Enhancing Intrusion Detection and Prevention Systems: A Deep Dive into Hybrid Learning Models

Vaneeta*[1], Sangeeta Rani[2]

[1]Research Scholar, Department of Computer Science, Punjabi University, Patiala, India
[2]Assistant Professor, Department of Computer Science, Mata Gujri College, Fatehgarh Sahib, India

Abstract:

In the modern digital landscape, protecting organizational information from cyber threats is essential. The continuous existence of complex attacks entails stronger security, comprehensive protocols, and monitoring systems. The proposed approach uses a hybrid Intrusion Detection and Prevention System (IDPS) deployed in the cloud environment, which impeccably consolidates Snort with a machine learning model such as "eXtreme Gradient Boosting – Whale Optimization Algorithm" (XGBoost-WOA), Isolation-forest, and external threat intelligence to enhance the detection capabilities. Our approach emphasizes real-time traffic to distinguish malicious and benign traffic by employing binary classification. The suggested work demonstrates a multistage detection and prevention of intrusions. For feasible analysis and functional validation of the proposed approach, experiments are being conducted in real-time, i.e., online and offline, using a dashboard to identify a wide range of attacks with an accuracy of 99.3% using XGBoost and 94% using Isolation Forest. It offers a novel perspective for the development of an intelligent, adaptive, and more effective IDPS capable of addressing evolving and complex threats in the cloud environment.

Keywords: Intrusion Detection System (IDS), Prevention System (IPS), Honey Network, Snort, Machine learning, DDoS, Port Scan, Brute Force, Network Attacks, XGBoost, Gradient Boosting, Whale Optimization, Isolation Forest, Real-time, MaxMind, AbuseIPDB, Dashboard.

1. Introduction

The evolving cyber threats pose a significant risk to organizational data and infrastructure in the era of pervasive digital connectivity [1]. With increasing complexity and frequency of network attacks such as Port Scanning, Distributed Denial of Service (DDoS), and Brute Force attacks, traditional systems often fail to provide adequate protection[2]. One of the primary obstacles lies in identifying the malicious network traffic from the excessive benign traffic[3]. Additionally, this issue is associated with the class imbalance of network traffic datasets, which includes the number of benign instances exceeding the attack instances[4].

Recent studies highlighted the challenges in transforming high-accuracy offline models into real-world operating environments[2]. Although numerous ML models outperform on existing datasets, integrating them with Snort is frequently constrained by rule compatibility and system overhead[5]. Over the recent past, DDoS attacks have grown in capacity, complexity, and purpose, posing a risk to the security and availability of resources[6]. NETSCOUT's DDoS Threat Intelligence report reveals that almost 7.9 million DDoS attacks were recorded in 2023, and the same trend continued in 2024 [7]. These attacks are commonly used to cause geopolitical disruption and are frequently used to target telecommunications and

government services. The increased frequency of these threats highlights the need for more effective and adaptive IDPS[8].

To address these gaps, our research proposes an innovative cloud-enabled hybrid IDPS by integrating Signature-based detection using Modern Honey Network (MHN)-deployed Snort [9] and Anomaly-based detection on real-time traffic using the machine learning model, i.e., XGBoost-WOA. Snort is an open-source, rule-based detection system backed by a large database, which enhances its efficiency in identifying only known attack patterns [10]. Therefore, it is recommended to use a combination of machine learning and Snort, which detects both known and unknown attacks.

The XGBoost classifier is to be used in conjunction with WOA to enhance the performance and precision for intrusion detection [11]. The critical properties of high-dimensional datasets can be handled through automatic weight optimization and feature selection using WOA. In the imbalanced and diverse datasets, WOA refines the hyperparameters of XGBoost, which leads to enhanced detection precision and reduced overfitting [12]. Using WOA and XGBoost together provides a flexible, robust, and efficient method for analyzing traffic in real-time. This model's exceptional performance in handling skewed and complicated datasets with multidimensional features led to its selection for use in network traffic analysis [13]. Beyond that, it also handles missing values, supports parallel processing, faster training, and more accurate predictions than other traditional classifiers [14]. Combining Snort and an ML model with optimization techniques shows promising results in evaluating the IDPS. This methodology also incorporates external threat intelligence, which enhances the accuracy and other performance parameters. Moreover, the system effectively manages suspicious IPs by leveraging geolocation data and the reputation score of suspected IPs using AbuseIPDB [15] and MaxMind[16] to strengthen the detection procedure.

