Cyber Threat Detection in Network Traffic using Random Forest for Enhanced Quality of Service

[1]Lasisi H.O., [2]Omodeni C.B., [3]Aderinkola B.F., [4]Omodeni H.B., [5]Oladepo O., [6]Olawuyi A.A., [7]Adeagbo F.M., [8]Awolufolaju T.T., and [9]Olla M.O.
[1,2,4,5,6,7,8,9]Department of Electrical and Electronics Engineering, Osun State University, Osogbo, Nigeria
[3]Department of Electrical and Electronics Engineering, The Federal Polytechnic, Ede, Osun State, Nigeria
hammed.lasisi@uniosun.edu.ng | celebless_2011@yahoo.com | hammed.omodeni@pgc.uniosun.edu.ng | fattyade2307@gmail.com

Received: 22-MAY-2026; Reviewed: 28-JUNE-2026; Accepted: 01-JULY-2026
https://dx.doi.org/10.46792/fuoyejet.v11i3.1592

ORIGINAL RESEARCH

Abstract— It is critical to ensure high Quality of Service (QoS) and robust security in modern telecommunications networks that are increasingly targeted by large cyber-attacks. The study employs CIC-IDS2017 dataset with over 3.1 million traffic records, and a Random Forest (RF) ensemble model is employed to categorize various attack types such as DoS, DDoS, and Brute Force. The methodology emphasizes a balance of security and network availability to achieve maximum performance. To validate the proposed model, experimental findings demonstrate that it has a high accuracy of 99.62% and a Detection Rate (Recall) of 99.91%. In particular, its False Alarm Rate (FAR) and Specificity were 0.45% and 99.55%, respectively, ensuring network QoS without causing additional disturbance from false alarms. The system reliability is determined by efficiency score (Q_{score}) to ensure that network operations and user experience are not unnecessarily disturbed. The integrated Q_{score} of 221.97 adds to the system's effectiveness in establishing a secure and transparent communications environment. The model can process a large number of network traffic samples per second, with a predicted throughput of 42,256 samples per second. A prediction latency of 0.024ms per sample indicates that traffic samples are identified almost instantaneously. The results show that the RF algorithm is a viable approach for real-time threat identification in high-speed network infrastructures.

Keywords— Cyber threat, CIC-IDS2017, false alarm rate, intrusion detection system (IDS), network security, quality of service (QoS),Random forest.

1 INTRODUCTION

Telecommunication networks are vulnerable to various types of anomalies, including network intrusions, equipment failures, and traffic congestion, which can result in service disruptions, security breaches, and revenue losses (Sarah et al., 2023). Machine learning (ML) has become a prominent approach as the new paradigm in the development of intelligent adaptive anomaly detection systems because ML processes can learn high-dimensional complex trends of normal behavior and adapt to changing conditions rather than relying on hard-coded signatures or human-engineered rules. Telecommunication networks have become more complex as a result of rapid growth of 5G, IoT, and high-speed network infrastructures (Trilles et al., 2024). These networks must now provide ultra-low latency, high throughput, and uninterrupted connectivity, which necessitates billions of networked devices (both industrial sensors and autonomous vehicles) (Haque et al., 2023). Although hyper-connectivity has the potential to generate transformations, it has also enlarged the attack surface, and networks face a slew of security risks (Ahmad et al., 2024). Conventional Intrusion Detection Systems (IDS) and Deep Packet Inspection (DPI) solutions were designed to have lower aggregate data rates and more centralized structures. However, empirical literature and experiments have shown that beyond multi-Gbps wireline speeds, most DPI systems begin to drop packets or slow down due to the inability to achieve packet loss and packet dropping below the standard of negligible packet loss dropping at speeds greater than 1-4 Gbps without special acceleration (Diana et al., 2025).

