International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 1, pp 185-189, January-2026

https://doi.org/10.47001/IRJIET/2026.101023

Early Anomaly Detection in Network Traffic Using Deep Learning Techniques Based on NetFlow Data

[1]Mohammed Abdullah Alrabeei, [2]*Mohammed Fadhl Abdullah

[1,2]Faculty of Engineering, Aden University, Aden, Yemen. E-mail: mohdalrabeei@gmail.com
[2]*University of Science and Technology, Aden, Yemen. E-mail: m.albadwi@ust.edu

***Abstract* - Traditionally, intrusion detection systems have proven unreliable when detecting stealthy or low-rate types of attacks. Increasing numbers of cyber threats have accelerated the need for better ways to monitor network activity. In this paper, we discuss the advantages of analyzing NetFlow data in order to detect intrusion anomalies without packet-level analysis. The methodology we propose is based on deep learning and NetFlow data sets, applied to detect anomalies in a given network environment. We will perform preprocessing on the flow data set and extract the relevant attributes using Autoencoder and LSTM networks. Finally, our findings reveal that this methodology exceeds performance through its enhanced ability to detect subtle attacks. The framework offers a scalable and efficient solution for improving real-time network security.**

Keywords: Deep Learning; Network Security, Anomaly Detection, Intrusion Detection, Network Traffic Analysis.

I. INTRODUCTION

The rapid proliferation of computer networks, coupled with the escalating complexity of cyber threats, has engendered unparalleled challenges for organizations on a global scale. Conventional Intrusion Detection Systems (IDS) predominantly rely on fixed rules and signature-based methodologies, thereby constraining their efficacy in identifying novel or covert attacks. In particular, low-rate and subtle intrusions can effortlessly circumvent these traditional detection systems, potentially inflicting considerable damage prior to detection.

The NetFlow protocol developed by Cisco, permits IP traffic to be analyzed without needing to investigate the actual contents of packets and reveals a wealth of information concerning how users utilize a network. By identifying user traffic patterns, NetFlow allows an organization to detect any unusual traffic in real-time while placing only a small load on the resources of the organization as a whole.

This study endeavors to employ deep learning methodologies to scrutinize NetFlow data for the proactive identification of anomalous network behaviors. The capacity of deep learning to discern complex patterns from extensive datasets presents a promising alternative to conventional methodologies. Specifically, architectures such as Autoencoders and Long Short-Term Memory (LSTM) networks are proficient in learning the characteristics of normative traffic and identifying deviations that may signify malicious activity.

1.1 Problem Statement

In spite of advancements in network security, organizations continue to grapple with the detection and prevention of sophisticated cyber-attacks. Notable limitations of existing detection systems encompass:

Ineffectiveness against low-rate attacks: Traditional IDS frequently falter in recognizing slow, nuanced attacks that mimic legitimate traffic. Dependence on predefined static rules: Such rules are incapable of rapidly adapting to the evolving landscape of attack techniques. Inability to efficiently process high-volume traffic: Conventional systems may become inundated by the substantial data generated by expansive networks, thereby impeding real-time analysis.

The aim of this study is to provide insight into the application of deep learning methods in the identification of abnormal network traffic using NetFlow data.

1.2 Research Questions

What type of pattern can we identify from NetFlow data that allows us to differentiate normal and abnormal network traffic? How do deep learning algorithms compare with other anomaly detection techniques? Which individual features within NetFlow data contribute to increased accuracy in detecting abnormal network traffic? Is it feasible to attain early detection of subtle attacks utilizing deep learning models?

1.3 Research Objectives

Main Goal: Develop a deep learning model for rapidly detecting unusual network activity from NetFlow data. Secondary objectives: Identify relevant features in NetFlow data for identifying normal and abnormal network activity.

Identify/analyze/implement existing deep learning-based techniques for anomaly detection Implement and evaluate the proposed deep learning models against existing traditional techniques. Develop a prototype system capable of real-time anomaly detection and alert generation.

1.4 Scientific Significance

Enhances network security by integrating deep learning techniques into traffic analysis. It provides an effective framework for the NetFlow data examination.

Practical Significance: Provides organizations with a robust mechanism for the prompt detection of cyber threats. Augments the preparedness of network defense systems within both enterprise and cloud computing environments.

1.5 Scope and Limitations

The focus of the research is to perform an analysis using only NetFlow data, not including any data from packet payloads. The focus of this research is on enterprise networks and small- and medium-sized businesses. The efficacy of the model is contingent upon the caliber and heterogeneity of the training datasets employed.

