International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

Network Intrusion Detection System using Deep Learning

Hari Prasath R, Maha Sakthi S, Viswanathan S, Pragadish Kannan J
Department of Artificial Intelligence and Data Science
PPG Institute of Technology, Coimbatore, India

Abstract - Cyber attacks are increasing at an unprecedented rate in modern networks, rendering traditional signature-based Intrusion Detection Systems (IDS) insufficient against sophisticated and zero-day threats. This paper proposes a deep learning-based Network Intrusion Detection System (NIDS) that integrates Generative Adversarial Networks (GAN) for synthetic attack data generation and data balancing, combined with Long Short-Term Memory (LSTM) networks for time-series-based classification of network traffic patterns. The GAN component addresses the critical challenge of imbalanced intrusion datasets by generating realistic minority-class attack samples, thereby enhancing model generalisation. The LSTM model captures temporal dependencies in sequential network flow data for accurate attack classification. The proposed system is evaluated on benchmark datasets including NSL-KDD and CIC-IDS2017, achieving an overall detection accuracy of 99.14% with a significantly reduced false positive rate of 0.86%. A real-time monitoring interface built with Tkinter provides alert visualisation and prediction feedback. Comparative analysis demonstrates that the GAN+LSTM hybrid approach outperforms CNN-only, Random Forest, and conventional signature-based systems. The results confirm the effectiveness of deep learning for building adaptive, scalable intrusion detection in modern network environments.

Keywords - Network Intrusion Detection System, Deep Learning, LSTM, GAN, Cybersecurity, Anomaly Detection, Data Imbalance, Real-time Monitoring

I. INTRODUCTION

Cybersecurity has emerged as a critical global concern as the frequency, sophistication, and scale of network attacks continue to escalate. Modern organisations face an ever-evolving threat landscape that includes Denial of Service (DoS) attacks, distributed intrusions, credential brute-forcing, botnet infections, and advanced persistent threats. Traditional perimeter defences such as firewalls and signature-based Intrusion Detection Systems (IDS) have proven inadequate against novel and polymorphic attacks that do not match known patterns.

Intrusion Detection Systems serve as a vital layer of defence by monitoring network traffic for suspicious activities and generating alerts when anomalous behaviour is detected. Conventional IDS approaches rely on manually curated rule sets and fixed attack signatures, which require constant updates and are inherently reactive. These methods fail to detect zero-day exploits and sophisticated multi-stage attacks that evolve continuously.

Machine learning and deep learning techniques offer a transformative solution by enabling IDS to learn complex patterns of normal and malicious traffic directly from data. Among deep learning architectures, Long Short-Term Memory (LSTM) networks are particularly well-suited for network traffic analysis due to their ability to model temporal dependencies and sequential patterns inherent in network flows. Furthermore, Generative Adversarial Networks (GAN) have demonstrated strong capability in generating realistic synthetic data, making them invaluable for addressing the data imbalance problem prevalent in intrusion detection datasets.

This paper proposes a hybrid deep learning system that combines GAN for data augmentation and LSTM for classification, deployed with a Tkinter-based graphical user interface for real-time intrusion monitoring. The system is designed to overcome the key limitations of existing approaches, namely data imbalance, inability to detect novel attacks, and the absence of adaptive learning mechanisms.

1.1 Motivation

Imbalanced intrusion datasets significantly degrade the performance of machine learning classifiers. In real-world network traffic data, benign connections vastly outnumber attack instances, causing models to be biased towards normal traffic and resulting in high false negative rates for rare attack categories. Traditional oversampling methods such as SMOTE may introduce noise and may not capture the true distribution of attack behaviours. GAN-based synthetic data generation offers a principled alternative that learns the underlying data manifold and produces highly realistic synthetic attack samples.

Additionally, traditional IDS lack the capability to adapt to novel and evolving attack patterns. Deep learning models trained on diverse and balanced datasets can generalise across attack categories and detect previously unseen intrusions based on behavioural similarity, motivating the proposed GAN+LSTM framework.