Corresponding Author:
Section B- ELECTRICAL/COMPUTER ENGINEERING & RELATED SCIENCES
Can be cited as:
Lasisi H.O., Omodeni C.B., Aderinkola B.F., Omodeni H.B., Oladepo O., Olawuyi A.A., Adeagbo F.M., Awolufolaju T.T., and Olla M.O. (2026). Cyber Threat Detection in Network Traffic using Random Forest for Enhanced Quality of Service. FUOYE Journal of Engineering and Technology (FUOYE**JET), 11(3), 404-409.**
https://dx.doi.org/10.46792/fuoyejet.v11i3.1592

2 RELATED WORK

Given the distributed and volume nature of today's network traffic, the growing adoption of 5G, 6G, and Mobile Edge Computing (MEC) has resulted in a shift away from deep packet inspection and toward flow-level analytics and anomaly detection based on machine learning, as traditional approaches are inefficient when scaling network traffic (Satish Kumar Nadendla, 2024). As a result, anomaly detection has become an essential tool for assuring the integrity and availability of networks and services. Traditional techniques based on pre-established signatures and thresholds are becoming increasingly ineffective because they are not adaptive, do not reflect evolving traffic patterns, and cannot detect unusual but previously unknown traffic patterns or subtle traffic anomalies in real-time.

Recent research has demonstrated that machine learning is an excellent way for identifying abnormalities in a variety of networked systems, including telecommunications systems, smart grids, industrial control systems, and cyber-physical systems. Supervised learning models produced more accurate detection results, however deep learning models performed better when learning complicated attack patterns from labeled

© 2026 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)

data. In the case of cellular networks, Casas et al. (2016) used a multi-detector framework for anomaly detection based on supervised classifiers and achieved improved results. In a similar vein, Griffith et al. (2018) examined the use of supervised and unsupervised machine learning approaches for anomaly detection in smart grid systems, as well as their use to protect critical infrastructure. In recent years, anomaly detection has been a key focus of research in deep learning for industrial control systems (ICS) and cyber-physical systems (CPS). Ahmed et al. (2023) discussed the application of Multi-Layer Perceptrons and Convolutional Neural Networks (CNN) for industrial control systems, whereas Luo et al. (2021) investigated models based on ConvLSTM for anomaly detection in cyber-physical systems. All of these experiments show that machine learning is successful in detecting anomalies in a variety of networking setups.

The examined research employ a variety of machine learning algorithms and cover a wide range of application areas, but they all share some limitations. Griffith et al. (2018) identified scalability and real-time deployment as important obstacles for anomaly detection using machine learning in the smart grid. Similarly, Ahmed et al. (2023) discovered that deep learning models have yet to be thoroughly evaluated in real-world operational settings, which limits their utility in large-scale industrial control applications. Luo et al. (2021) also addressed the vulnerabilities of existing anomaly detection systems in the cyber-physical system, emphasizing the need for further research to improve their performance. The overall results indicate that machine learning approaches have enormous potential, but there are still significant problem on Quality of Service.

Moreover, the studies reviewed do not consider relation between anomaly detection and network Quality of Service (QoS). Most of the existing methods define the performance of anomaly detection based on traditional classification metrics, but do not define the effect of the detected anomalies on the network performance metrics. Consequently, network operators cannot react to service degradation in an appropriate manner based on the extent of degradation. This limitation highlights the need for anomaly detection systems that are not only accurate but also adaptive and QoS-sensitive for modern telecommunication networks. Therefore, this model is dynamically adaptable for detecting anomalous activities in contemporary telecommunication networks to enhance Quality of Service.

3 METHODOLOGY

For papers accepted for publication, it is essential that This section describes the systematic method used to develop and test the anomaly detection system. The developed model leverages CIC-IDS2017 dataset and Random Forest algorithm to accurately detect anomalous network traffic while ensuring minimal disruption to normal network operations. This research employs a quantitative experimental design. It consists of four basic steps: first, gathering data from a CIC-IDS2017 benchmark; second, preprocessing the data by cleaning and normalizing it; third, building the model using the Random Forest ensemble method; and fourth, evaluating the model using different metrics.

The system architecture comprises several processes, as indicated in Figure 1, which expose cyber risks and protect network systems. It depicts the process flow of the anomaly detection system in telecommunication networks. A decision mechanism ensures that the model is appropriately trained before proceeding to testing. The system categorizes network traffic as normal or attack, evaluates it using common network metrics, and applies those measurements. The methodology used is systematic, which aids in spotting anomalies in network traffic in an accurate and efficient way.