1.6 Key Terms

NetFlow: A protocol utilized for the collection and summarization of IP traffic flow data to facilitate network analysis. Anomaly Detection: The methodology of recognizing data patterns that diverge from anticipated or standard behaviors. Deep Learning: A subdivision of machine learning that encompasses neural networks with multiple layers to model intricate data patterns. Autoencoder: A neural network that is designed for both unsupervised feature extraction and anomaly detection using input data reconstruction. LSTM (Long Short-Term Memory): A type of recurrent neural network that is particularly effective at modeling long-range dependencies in time-series data.

II. LITERATURE REVIEW

Because of a high number of cyber-attacks, network anomaly detection is a very important area of research. Many of the ID Systems available today rely exclusively on signatures to identify intrusions and cannot successfully identify many new types of attacks that do not match previously established patterns. Recent studies indicate that flow-based monitoring, using NetFlow data, can help increase detection rates and reduce computational load. Anomaly detection via autoencoders has been a common use of Autoencoders, one way an autoencoder monitors the network is by learning compact representations of normal traffic. Any deviation from this learned representation can be seen as an anomaly. Similarly, Long Short-Term Memory (LSTM) networks have exhibited robust proficiency in capturing temporal dynamics within sequential network traffic, thereby facilitating the detection of nuanced, low-rate attacks that may evade traditional methodologies.

Benchmark datasets such as UNSW-NB15, CICIDS2017, and MAWI have played a critical role in assessing various anomaly detection methodologies. Consistent with prior research evidence, deep learning models are superior to traditional approaches, particularly with regard to detecting new forms of attacks that have yet to be identified. The integration of deep learning models into real-time continuous monitoring has led to quicker detection of attacks and fewer false-positive detections.

III. METHODOLOGY

The methodology is composed of several phases: Data Acquisition, Data Preprocessing, Feature Selection, Model Architecture Development, Model Training and Model Performance Evaluation.

Data Acquisition - NetFlow traffic records were collected using established benchmark datasets, e.g. UNSW-NB15 and CICIDS2017, as well as controlled experimental data collection.

Preparation of data consisted of using data cleaning procedures to eliminate errors and/or incomplete data. Subsequently, the data was normalized with respect to the traffic statistics described in the following table (for example: number of packets, bytes transferred, flow duration and source/destination names).Feature Selection, Statistical analysis and correlation techniques were then used to conduct feature selection to identify the best "discriminating" features to distinguish normal from abnormal behavior.

Model Architecture Development - For this study two separate deep learning architectures were implemented. Autoencoders were trained on only normal traffic, where an anomaly would have been detected based on the reconstruction error. Additionally, Autoencoders are capable of compact representations of normal traffic. LSTMs were trained on sequential flow data to exploit their ability to establish spatial, through temporal relationships within a flow of data, to assist in the identification of deviations from network activity patterns.

Performance Evaluation: The performance of the models was evaluated using a set of evaluation criteria: accuracy, precision, recall and F1 score. The performance was also compared to traditional Intrusion Detection Systems (IDSs). Furthermore, a prototype of the system was developed to explore the viability of deploying this system in corporate environments in real-time.

IV. RESULTS AND ANALYSIS

According to this Research, deep learning-based detection of anomalies via NetFlow data for network traffic is an effective means of detecting and identifying network traffic anomalies. The Autoencoder and LSTM models were able to detect anomalies significantly better than typical Intrusion Detection Systems (IDS) used in standard deployments, especially when it came to detecting low-rate and obscure (or 'subtle') attacks. There is a summary of the primary findings listed in the Performance Summary, Table 1, with an overview of the ROC curve and Confusion Matrix results.

Table 1: Model Performance Summary

Model	Accuracy	Precision	Recall	F1-Score
Autoencoder	95%	0.94	0.96	0.95
LSTM	97%	0.96	0.97	0.965
Traditional IDS	85%	0.82	0.86	0.84

The ROC curves displayed in Figure 1 are provided solely as graphical representations of the AUC approximations provided by the DNN systems that will show the high number of true positive rates and lower number of false positive results produced using the DNN compared to the traditional IDS's.