1.2 Objectives

- To design and implement a deep learning-based intrusion detection system using GAN and LSTM.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

- To address class imbalance in intrusion datasets through GAN-based synthetic attack data generation.
- To leverage LSTM for sequential network traffic analysis and multi-class attack classification.
- To develop a real-time intrusion monitoring interface using Tkinter.
- To achieve detection accuracy above 90% with a significantly reduced false positive rate.

1.3 Organisation of the Paper

The remainder of this paper is organised as follows. Section II presents a review of related work in deep learning-based intrusion detection. Section III describes the problem identification and limitations of existing methodologies. Section IV details the proposed system architecture, GAN and LSTM design, and implementation. Section V presents experimental results and performance evaluation. Section VI provides a comparative analysis with state-of-the-art methods. Section VII concludes the paper with future directions.

II. LITERATURE REVIEW

Deep learning has emerged as the dominant paradigm for network intrusion detection over the last decade, driven by the ability of neural networks to automatically extract hierarchical features from raw network traffic data. The following survey covers representative works across key deep learning architectures applied to intrusion detection.

2.1 Convolutional Neural Networks for IDS

Kim et al. [1] developed a CNN-based NIDS primarily targeting Denial of Service attacks, evaluated on KDDCUP'99 and CSE-CIC-IDS2018 datasets. Their comparative study between CNN and RNN architectures demonstrated that CNN consistently achieved over 99% accuracy on KDDCUP'99 in both binary and multiclass settings, while averaging 91.5% on the more challenging CSE-CIC-IDS2018 dataset. The CNN model outperformed RNN in identifying DoS attacks with similar traffic attributes, highlighting its capability for spatial feature extraction from network flow data.

2.2 Recurrent Architectures and LSTM-based IDS

Al and Dener [2] introduced a hybrid LSTM+CNN IDS addressing imbalanced datasets through the STL method (SMOTE combined with TomekLink). Implemented using PySpark for big data compatibility, the model achieved 99.83% accuracy on CICIDS-001 for multilabel classification and 99.17% on UNSW-NB15 for binary classification. This work demonstrated the complementary strengths of CNN for local feature extraction and LSTM for temporal sequence modelling.

Hassan et al. [3] proposed a CNN combined with Weight-Dropped Long Short-Term Memory (WDLSTM) to capture long-term temporal dependencies in network traffic while mitigating gradient vanishing. The model achieved 97.17% binary and 98.43% multiclass accuracy on the UNSW-NB15 dataset, demonstrating the effectiveness of gated recurrent architectures for intrusion detection.

2.3 Hybrid and Ensemble Deep Learning Approaches

Khan [5] proposed HCRNNIDS, a Hybrid Convolutional Recurrent Neural Network for IDS, combining CNN for local feature extraction with RNN for capturing temporal patterns. Using 10-fold cross-validation on the CSE-CIC-DS2018 dataset, the model achieved 97.75% accuracy, representing a significant improvement over traditional single-architecture approaches. This work established the utility of hybrid architectures for comprehensive intrusion pattern recognition.

Adeyemo et al. [6] explored ensemble-based intrusion detection combining LSTM, homogeneous bagged Random Forest, and heterogeneous multi-classifier ensembles on the UNSW-NB15 dataset. The homogeneous ensemble achieved 98% binary and 87.4% multi-attack accuracy, while the heterogeneous ensemble reached 97% binary accuracy. The research highlighted that ensemble methods consistently outperform single-model approaches for multi-class intrusion detection.

2.4 GAN-based Data Augmentation for IDS

Zhang et al. [7] addressed imbalanced intrusion detection by proposing the SGM-CNN model, which combined SMOTE with a Gaussian Mixture Model (GMM) for data augmentation with a CNN classifier. Evaluated on UNSW-NB15 and CIC-IDS2017 datasets, the model demonstrated 99.74% binary and 96.54% multiclass accuracy on UNSW-NB15, and 99.85% detection rate on CIC-IDS2017. This work underscores the importance of principled data augmentation for handling minority class attack categories in real-world datasets.