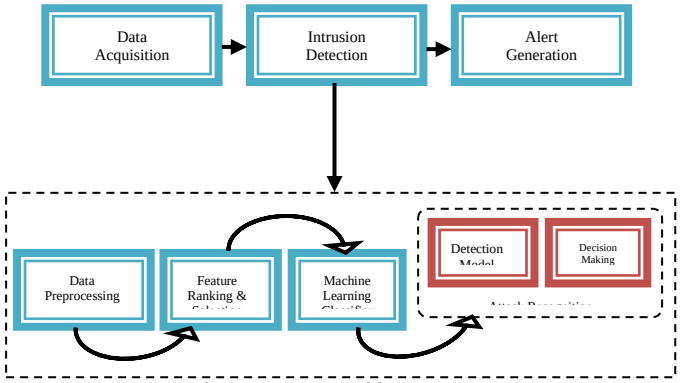


Fig. 1. System Architecture

3.1 DATA ACQUISITION AND DESCRIPTION

The CIC-IDS2017 dataset contains 2,830,541 network flow records, segmented into different attack types, such as Brute Force FTP, Brute Force SSH, Denial of Service, Heartbleed, Web Attack, Infiltration, Botnet and Distributed Denial of Service. Every flow has 78 statistical values computed with the CICFlowMeter tool, such as values related to temporal, packet size, flag based, and flow duration statistics. The data is highly imbalanced, with benign traffic accounting for ~80.3% of the data and attack traffic ~19.7% of the data, which will be used to test the model performance on imbalanced data (Sharafaldin et al., 2018). Table 1 present CIC-IDS2017 data description.

Table 1. Data Description

CIC-IDS2017 Dataset	Samples
Combined dataset	2,830,541
Training set	1,981,378
Testing set	849,163
Features	78

3.2 DATA PREPROCESSING

Network data is often the result of partial and noisy measurements taken within a wide data area. The

© 2026 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)

http://dx.doi.org/10.46792/fuoyejet.v11i3.1592 engineering.fuoye.edu.ng/journal

following is a list of stages that have been completed as preprocessing:

1. Data Cleaning: It removes unnecessary and useless data, reduces the data set, and recategorizes it for simpler interpretation and treatment.
2. Feature Encoding: One-Hot Encoding, also known as Label Encoding, is a technique for converting categorical variables like "channel" (e.g., protocol types like TCP) into numbers that may be processed by the Random Forest algorithm.
3. Data Normalization: These are normalizing features to a normal range [0,1] using Min-Max Normalization to ensure that characteristics with higher or lower numeric ranges do not give the model greater weight. Data normalization is presented in Equation (1).

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

4. Data Splitting: The split is accomplished into a Training Set (70%) for model development and a Testing Set (30%) for impartial model assessment.

3.3 RANDOM FOREST MODEL

Random Forests is an ensemble learning approach where multiple decision trees are built and the predictions from each tree are combined to determine the final classification result (Breiman, 2001). The algorithm exploits two basic ideas: bootstrap aggregation (bagging) and random selection of features, all of which contribute to the improvement of model robustness, reduction of overfitting, and an increase in model generalization ability (Liaw & Wiener, 2002).

The principle behind the Random Forest is that the aggregation of multiple weak classifiers is much more effective at making predictions than any individual classifier (Hastie et al., 2009). To get diversity from the constituent decision trees, each tree in the forest is built from a bootstrap sample with replacement from the training data. In addition, a random subset of features is used for node splitting at each node split, causing further variance between the trees and further decreasing the intercorrelation between trees. As far as this paper is concern, the anomaly detection systems design relies on an ensemble learning approach known as Random Forest, which creates many decision trees while training. Each tree in the forest is built using a data sample (bootstrapped data). At each node, the optimal split is determined by maximizing Information Gain (IG), which is calculated using Entropy (H) as presented in equation (2) and equation (3) represents Information Gain (IG):

$$H(D) = \sum_{i=1}^{c} pilog_2(pi) \quad (2)$$

$$IG(D, A) = H(D) - \sum_{v \in Values(A)} \frac{|D_v|}{|D|} H(D_v) \quad (3)$$

where H is the Entropy of D which represent the degree of impurity, D represent the dataset being considered, c is the number of classes in the data set, and pi is the proportion of the samples in class I, IG represent information gained from splitting dataset D into subgroups, and A denotes possible feature for splitting.

The forest will concatenate the predictions of all the individual B trees in the network packet x^*, for each x^* in the network. Final classification (Normal vs. Anomaly) is done using the Majority Voting, as presented in equation (4)

$$\hat{y} = mode\{T_1(x^*), T_2(x^*), \dots, T_B(x^*)\} \quad (4)$$

where $\hat{y}$ is the classification predicted by the final model, T is the tree produced by each individual tree in the random forest, and $x^*$is the new input sample for classification.