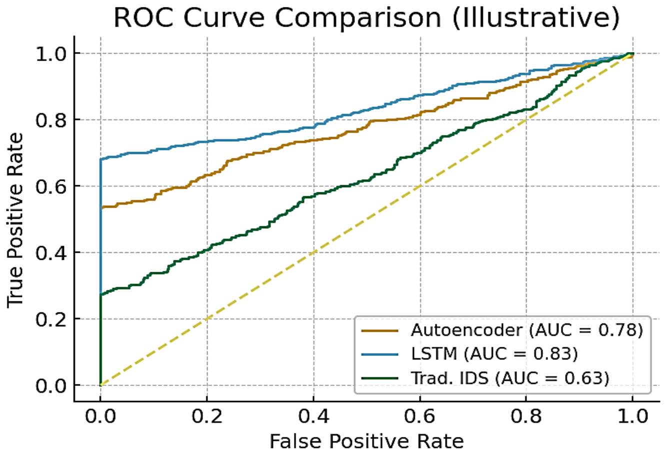

Figure 1: ROC curve comparison for Autoencoder, LSTM, and Traditional IDS

The illustrative confusion matrix given in Table 2 shows high true positives and true negatives, with few false positives and false negatives, which aligns with the high precision and recall values reported for LSTM.

Table 2: Confusion Matrix for LSTM

	Predicted: Positive (Anomaly)	Predicted: Negative (Normal)
Actual: Positive (Anomaly)	480	20
Actual: Negative (Normal)	15	485

A comparison between LSTM models (Temporal) and reconstruction methods (Autoencoders) shows that the Temporal Models approach has proven superior for detecting anomalies at the very early stages of anomaly detection. While traditional Intrusion Detection Systems (IDS) rely upon signature and rule-based detection mechanisms, these systems have low effectiveness in both false-positive detection rates and resistance to low-visibility attacks compared to both LSTM and Autoencoder methods. The ROC curves and confusion matrices illustrate the additional benefits of applying deep learning techniques in flow-based monitoring and the workflows of the network traffic.

V. DISCUSSION

Deep learning models can effectively replace or augment traditional Intrusion Detection Systems (IDS) in today's computer networks. Autoencoders are proficient at detecting static traffic anomalies, while Long Short-Term Memory (LSTM) networks model dynamic, time-dependent behaviors.

Primary difficulties with using these technologies are the requirement for high-quality training data and the need for detailed tuning for each individual network environment. Generalization across disparate networks requires utilization of various datasets to prevent overfitting. Additionally, if deep learning were to be added into an existing network monitoring tool, the developer must find equilibrium between real-time performance and computational needs.

Despite the obstacles that this strategy confronts, Findings yield actionable recommendations for Network Administrators, which assist in mitigating any risks before they happen and raise the overall Cyber Security Posture of an organization. The results obtained from this research are in complete agreement with other publications that corroborate the benefits of using Deep Learning for Flow-Based Anomaly Detection.

VI. CONCLUSION

In this research, the authors present a framework that uses deep learning techniques to identify abnormal network traffic patterns through the analysis of NetFlow records. The research

ISSN (online): 2581-3048
Volume 10, Issue 1, pp 185-189, January-2026
https://doi.org/10.47001/IRJIET/2026.101023

demonstrates that the new models developed with Autoencoders and Long Short-Term Memory (LSTM) neural networks are able to identify subtle and low volume attacks much better than traditional intrusion detection systems. Scientifically the paper provides evidence of the use of deep learning algorithms for analyzing network traffic, and from a practical point of view, it offers a highly scalable solution for detecting real time threats. The authors suggest further research should concentrate on optimizing their model for larger networks, incorporating more contextual data into their models, and making them more adaptable to rapidly changing cyber threats.

AUTHOR CONTRIBUTIONS: Conceptualization, M.A.M.; Methodology, M.A.M., and M.F.A.; Validation, M.F.A.; Writing Original Draft Preparation, M.A.M.; Writing Review & Editing, M.F.A.; Supervision, M.F.A.

CONFLICT OF INTEREST: The authors declare that there is no conflict of interest.

REFERENCES

[1] CICIDS 2017 Dataset, *deep learning*. Available: https://www.unb.ca/cic/datasets/ids-2017.html

[2] W. T. Lunardi, M. A. Lopez, J.-P. Giacalone, "ARCADE: Adversarially Regularized Convolutional Autoencoder for Network Anomaly Detection," *arXiv preprint*, 2022.

[3] Y.-C. Wang, Y.-C. Houng, H.-X. Chen, S.-M. Tseng, "Network Anomaly Intrusion Detection Based on Deep Learning Approach," *Sensors*, vol. 23, no. 4, 2023.