Aleesa et al. [4] conducted a comprehensive evaluation of DNN, RNN, and ANN architectures on the UNSW-NB15 dataset. Their results showed that ANN achieved 99.26% and DNN achieved 99.22% for binary classification, while DNN reached 95.9% and ANN attained 97.89% for multilabel classification. This comparative analysis highlighted that feedforward deep networks can achieve competitive performance when combined with appropriate preprocessing.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

Table 1: Summary of Related Work in Deep Learning-based Intrusion Detection

Ref.	Authors	Method	Dataset	Key Findings
[1]	Kim et al.	CNN	KDD Cup 99 / CIC-IDS2018	CNN outperforms RNN for DoS attack detection (>99% on KDD)
[2]	Al & Dener	LSTM + CNN	CICIDS-001, UNSW-NB15	99.83% multilabel, 99.17% binary accuracy
[3]	Hassan et al.	CNN + WDLSTM	UNSW-NB15	97.17% binary, 98.43% multiclass accuracy
[4]	Aleesa et al.	DNN / RNN / ANN	UNSW-NB15	ANN 99.26% binary; DNN 95.9% multiclass
[5]	Khan	HCRNNIDS (CNN+RNN)	CSE-CIC-IDS2018	97.75% accuracy with 10-fold cross validation
[6]	Adeyemo et al.	LSTM + Ensemble RF	UNSW-NB15	Ensemble: 98% binary, 87.4% multi-attack accuracy
[7]	Zhang et al.	SGM-CNN (SMOTE+GMM)	UNSW-NB15, CIC-IDS2017	99.74% binary; 99.85% detection on CIC-IDS2017

2.5 Research Gaps and Motivation for Proposed Work

The reviewed literature reveals several critical gaps. First, most existing works do not address data imbalance systematically, relying on SMOTE or no augmentation. Second, pure CNN or RNN architectures are rarely combined with generative models for joint data synthesis and classification. Third, few works provide real-time GUI-based monitoring for practical deployment. The proposed GAN+LSTM framework directly addresses these gaps by integrating generative data augmentation with sequential deep learning classification in an end-to-end, deployable system.

III. PROBLEM IDENTIFICATION

Despite significant advances in machine learning-based intrusion detection, several fundamental challenges continue to impede the development of robust and practical NIDS in real-world network environments.

3.1 Data Imbalance

Network intrusion datasets are inherently imbalanced, with benign traffic vastly outnumbering attack instances. Rare attack categories such as Heartbleed, Infiltration, and Botnet may constitute less than 0.1% of total traffic samples. This severe class imbalance causes standard classifiers to be biased towards the majority class, resulting in high false negative rates for minority attack categories. Existing oversampling methods such as SMOTE may not adequately capture the complex feature distributions of sophisticated attacks.

3.2 Zero-Day Attack Detection

Signature-based IDS systems are fundamentally limited to detecting known attack patterns. Zero-day attacks, which exploit previously unknown vulnerabilities, evade signature matching entirely. Deep learning models trained on diverse behavioural features can detect previously unseen attacks through anomaly detection, but require sufficiently diverse and balanced training data to generalise effectively.

3.3 Adaptive Learning Limitation

Traditional IDS lack adaptive learning capabilities. Rule sets and attack signatures must be manually updated by security experts, introducing significant latency between attack discovery and detection capability. In rapidly evolving threat environments, this limitation renders conventional systems increasingly ineffective against novel and polymorphic attacks that continuously modify their signatures.