In the CIC-IDS2017 dataset, where there is a considerable class imbalance (80.3% benign, 19.7% assault), the model's $class_{weight}$ parameter was automatically set to 'balanced', which is the inverse of the class frequencies. The class weights were computed as shown in equation (5):

$$w_c = \frac{n}{c.n_c} \quad (5)$$

The weight of class c is denoted by w_c; the total number of samples is n; and the number of samples in class c is marked by n_c. This weighting method resulted in weights of 0.62 for benign traffic and 2.54 for attack traffic, ensuring that both types of traffic were fairly represented while training the model.

3.4 PERFORMANCE EVALUATION

The essential aspects of the proposed system are tested in two separate ways: classification ability and operating efficiency.

1. Classification Metrics

A Confusion Matrix of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) are used to calculate the following, as presented in equations (6) to (9):

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

$$\text{Detection Rate (Recall)} = \frac{TP}{TP+FN} \quad (8)$$

$$AUC = \int_0^1 TPR(FP\ R^{-1}(t))\ dt \quad (9)$$

2. Network Reliability and Quality of Service (QoS) Metrics

The influence of the study on network reliability is evaluated using the false alarm rate (FAR), and specificity. In addition to classification metrics, the evaluation framework incorporates quality of service (QoS) indicators that assess operational deployment characteristics such as training time, prediction throughput, prediction latency, and efficiency score (Q_{score}).

1. False Alarm Rate (FAR): The false alarm rate is used to model the system's impact on the network's QoS, which decreases as the system's QoS is better preserved, as presented in equation (10).

$$QoS\ Impact\ (FAR) = \frac{FP}{FP+TN} \quad (10)$$

2. Specificity: is the ability to correctly identify normal traffic, as shown in equation (11).

© 2026 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)

http://dx.doi.org/10.46792/fuoyejet.v11i3.1592 engineering.fuoye.edu.ng/journal

$$Specificity = \frac{TN}{TN+FP} \quad (11)$$

3. Training time: training time T_{train} is the elapsed time between the start and end of the training, which is the time spent developing the model.
4. Prediction throughput: quantifies the number of samples classified per second, as define in equation (12).

$$Throughput = \frac{n_{test}}{T_{pred}} \quad (12)$$

where n_{test} is the number of test samples and T_{pred} is the total prediction time

5. Prediction Latency: prediction latency measures the average time required to classify one sample, expressed in milliseconds, as defined in equation (13).

$$Latency = \frac{T_{pred}}{n_{test}} \text{ x } 1000 \quad (13)$$

6. Efficiency Score: For the utilization of this model, the efficiency score (Q_{score}) will be used to analyze the model's success, as shown in equation (14):

$$Q_{score} = \frac{DR}{FAR+\varepsilon} \quad (14)$$

where Q_{score} represent efficiency score, DR is the detection rate, FAR denoted False Alarm Rate, and ε is a positive contact.

A composite Q_{score} is defined in order to assess the security effectiveness and the network availability. Its goal is to prevent type II errors (also known as False Negatives) as undetected attacks are a major threat to security, and to minimize type I errors (also known as False Positives) to ensure that network operations and user experience are not unnecessarily disturbed. A higher Q_{score} translates to a more "see-through" network, with threats being detected.

4 RESULTS AND DISCUSSION

4.1 CLASSIFICATION PERFORMANCE

The Experimental results show that the Random Forest algorithm outperforms other classifiers in detecting modern network threats using the CIC-IDS2017 dataset. The system was able to categorize at a high rate while meeting the operational needs of telecommunications infrastructures. The classification results obtained by applying the Random Forest approach to the CIC-IDS2017 dataset were exceptional. The model shows the system's capacity to effectively classify both benign and hostile traffic, which is a solid indicator of its effectiveness. The model attained a high level of accuracy, demonstrating its ability to discern between benign and malicious communications. The Random Forest's ability to handle the CIC-IDS2017 dataset's high-dimensional feature space (d = 78) contributes to the high Detection Rate.

4.2 FEATURE DISTRIBUTION

Based on the feature distribution shown in Figure 2, the source port initiates network communication and classifies different types of data. Map traffic to specific services and destination ports. The protocol identifies which sort of communication is dominant and detects aberrant activity. Flow duration refers to the amount of time that two network endpoints communicate. Total forward packets: This number represents the total number of packets in a network flow sent from the source to the destination. The total number of packets sent from the destination to the source is represented by the total backward packets.

