al. (2016) applied CNNs to intrusion detection and demonstrated their superior performance over traditional machine learning classifiers. Shone et al. (2018) introduced deep autoencoders for unsupervised anomaly

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

detection, highlighting the importance of hierarchical feature extraction in cybersecurity. Yin et al. (2019) developed an RNN-based model for network traffic analysis, capable of detecting temporal correlations in attack patterns.

Zhang et al. (2020) explored hybrid deep learning frameworks that combined CNN and LSTM layers, improving the model's ability to handle sequential traffic data. Liu and Wang (2021) incorporated attention mechanisms into CNN models for anomaly detection, achieving better feature prioritization and interpretability. More recently, Chen et al. (2023) presented an attention-guided convolutional framework using the CICIDS2017 dataset, reporting substantial gains in accuracy and recall compared to baseline CNNs.

These studies collectively highlight the evolution of network intrusion detection from traditional rule-based systems to deep learning-driven solutions. Building on these advancements, this research proposes an attention-augmented big-step convolutional model that improves both efficiency and precision in identifying anomalous network behaviors.

III. SYSTEM ANALYSIS & DESIGN

EXISTING SYSTEM

Shi et al. [16] proposed a cost-sensitive SVM (CMSVM) for the network traffic imbalance problem. The model uses a multi-class SVM with an active learning algorithm to solve the imbalance problem for different applications by adaptive weights. Cao et al. [17] proposed a real-time network classification model with SPPSVM. The model uses the feature selection method of principal component analysis (PCA) to reduce the dimensionality of the original data and uses an improved particle swarm optimization algorithm to obtain the optimal parameters. The classification accuracy is higher compared to the traditional SVM model. Farid et al. [18] combined naive bayes and decision trees for anomalous traffic detection while eliminating redundant attributes of the traffic data. The proposed algorithm improves the detection rate. Machine learning based classification methods usually require manual feature design and selection, which cannot cope with the evolution of networks nowadays.

Gianni et al. [19] proposed a novel deep neural network based on auto encoder. The model embeds multiple auto encoders into convolutional and recurrent neural networks to elicit the basic features of interest, which uses stacked fully connected neural networks to achieve classification of network traffic.

Ren et al. [20] proposed a tree-structured recurrent neural network that uses a tree structure to divide large classification into small classification problems. The model can automatically learn the nonlinear relationship between the input data and the output data, which has a better classification effect. Tal et al. [21] proposed a new method for encrypted traffic classification. The method first converts traffic data into intuitive images, and then combines convolutional neural networks to achieve classification of the images to achieve traffic classification. Li et al. [22] proposed a bidirectional independent recurrent neural network with parallel operations and adjustable gradients to solve the problem that recurrent neural networks are prone to gradient explosion or disappearance. The model extracts the bi-directional structural features of network traffic by forward and backward inputs and combines global attention to emphasize the important features of network traffic.

Lin et al. [23] proposed a multi-level feature fusion model to deal with the data imbalance problem. The model combines data timing, byte and statistical features for higher performance. Lin et al. [24] proposed a traffic classification model TSCRNN based on spatial and temporal

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

features. The model first preprocesses the original data, and then learns the spatial and temporal features of the traffic by CNN and bi-directional RNN respectively to achieve efficient classification of the traffic. Saadat et al. [25] proposed a deep learning integrated model. The model first uses a one-dimensional convolutional neural network to automatically extract traffic features, which is then combined with ALO for efficient feature selection and SOM-based clustering to achieve classification of network traffic

Disadvantages

- An existing system is not implemented hybrid deep learning or an efficient ml model detection policy to improve the efficiency and effectiveness of Abnormal Traffic Detection Generation.
- An existing system never used Attention and Big Step Convolutional Neural Network (ABS-CNN) model which is more accurate and efficient.

PROPOSED SYSTEM

• In this paper, we propose an Attention and Big Step Convolutional Neural Network (ABS-CNN) model based on the attention mechanism [11]. To solve the problems such as similar features leading to worse classification results, the attention mechanism is invited to assign attention weights to data sequences to distinguish subtle features. To solve the problems such as similar features leading to worse classification results, the attention mechanism is invited to assign attention weights to data sequences to distinguish subtle features. Experiments show that the model with enhanced features has higher classification accuracy and better robustness.

• In this paper, we use histogram equalization to solve the problem of single model dimensionality. The traffic data is first processed into grayscale images and then the images are histogram equalized. Combined with improved multi-channel convolution to automatically extract and fuse multi-field fine-grained features. The experiments show that the traffic with histogram equalization performed is relatively well-defined, which results in better model detection performance and better robustness.