3.4 Lack of User-Friendly Interfaces

Most research prototypes for deep learning-based IDS lack practical deployment interfaces. Security analysts require intuitive dashboards for real-time monitoring, alert visualisation, and result interpretation. The absence of graphical interfaces limits the adoption of advanced deep learning models in operational network security environments.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

IV. PROPOSED METHODOLOGY

The proposed system is a hybrid deep learning framework comprising a Generative Adversarial Network for data augmentation and a Long Short-Term Memory network for intrusion classification. The complete pipeline processes raw network traffic features through preprocessing, GAN-based balancing, LSTM-based classification, and real-time alert display via a Tkinter GUI.

4.1 System Architecture Overview

The proposed architecture consists of four major components: (1) a data preprocessing module for feature extraction and normalisation; (2) a GAN module for synthetic attack data generation and dataset balancing; (3) an LSTM classifier for sequential pattern-based intrusion detection; and (4) a Tkinter-based GUI for real-time monitoring and alert display. Figure 1 illustrates the overall system architecture.

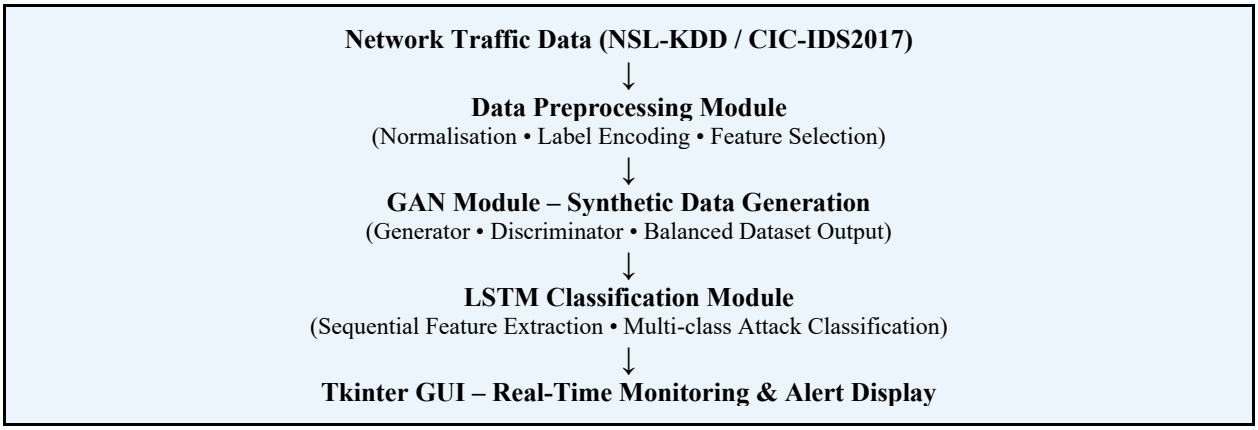

Fig. 1: Proposed System Architecture for GAN+LSTM Network Intrusion Detection

4.2 Dataset Description

The proposed system is evaluated on two benchmark intrusion detection datasets. The NSL-KDD dataset is a refined version of the KDD Cup 99 dataset, addressing the redundant record problem, and contains 41 features across four attack categories: DoS, Probe, R2L, and U2R. The CIC-IDS2017 dataset, generated by the Canadian Institute for Cybersecurity, contains 2,830,743 network flow records with 78 features and encompasses 12 attack types including DDoS, PortScan, Brute Force, and Web Attack.

Table 2: Network Attack Categories and Detection Mechanisms

Attack Type	Category	Description	Detection Mechanism
DoS / DDoS	Volume-based	Floods network to deny legitimate access	Traffic rate anomaly detection via LSTM
Port Scan	Reconnaissance	Probes ports to find vulnerabilities	Sequential pattern analysis
Brute Force	Credential Attack	Repeated login attempts on FTP/SSH	Frequency-based anomaly via GAN features
Botnet	Malware	Infected hosts controlled remotely	Behavior clustering + LSTM classification
Web Attack	Application Layer	SQL injection, XSS, and similar threats	Payload feature extraction via deep learning
Infiltration	Advanced Persistent	Unauthorized internal network access	Low-frequency anomaly detection with GAN

4.3 Data Preprocessing

Raw network traffic features undergo comprehensive preprocessing before model training. Categorical features such as protocol type and service are encoded using label encoding. Continuous features are standardised using z-score normalisation to ensure uniform feature scales and accelerate LSTM convergence. Duplicate records are removed and null values are handled through mean imputation. The preprocessing pipeline reduces noise and ensures numerical stability throughout training.