The hybrid approach begins with MHN-Snort, a Signature-based detection, subsequently followed by the WOA-XGBoost model for anomaly-based classification, and concluding with alert prioritization using external threat intelligence. Thus, the proposed approach involves a multi-stage detection and prevention process. Finally, a monitoring tool, i.e., a dashboard that visualizes traffic with signature IDs, predicts data based on the model, and provides high-risk notifications in real-time, thereby facilitating analysis and prevention of attacks. A novel approach that incorporates Signature-based, anomaly-based, and reputation models that can be developed to address the growing need for a secure and intelligent IDPS. The proposed approach enhances accuracy, reduces false alarms, and makes dynamic predictions by integrating the XGBoost model and WOA features.

The primary contributions of this research include:

(1) The integration of a hybrid approach with a centralized trust and reputation system within an IDPS.

(2) WOA application for automated feature selection and hyperparameter tuning of the XGBoost model.

(3) The validation of the approach using a dashboard in a simulated cloud-based environment using real-time traffic enriched with threat intelligence.

The remainder of the paper is structured as follows: Section 2 covers the literature review, Section 3 outlines the proposed methodology, Section 4 presents the experimental setup, Section 5 covers the results, and Section 6 includes the conclusion with future research directions.

2. Literature Review

The latest innovations of intrusion detection and prevention systems (IDPS) increasingly emphasize hybrid approaches that integrate signature-based with machine learning (ML) techniques for enhancing detection efficiency, especially when confronting evolving and complicated cyberattacks. This section covers recent studies in this field and contrasts their approaches and findings with our proposed hybrid approach.

Amouri et al. [17] presented a hybrid approach to IDS that effectively integrates KAN with the XGBoost algorithm. KANs use learnable activation functions to represent complex links in data, whereas XGBoost has been recognized for its extraordinary performance in classification functions. This ensemble methodology attained over 99% accuracy, precision, recall, and F1-scores inside IoT instances, demonstrating its effectiveness in managing dynamic and complex network traffic.

Akif et al. [18] proposed a hybrid approach that integrates Random Forest, XGBoost, K-Nearest Neighbours, and AdaBoost to form a voting-based ensemble classifier. The approach uses the IoT-23 dataset; the model attained outstanding performance, with precision, accuracy, recall, and F1-scores over 99% in binary classification.

Bamber et al. [19] conducted a comparative analysis of various deep learning classifiers, including LSTM, ANN, CNN-LSTM, BiLSTM, GRU, and BiGRU, and implemented Recursive Feature Elimination (RFE) with a Decision Tree (DT) classifier for feature selection. The CNN_LSTM exhibited exceptional performance on the NSL-KDD dataset, attaining an accuracy of 95%, a recall of 89%, and an F1-score of 94%. Despite demonstrating competence in offline classification, their approach was found to be insufficient for anomaly detection functionalities and real-time alert processing.

The AutoIDS model was suggested by Gharib et al. [20]. It is an unsupervised intrusion detection system that employs autoencoders to identify the fundamental patterns of network traffic. The model was designed to recognize zero-day attacks by exclusively training the model on benign data and analyzing reconstructive errors during testing. When the reconstructed error of an input surpasses a predetermined threshold, events are classified as intrusions. The model doesn't rely on labeled datasets, which can sometimes be inadequate or outdated, and it allows for flexibility in changing situations. The study showed that AutoIDS works well with the NSL-KDD and UNSW-NB15 datasets. It had accuracy that was as good as or better than existing ML models and very few false positives.