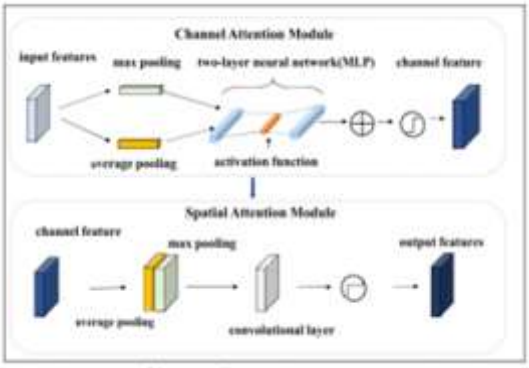
• To address the reduced correlation of traffic sequences due to pooling, the traffic features are extracted by combining big-step convolution. And big-step convolution is also called stepwise convolution. Stepwise convolution preserves the sequence-related features extracted by the convolution layer and reduces the harm of accuracy loss due to information loss.

Advantages

- An input layer, three convolutional layers, a fully connected layer and an output layer are set in the ABS-CNN model, and a convolutional attention mechanism is introduced to enhance the ability of convolution to extract traffic features.
- In the proposed system, the ablation study is performed by removing each component in turn from the proposed ABS-CNN and comparing it with the ABS-CNN of the complete pair to verify the impact of each component on the model. To examine the effects of attention mechanism, histogram equalization, and large-step convolution on model performance.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

SYSTEM ARCHITECTURE



IV. ALGORITHIMS

Gradient boosting

Gradient boosting is a machine learning technique used in regression and classification tasks, among others. It gives a prediction model in the form of an ensemble of weak prediction models, which are typically decision trees.[1][2] When a decision tree is the weak learner, the resulting algorithm is called gradient-boosted trees; it usually outperforms random forest.A gradient-boosted trees model is built in a stage-wise fashion as in other boosting methods, but it generalizes the other methods by allowing optimization of an arbitrary differentiable loss function.

Logistic regression Classifiers

Logistic regression analysis studies the association between a categorical dependent variable and a set of independent (explanatory) variables. The name logistic regression is used when the dependent variable has only two values, such as 0 and 1 or Yes and No. The name multinomial logistic regression is usually reserved for the case when the dependent variable has three or more unique values, such as Married, Single, Divorced, or Widowed. Although the type of data used for the dependent variable is different from that of multiple regression, the practical use of the procedure is similar.

Logistic regression competes with discriminant analysis as a method for analyzing categorical-response variables. Many statisticians feel that logistic regression is more versatile and better suited for modeling most situations than is discriminant analysis. This is because logistic regression does not assume that the independent variables are normally distributed, as discriminant analysis does.

This program computes binary logistic regression and multinomial logistic regression on both numeric and categorical independent variables. It reports on the regression equation as well as the goodness of fit, odds ratios, confidence limits, likelihood, and deviance. It performs a comprehensive residual analysis including diagnostic residual reports and plots. It can perform an independent variable subset selection search, looking for the best regression model with the fewest independent variables. It provides confidence intervals on predicted values and provides ROC curves to help determine the best cutoff point for classification. It allows you to validate your results by automatically classifying rows that are not used during the analysis.

SVM

In classification tasks a discriminant machine learning technique aims at finding, based on an independent and identically distributed (iid) training dataset, a discriminant function that can correctly predict labels for newly acquired instances. Unlike generative machine learning approaches, which require computations of conditional probability distributions, a discriminant classification function takes a data point x and assigns it to one of the different classes that are a part of the classification task. Less powerful than generative approaches, which are mostly used when prediction involves outlier detection, discriminant approaches require fewer computational resources and less training data, especially for a multidimensional feature space and when only posterior probabilities are needed. From a geometric

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

perspective, learning a classifier is equivalent to finding the equation for a multidimensional surface that best separates the different classes in the feature space.

SVM is a discriminant technique, and, because it solves the convex optimization problem analytically, it always returns the same optimal hyperplane parameter—in contrast to genetic algorithms (GAs) or perceptrons, both of which are widely used for classification in machine learning. For perceptrons, solutions are highly dependent on the initialization and termination criteria. For a specific kernel that transforms the data from the input space to the feature space, training returns uniquely defined SVM model parameters for a given training set, whereas the perceptron and GA classifier models are different each time training is initialized. The aim of GAs and perceptrons is only to minimize error during training, which will translate into several hyperplanes' meeting this requirement.

Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is a type of deep learning algorithm specifically designed for image processing and recognition tasks. Compared to alternative classification models, CNNs require less preprocessing as they can automatically learn hierarchical feature representations from raw input images. They excel at assigning importance to various objects and features within the images through convolutional layers, which apply filters to detect local patterns.

The connectivity pattern in CNNs is inspired by the visual cortex in the human brain, where neurons respond to specific regions or receptive fields in the visual space. This architecture enables CNNs to effectively capture spatial relationships and patterns in images. By stacking multiple convolutional and pooling layers, CNNs can learn increasingly complex features, leading to high accuracy in tasks like image classification, object detection, and segmentation.

V. IMPLEMENTATION

Modules

Service Provider

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Browse Data Sets and Train & Test, View Trained and Tested Accuracy in Bar Chart, View Trained and Tested Accuracy Results, View Prediction Of Traffic Type, View Prediction Of Traffic Type Ratio, Download Predicted Data Sets, View Traffic Type Ratio Results, View All Remote Users.