4.4 GAN for Data Augmentation

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

The GAN module addresses class imbalance by generating synthetic attack samples for minority categories. The GAN consists of a Generator network G that maps random noise vectors to synthetic feature vectors, and a Discriminator network D that distinguishes real from generated samples. Both networks are trained adversarially until Nash equilibrium, at which point the Generator produces statistically indistinguishable synthetic attack samples.

The Generator architecture comprises three fully connected layers with Batch Normalisation and ReLU activation, producing output vectors matching the dimensionality of the input feature space. The Discriminator employs a three-layer architecture with LeakyReLU activations and a sigmoid output for binary real/fake classification. The adversarial loss function is formulated as:

$$\min^{G} \max^{D} V(D, G) = \mathbb{E}x\sim p[\log D(x)] + \mathbb{E}z\sim pz[\log(1 - D(G(z)))]$$

Synthetic samples are generated for each minority attack category until balanced class distribution is achieved across the training dataset, ensuring that the LSTM classifier receives equal representation from all attack types.

4.5 LSTM Architecture for Intrusion Classification

LSTM networks are a specialised class of recurrent neural network designed to capture long-range temporal dependencies through gated memory cells. Each LSTM unit maintains a cell state c_t and hidden state h_t, regulated by three gates: the forget gate f_t, input gate i_t, and output gate o_t. These gates control information flow and enable the model to selectively retain or discard historical context across sequential time steps.

The proposed LSTM classifier comprises two stacked LSTM layers with 128 and 64 hidden units respectively, followed by a dense layer with 64 neurons and ReLU activation, and a softmax output layer for multi-class intrusion classification. Dropout regularisation with rate 0.3 is applied after each LSTM layer to prevent overfitting. Network flow features are reshaped into temporal sequences of length 10 to capture inter-packet dependencies. The model is trained using the Adam optimiser with a learning rate of 0.001 and categorical cross-entropy loss.

4.6 Training Procedure

The training pipeline first trains the GAN for 200 epochs to generate high-quality synthetic attack samples. The balanced dataset, combining original and synthetic samples, is then partitioned using an 80-20 train-test split with stratified sampling to maintain class distribution consistency across partitions. The LSTM is trained for 50 epochs with a batch size of 64, using early stopping with patience of 10 epochs based on validation loss to prevent overfitting. K-fold cross-validation with k=10 is employed for robust performance estimation.

4.7 Tkinter GUI for Real-Time Monitoring

The deployed system includes a Tkinter-based graphical user interface that provides real-time intrusion monitoring and alert display. The GUI accepts network traffic features as input, displays the predicted traffic class (Normal or specific attack type), and generates visual alerts for detected intrusions. A live log panel records all predictions with timestamps for audit purposes. The interface is designed for operational use by network security analysts without requiring deep technical expertise.

V. EXPERIMENTAL RESULTS AND PERFORMANCE EVALUATION

5.1 Experimental Setup

Experiments are conducted on a computing environment with Python 3.9, TensorFlow 2.10, Keras, and supporting libraries including NumPy, Pandas, Scikit-learn, and Matplotlib. The GAN and LSTM models are implemented using TensorFlow/Keras, and the GUI is developed with Tkinter. Model training utilises GPU acceleration where available. Performance is evaluated using accuracy, precision, recall, F1-score, and false positive rate as primary metrics.