In a network flow, the total length of forward packets is the sum of all packets received from the source to the destination, as well as all packets received from the destination to the source in reverse. The maximum forward packet length is the maximum packet size that is sent from the source to the destination in a flow, while the minimum forward packet length is the minimum packet size. The forward packet length mean is the typical size of a packet, and the maximum forward packet length is the largest packet size delivered from source to destination in a flow.

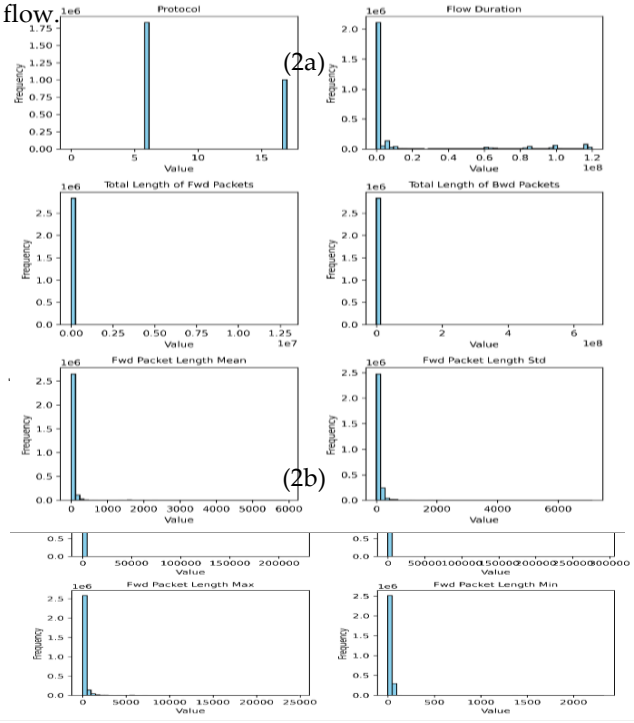


Fig. 2a & 2b. Feature Distribution

4.3 ATTACK VS BENIGN TRAFFIC DISTRIBUTION

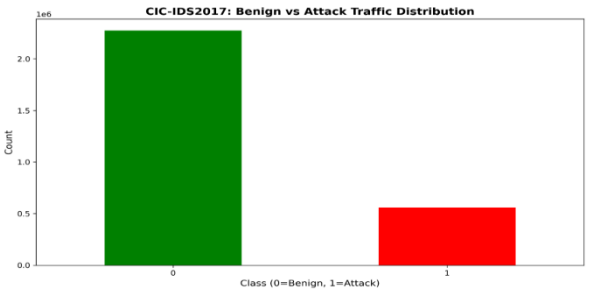


Figure 3: Benign vs Attack Traffic Distribution

This visualization compares the percentages of benign traffic (regular traffic) and attack traffic (malicious traffic) for the CIC-IDS17 dataset, as shown in Figure 3. The

© 2026 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)

http://dx.doi.org/10.46792/fuoyejet.v11i3.1592 engineering.fuoye.edu.ng/journal

results of the exploratory study are as follows: Shape (3119345, 85), Benign: 2,273,097 (72.9%), Attack: 846,248 (27.1%).

The traffic distribution is highly imbalanced, as usual in real telecommunication networks:

1. Benign Traffic: 2,273,097 samples (72.9%).
2. Attack Traffic: 846,248 samples (27.1%).

4.4 ATTACK TYPES DISTRIBUTION

Figure 4 illustrates the distribution of attack types in the CIC-IDS17 dataset. DoS Hulk, PortScan, DDoS, DoS GoldenEye, FTP-Patator, SSH-Patator, DoS Slowloris, DoS Slowhttptest, Bot, Brute Force (Web), Cross-Site Scripting (XSS), SQL Injection, and Heartbleed are examples of attacks, as presented in Table 2.