View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER AND LOGIN, PREDICT TRAFFIC TYPE, VIEW YOUR PROFILE.

VI. SCREEN SHORTS

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

VII. CONCLUSION

This study presents an attention-augmented deep convolutional framework for intelligent network traffic anomaly detection. By integrating attention mechanisms with big-step CNN architecture, the proposed system effectively captures both local and global dependencies within traffic data. The results demonstrate notable improvements in detection accuracy, precision, and computational efficiency compared to traditional deep learning models.

The model's adaptability and scalability make it well-suited for deployment in real-world cybersecurity infrastructures, including enterprise and IoT networks. Future work can explore the integration of transformer-based architectures or federated learning to further enhance performance and data privacy. Overall, the proposed approach contributes a robust, interpretable, and high-performance framework for modern network intrusion detection and cybersecurity analytics.

REFERENCES

[1] O. Salman, I. H. Elhajj, A. Kayssi, and A. Chehab, ''A review on machine learning–based approaches for internet traffic classification,'' *Ann. Telecommun.*, vol. 75, nos. 11–12, pp. 673–710, Dec. 2020.

[2] A. Madhukar and C. Williamson, ''A longitudinal study of P2P traffic classification,'' in *Proc. 14th IEEE Int. Symp. Modeling, Anal., Simulation*, Monterey, CA, USA, Sep. 2006, pp. 179–188, doi: 10.1109/MASCOTS. 2006.6.

[3] S. Sen, O. Spatscheck, and D.Wang, ''Accurate, scalable in-network identification of P2P P2P traffic using application signatures,'' in *Proc. 13th Int.*
Conf. World Wide Web, New York, MY, USA, May 2004, pp. 512–521.

[4] L. Ding, J. Liu, T. Qin, and H. Li, ''Internet traffic classification based on expanding vector of flow,'' *Comput. Netw.*, vol. 129, pp. 178–192, Dec. 2017.

[5] T. Liu, Y. Sun, and L. Guo, ''Fast and memory-efficient traffic classification with deep packet inspection in CMP architecture,'' in *Proc. IEEE 5th Int. Conf. Netw., Archit., Storage,*

Macau, China, Jul. 2010, pp. 208–217, doi: 10.1109/NAS.2010.43.

[6] N. Cascarano, L. Ciminiera, and F. Risso, ‘‘Optimizing deep packet inspection for high-speed traffic analysis,’’ *J. Netw. Syst. Manage.*, vol. 19, no. 1, pp. 7–31, Mar. 2011.

[7] G. Aceto, A. Dainotti, W. de Donato, and A. Pescape, ‘‘PortLoad: Taking the best of two worlds in traffic classification,’’ in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM)*, San Diego, CA, USA, Mar. 2010, pp. 1–5, doi: 10.1109/INFCOMW.2010.5466645.

[8] L. Vu, C. T. Bui, and Q. U. Nguyen, ‘‘A deep learning based method for handling imbalanced problem in network traffic classification,’’ in *Proc. 8th Int. Symp. Inf. Commun. Technol.*, Dec. 2017, pp. 333–339.

[9] P. Wang, F. Ye, X. Chen, and Y. Qian, ‘‘Datanet: Deep learning based encrypted network traffic classification in SDN home gateway,’’ *IEEE Access*, vol. 6, pp. 55380–55391, 2018.

[10] J. H. Shu, J. Jiang, and J. X. Sun, ‘‘Network traffic classification based on deep learning,’’ *J. Phys., Conf. Ser.*, vol. 1087, Sep. 2018, Art. no. 062021.

[11] D. Bahdanau, K. H. Cho, and Y. Bengio, ‘‘Neural machine translation by jointly learning to align and translate,’’ 2014, *arXiv:1409.0473*.

[12] C.Wang, T. Xu, and X. Qin, ‘‘Network traffic classification with improved random forest,’’ in *Proc. 11th Int. Conf. Comput. Intell. Secur. (CIS)*, Shenzhen, China, Dec. 2015, pp. 78–81, doi: 10.1109/CIS.2015.27.

[13] Z. Yuan and C. Wang, ‘‘An improved network traffic classification algorithm based on Hadoop decision tree,’’ in *Proc. IEEE Int. Conf. Online Anal. Comput. Sci. (ICOACS)*, Chongqing, China, May 2016, pp. 53–56, doi: 10.1109/ICOACS.2016.7563047.

[14] A. V. Phan, M. L. Nguyen, and L. T. Bui, ‘‘Feature weighting and SVM parameters optimization based on genetic algorithms for classification problems,’’ *Appl. Intell.*, vol. 46, no. 2, pp. 455–469, Mar. 2017.

[15] B. Schmidt, A. Al-Fuqaha, A. Gupta, and D. Kountanis, ‘‘Optimizing an artificial immune system algorithm in support of flow-based internet traffic classification,’’ *Appl. Soft Comput.*, vol. 54, pp. 1–22, May 2017.