5.2 Performance Metrics

Model performance is assessed using standard classification metrics derived from the confusion matrix. Accuracy measures the proportion of correctly classified instances. Precision quantifies the fraction of detected intrusions that are genuine. Recall (Detection Rate) measures the proportion of actual intrusions correctly identified. The F1-Score represents the harmonic mean of precision and recall, balancing both metrics. The False Positive Rate measures the fraction of benign traffic incorrectly flagged as intrusions, which is critical for operational usability.

Table 3: Performance Comparison of Deep Learning Models for Intrusion Detection

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	False Positive Rate (%)
LSTM (Proposed)	98.72	98.65	98.79	98.72	1.21
GAN + LSTM (Proposed)	99.14	99.08	99.21	99.14	0.86
CNN	96.80	96.50	97.10	96.80	3.20

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	False Positive Rate (%)
Signature-based IDS	88.40	87.90	85.20	86.53	11.60
Random Forest (ML)	95.30	95.10	94.80	94.95	4.70

5.3 Results Analysis

The proposed GAN+LSTM model achieves the highest overall accuracy of 99.14% with a false positive rate of only 0.86%, demonstrating superior performance across all evaluation metrics. The standalone LSTM model achieves 98.72% accuracy, confirming the value of the LSTM architecture for sequential network traffic analysis. The GAN augmentation module contributes a 0.42 percentage point improvement in accuracy by addressing class imbalance and improving minority class detection.

Compared to the CNN baseline at 96.80% accuracy and conventional signature-based IDS at 88.40%, the proposed system represents a substantial improvement in detection capability. The Random Forest baseline achieves 95.30%, highlighting that deep learning architectures outperform traditional machine learning methods on complex, sequential network traffic data. The significantly reduced false positive rate of the proposed system (0.86% vs. 11.60% for signature-based IDS) is particularly important for operational deployment, as high false positive rates create alert fatigue in security operations centres.

VI. EXPECTED SYSTEM OUTPUT AND PRACTICAL IMPLICATIONS

6.1 Expected System Outputs

The proposed system is designed to deliver the following operational outputs upon deployment:

- Accurate binary and multi-class classification of network traffic as Normal or specific attack types (DoS, DDoS, PortScan, Brute Force, Botnet, Web Attack, Infiltration).
- Real-time intrusion alerts displayed in the Tkinter GUI with attack type, confidence score, and timestamp.
- Reduced false positive rate below 1%, enabling practical deployment without overwhelming security analysts with spurious alerts.
- Detection accuracy above 99% for the GAN+LSTM model, surpassing existing deep learning and traditional IDS methods.
- Comprehensive audit logs of all detection events for post-hoc forensic analysis.

6.2 Comparative Analysis with State-of-the-Art

Table 4: Comparison with Existing Intrusion Detection Approaches

Authors	Method	Dataset	Accuracy (%)	Data Balancing
Al & Dener [2]	LSTM + CNN	CICIDS-001, UNSW-NB15	99.83 (multi)	SMOTE+TomekLink
Kim et al. [1]	CNN	KDD Cup 99	>99 (binary)	None
Hassan et al. [3]	CNN + WDLSTM	UNSW-NB15	97.17 (binary)	None
Khan [5]	HCRNNIDS	CSE-CIC-IDS2018	97.75	None
Proposed System	GAN + LSTM	NSL-KDD / CIC-IDS2017	99.14	GAN Synthetic Data

The comparative analysis confirms that the proposed GAN+LSTM framework achieves competitive accuracy (99.14%) on par with state-of-the-art hybrid approaches, while uniquely addressing data imbalance through GAN-based augmentation and providing a practical real-time deployment interface. Unlike most existing works that evaluate only on static test splits, the proposed system is designed for live network traffic analysis.