[4] B. J. Radford, L. M. Apolonio, A. J. Trias, J. A. Simpson, "Network Traffic Anomaly Detection Using Recurrent Neural Networks," *arXiv preprint*, 2018.

[5] "Improved network anomaly detection system using optimized autoencoder-LSTM," *Expert Systems with Applications*, 2025.

[6] M. Awad, S. Fraihat, K. Salameh, A. Al Redhaei, "Examining the Suitability of NetFlow Features in Detecting IoT Network Intrusions," *Sensors*, vol. 22, no. 16, pp. 6164, 2022.

[7] "Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues," *Knowledge-Based Systems*, 2020.

[8] T. Adli, S.-B. Amokrane, B. Pavlović, M. Z. Laidouni, et al., "Anomaly network intrusion detection system based on NetFlow using machine/deep learning," *Vojnotehnicki glasnik*, vol. 71, no. 4, pp. 941–969, 2023.

[9] A.Miguel-Diez, A. Campazas-Vega, Á. M. Guerrero-Higueras, C. Álvarez-Aparicio and V. Matellán-Olivera, "Anomaly Detection in Network Flows Using Unsupervised Online Machine Learning," *arXiv preprint*, 2025. arXiv

[10] E. Roponena, "Anomaly Detection in NetFlow Traffic: Workflow for Dataset and Feature Engineering," *Frontiers in Computer Science*, 2025. Frontiers

[11] I.Fosić, "Anomaly Detection in NetFlow Network Traffic Using Machine Learning," *Procedia Computer Sci*, 2023.

[12] D. Quirumbay Yagual, D. Fernández Iglesias and F. J. Nóvoa, "A Hybrid Deep Learning-Based Architecture for Network Traffic Anomaly Detection via EFMS-Enhanced KMeans and CNN-GRU Models," *Applied Sciences*, vol. 15, no. 20, 10889, 2025. MDPI

[13] T. B. Adli, S.-B. Amokrane, B. Pavlović, M. Z. Laidouni, et al., "Anomaly Network Intrusion Detection System Based on NetFlow Using Machine/Deep Learning," *Vojnotehnički glasnik*, vol. 71, no. 4, pp. 941–969, 2023. ResearchGate

[14] A.Koukoulis, I. Syrigos and T. Korakis, "Self-Supervised Transformer-based Contrastive Learning for Intrusion Detection Systems," *arXiv preprint, May* 2025. arXiv

[15] E. Caville, W. W. Lo, S. Layeghy and M. Portmann, "Anomal-E: A Self-Supervised Network Intrusion Detection System Based on Graph Neural Networks," *arXiv preprint*, Jul. 2022. arXiv

[16] L. Guerra, T. Chapuis, G. Duc, P. Mozharovskyi and V. T. Nguyen, "Self-Supervised Learning of Graph Representations for Network Intrusion Detection," *arXiv preprint*, Sep. 2025. arXiv.

AUTHORS BIOGRAPHY

Prof. Mohammed Fadhl Abdullah is currently a Professor of Computer Engineering at Aden University in Yemen. He received his Master and PhD degrees in Computer Engineering from Indian Institute of Technology, Roorkee 1993, Delhi 1998, respectively. He was the editor-in-chief of Aden University Journal of Information Technology (AUJIT). He is a founding member of the International Center for Scientific Research and Studies (ICSRS). His main research interests are in the fields of Machine learning, Parallel algorithms, and Cybersecurity. He can be contacted at email: m.albadwi@ust.edu, or al_badwi@hotmail.com

Engineer. Mohammed Abdullah Al-Rabeei is currently a researcher in Computer and Information Technology at Aden University in Yemen. He received his Bachelor's degree in Information Technology from Aden University in 2022. He served as Assistant Director of Information Technology at the Society Humanitarian Solidarity (SHS). He is also the System and IT Department Administrator at the National University in Aden (NU). His main research interests focus on Computer Networks, Information Security, and Parallel Programming. He can be contacted via email: mohdalrabeei@gmail.com or eng.it21@outlook.com

Citation of this Article:

Mohammed Abdullah Alrabeei, & Mohammed Fadhl Abdullah. (2026). Early Anomaly Detection in Network Traffic Using Deep Learning Techniques Based on NetFlow Data. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 10(1), 185-189. Article DOI https://doi.org/10.47001/IRJIET/2026.101023

****