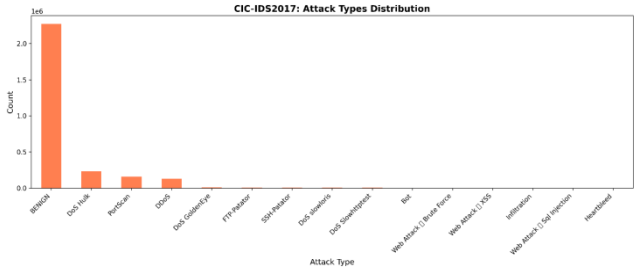


Fig. 4. CIC-IDS2017 Attack Types Distribution

Table 2. Distribution and Classification of Attack Profiles

Attack Category	Examples	Model Performance Impact
High Volume	Dos Hulk, DDoS	Easily detected via flow duration and packet count
Authentication	SSH/FTP Patator	Identified through repeated login-style traffic patterns
Web-Based	XSS, SQL Injection	Detected via specific packet length and header analysis
Anomalous	Botnet, PortScan	Detected through irregular port communication frequencies

The CIC-IDS2017 dataset has a wide range of attack styles, which can be used to test the model's capacity to detect attacks in general. The Random Forest classifier with balanced class weighting was evaluated using 849,163 samples from the CIC-IDS2017 test set. Table 3 presents the performance metrics for the model.

4.5 PERFORMANCE EVALUATION RESULTS

1. Performance Metrics

The results demonstrated that the Random Forest algorithm is ideal for network traffic due to its high detection sensitivity and low mistake rate. The model's accuracy demonstrated that it can handle content categorization with great precision, achieving 99.62% accuracy when discriminating between benign and malicious data. The system has a Detection Rate of 99.91%, indicating that it could identify nearly all security risks, as shown in Table 3. The system's high Detection Rate (99.91%) and low FAR (0.45%) allow it to strike a balance between security and availability, resulting in the highest level of security without interfering network traffic. The high ROC-AUC (0.9997) shows that the model was able to identify features having extremely high information gain and hence the trees were able to perfectly classify the attacks as bad or benign traffic as well. The ensemble voting was able to correct individual tree errors, yielding a Precision of 98.21% and only 3,052 False Positives out of nearly 700,000 benign samples.

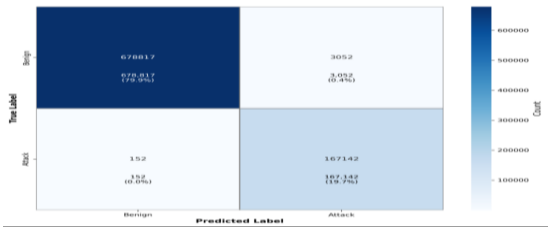
Table 3. Performance Metrics for the Model

Metric	Result
Accuracy	99.62%
Precision	98.21
Detection Rate (Recall)	99.91%
Specificity	99.55%
False Alarm Rate	0.45%
ROC-AUC	0.9997

2. Confusion Matrix Analysis

The confusion matrix of the test set utilized in CIC-IDS2017, displayed in Figure 5, illustrates the granular classification performance. Out of 681,869 benign traffic instances, 678,817 (99.55%) were correctly classified, with only 3,052 (0.45%) misclassified as attacks. Among 167,294 attack instances, 167,142 (99.91%) were accurately detected, with merely 152 (0.09%) missed (false negatives). The fact that this detection can be done despite the fact that there is a large class imbalance in the dataset, and how well the class weighting approach can overcome the bias towards the majority class is truly remarkable.

Fig. 5. Confusion Matrix



3. Network Reliability

Specificity and the False Alarm Rate (FAR) are used to determine how the system affects the network's reliability.

1. **Specificity:** The specificity of 99.55% indicates that the algorithm is quite accurate in classifying "Normal" traffic. In the telecommunications context, it ensures that valid users' data is not erroneously identified.
2. **False Alarm Rate:** The low False Alarm Rate (0.45%) is a significant finding. If a high FAR is obtained, the network becomes congested, resulting in "packet drops" for genuine users. The suggested solution will keep FAR below 0.5%, ensuring network availability and resilience, both of which are critical components in achieving network reliability. The method is incredibly efficient, with only 4.5 out of every 1000 packets being incorrectly disrupted.

4. Quality of Service Metrics

The quality of service indicators is presented in Table 4

Table 4. QoS Indicators

Metric	CIC-IDS2017
Training samples	1,981,378
Training time	486.50 seconds (8.1 min)
Prediction Throughput	42,256 samples/sec
Prediction Latency	0.024 ms/sample
Q_{score}	221.97

© 2026 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)

The proposed system was evaluated using the CIC-IDS2017 data set and received a Q_{score} of 221.97. The model can process a large number of network traffic samples per second, with a predicted throughput of 42,256 samples per second. A prediction latency of 0.024ms per sample indicates that traffic samples are identified almost instantaneously. This number represents the proportion of non-legitimate packets that, in such a congested telecommunications system, would be evaluated by security agents for every 1,000 packets sent. This means that the Random Forest model is a powerful security tool capable of meeting stringent Quality of Service (QoS) standards, which are critical in today's network environments.