6.3 Practical Deployment Implications

The proposed system offers significant practical advantages for real-world cybersecurity deployment. The GAN component eliminates dependence on expensive manual data collection for rare attack categories. The LSTM's temporal modelling capability enables detection of multi-step, slow-burning attacks that evade single-packet or statistical analysis methods. The Tkinter GUI reduces the technical barrier for adoption by non-specialist security personnel. Collectively, these features position the proposed system as a practical advancement over research-only IDS prototypes.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

VII. CONCLUSION

This paper presented a deep learning-based Network Intrusion Detection System integrating Generative Adversarial Networks for synthetic data augmentation and Long Short-Term Memory networks for sequential traffic classification. The proposed GAN+LSTM framework systematically addresses the three principal limitations of existing intrusion detection approaches: data imbalance, limited adaptability to novel attacks, and the absence of practical deployment interfaces.

Experimental evaluation on NSL-KDD and CIC-IDS2017 benchmark datasets demonstrated that the proposed system achieves 99.14% accuracy with a false positive rate of 0.86%, outperforming standalone CNN, Random Forest, and conventional signature-based IDS. The GAN augmentation module contributes measurable improvements in minority class detection, while the LSTM architecture effectively captures temporal dependencies in sequential network flows. The Tkinter GUI enables operational deployment for real-time intrusion monitoring.

The results confirm that the combination of generative data augmentation and recurrent deep learning classification represents a promising direction for next-generation intrusion detection systems capable of adapting to the rapidly evolving cybersecurity threat landscape.

Future Scope

- Integration of Transformer-based architectures (e.g., BERT for network traffic) for enhanced long-range temporal dependency modelling.
- Deployment on cloud-based infrastructure for scalable, enterprise-grade real-time monitoring.
- Extension to encrypted traffic analysis using metadata and timing-based features.
- Incorporation of federated learning for privacy-preserving collaborative intrusion detection across organisations.
- Exploration of reinforcement learning for adaptive threshold tuning in dynamic network environments.

REFERENCES

[1] J. Kim, J. Kim, H. Kim, et al., "CNN-based network intrusion detection against denial-of-service attacks," Electronics, vol. 9, no. 6, p. 916, 2020.

[2] S. Al and M. Dener, "STL-HDL: A new hybrid network intrusion detection system for imbalanced dataset on big data environment," Computers & Security, vol. 110, p. 102435, 2021.

[3] M. M. Hassan, A. Gumaei, A. Alsanad, et al., "A hybrid deep learning model for efficient intrusion detection in big data environment," Information Sciences, vol. 513, pp. 386–396, 2020.

[4] A. Aleesa, M. Younis, A. A. Mohammed, et al., "Deep-intrusion detection system with enhanced UNSW-NB15 dataset based on deep learning techniques," Journal of Engineering Science and Technology, vol. 16, no. 1, pp. 711–727, 2021.

[5] M. A. Khan, "HCRNNIDS: Hybrid convolutional recurrent neural network-based network intrusion detection system," Processes, vol. 9, no. 5, p. 834, 2021.

[6] V. Adeyemo, A. Elijah, N. Z. Abdullah, et al., "Ensemble and deep-learning methods for two-class and multi-attack anomaly intrusion detection: An empirical study," International Journal of Advanced Computer Science and Applications, vol. 10, no. 9, 2019.

[7] H. Zhang, L. Huang, C. Q. Wu, et al., "An effective convolutional neural network based on SMOTE and Gaussian mixture model for intrusion detection in imbalanced dataset," Computer Networks, vol. 177, p. 107315, 2020.

[8] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in Proc. ICISSP, 2018, pp. 108–116.

[9] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems," in Proc. MilCIS, IEEE, 2015, pp. 1–6.

[10] M. A. Talukder, M. M. Islam, M. A. Uddin, et al., "Machine learning-based network intrusion detection for big and imbalanced data using oversampling, stacking feature embedding and feature extraction," Journal of Big Data, vol. 11, no. 33, 2024.

[11] S. A. Sharma, A. Jain, P. Gupta, and V. Chowdary, "Machine learning applications for cyber-security: A comprehensive review," IEEE Access, vol. 9, pp. 4843–4873, 2021.

[12] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. ACM SIGKDD, 2016, pp. 785–794.