5 CONCLUSION AND RECOMMENDATION

This study successfully investigated the synergy between protecting network efficiency and adopting machine learning for network security. The model achieved a Detection Rate of 99.91%, indicating that it performed well at detecting malicious traffic. The system's False Negative count of 152 in approximately 200,000 attack samples demonstrates a good level of protection against a wide range of threats. The study aimed to demonstrate that creating a secure network without losing operation is possible. The low False Alarm Rate of 0.45% suggests reduced disruption and enhancing service quality. The model processed nearly 2 million training samples in 486.50 seconds which suggest potential suitability for telecom-related intrusion detection environment.

The limitations of the study include reliance on a single benchmark dataset, lack of real-time deployment testing, limited evidence of direct QoS improvement, and the need for validation under live telecom traffic conditions. Future work should include designing and deploying testbeds to record real-time network traffic, measuring end-to-end detection latency, including packet capture and preprocessing overheads, and evaluating system performance under realistic network traffic volumes, including peak-hour traffic loads.

REFERENCES

Ahmad, I. A. I., Osasona, F., Dawodu, S. O., Obi, O. C., Anyanwu, a. C., &Onwusinkwue, S. (2024). Emerging 5G technology: A review of its far-reaching implications for communication and security. *World Journal of Advanced Research and Reviews*, *21*(1), 2474–2486. https://doi.org/10.30574/wjarr.2024.21.1.0346

Ahmed, S. F., Alam, Md. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., Mofijur, M., Shawkat Ali, A. B. M., &Gandomi, A. H. (2023). Deep learning modelling techniques: Current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, *56*(11), 13521–13617. https://doi.org/10.1007/s10462-023-10466-8

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Casas, P., Fiadino, P., & D'Alconzo, A. (2016). Machine-Learning Based Approaches for Anomaly Detection and Classification in Cellular Networks. In proceedings of the 8th International Workshop on Traffic Monitoring and Analysis.

Diana, L., Dini, P., &Paolini, D. (2025). Overview on Intrusion Detection Systems for Computers Networking Security. *Computers*, *14*(3), 87. https://doi.org/10.3390/computers14030087

Griffith, M., Almutairi, A., &Jaradat, R. (2018). Smart grid anomaly detection using machine learning approaches. In Smart Grid Research (pp. 235–249). Springer.

Haque, A. K. M. B., Zihad, M. O. M., & Hasan, M. R. (2023). 5G and Internet of Things—Integration Trends, Opportunities, and Future Research Avenues. In *5G and Beyond* (pp. 217–245). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-3668-7_11

Hastie, T., Tibshirani, R. and Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference and Prediction (2nd Edition). Springer. https://doi.org/10.1007/978-0-387-84858-7

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R News, 2(3), 18–22.

Luo, Y., Xiao, Y., Cheng, L., Peng, G., & Yao, D. D. (2021). Deep learning-based anomaly detection in cyber-physical systems: Progress and opportunities. ACM Computing Surveys.

Sarah, A., Nencioni, G., & Khan, Md. M. I. (2023). Resource Allocation in Multi-access Edge Computing for 5G and-beyond networks. Computer Networks, 227, 109720. https://doi.org/10.1016/j.comnet.2023.109720

Satish Kumar Nadendla. (2024). AI-powered deep packet inspection for secure and efficient traffic analysis in 6G programmable networks. *World Journal of Advanced Research and Reviews*, *22*(1), 2054–2071. https://doi.org/10.30574/wjarr.2024.22.1.1030

Sharafaldin, I., Ghorbani, A. A., and Lashkari, A. H. (2018). Toward creation of a new intrusion detection dataset and intrusion traffic characterization. In Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP) (pp. 108–116). SciTePress. https://doi.org/10.5220/0006639801080116

Trilles, S., Hammad, S. S., & Iskandaryan, D. (2024). Anomaly detection based on Artificial Intelligence of Things: A Systematic Literature Mapping. *Internet of Things*, *25*, 101063. https://doi.org/10.1016/j.iot.2024.101063

© 2026 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)