Song et al. [11] proposed an IDS by combining the WOA with XGBoost. The KDD Cup 99 dataset was utilized for the experiment. The author employed Principal Component Analysis (PCA) for dimensionality reduction and to simplify the process of feature selection. Employing WOA to optimize feature subsets and modifying the XGBoost hyperparameters to enhance the detection efficiency and accuracy can improve the model.

Modi and Patel [21] illustrated a security system specifically designed for a Cloud environment. The system uses Snort with Bayesian, Decision Tree, and Associative classifiers for the detection and prevention of intrusion. The system deploys IDS sensors on each host computer to detect coordinated attacks, thereby enabling distributed monitoring and correlating alerts over different cloud regions. The effectiveness of the system in detecting complex attacks was demonstrated through real-time traffic analysis and offline simulations. The approach highlights the significance of distributed detection and multi-classifier integration in enhancing the functionality of IDS in a dynamic environment.

Pramudya et al. [22] proposed a signature-based detection system by using Snort to protect network servers against attacks like DoS and network scanning attacks. The MIT-DARPA 1999 dataset was employed to evaluate the system. The system analysed over 1.2 million packets and operated at a speed of 83,494 packets per second. The results showed an accuracy of 98.10% and a 100% true positive rate, indicating the effectiveness of Snort in detecting known attack patterns in static datasets.

Devan et al. [23] suggested a hybrid detection system that combines a Deep Neural Network (DNN) for feature selection and XGBoost for classification. The system was trained using the Adam optimizer, which included normalization, feature selection, and classification. It was implemented using Python and TensorFlow, and evaluated on the NSL-KDD dataset. The model outperformed traditional ML algorithms such as logistic regression, Naive Bayes, and SVM in detecting intrusion.

Abdulganiyu et al. [24] Introduced a multi-model architecture, CWFLAM-VAE, to address the class imbalance issue in NIDS. The framework combines XGBoost, Variational Autoencoder(VAE), and Class-Wise Focal Loss to generate attack samples while preserving the original features distribution. This approach improves the performance, especially for the least common attacks. NSL-KDD and CSE-CIC-IDS2018 datasets were used for the evaluation of the model. The model outperformed standard resampling algorithms like SMOTE, ADASYN, ROS, and RUS, and classifiers like KNN, SVM, and CNN due to the majority of benign traffic and significant class imbalance. The model achieved low false positive rates of 0.17 and 0.27 with F-scores of 97.6% and 98.1%, indicating its robustness in real-time anomaly detection tasks where accurate classification of rare attacks is critical.

We identify research gaps as mentioned in Table 1. Existing approaches [18]–[24] mainly focus on either signature-based detection or machine learning models independently, lacking integration that leverages both methods for enhanced detection. Signature-based systems such as Snort [21], [22] effectively detect known attacks but fail to identify unknown or zero-day intrusions, which are addressed partially by anomaly detection models in [11], [18]–[20], [23], [24]. However, most anomaly-based proposals [11], [18]–[20], [23], [24] do not incorporate real-time processing or threat intelligence integration, limiting their practical deployment in dynamic environments like the Cloud. Furthermore, proposals like [11], [18], [19] optimize machine learning classifiers (e.g., WOA-XGBoost, CNN-LSTM) [14] but often do not address high-volume traffic inspection efficiency or deployment scalability across distributed virtual machines, leading to potential packet loss or sensor overload. Additionally, the lack of correlation between alerts from multiple sensors in distributed setups [18], [19], [24] hinders the detection of coordinated or distributed attacks. Ultimately, hybrid IDPS frameworks are required, which integrate signature-based detection with optimized ML models and enriched threat intelligence, providing robust real-time detection and prevention capabilities while minimizing false positives and resistance to sensor compromise. Our proposed hybrid approach addresses these gaps by integrating Snort with WOA-XGBoost and Isolation Forest models, enriched with IP and geolocation threat intelligence, deployed in a cloud VM environment to enable scalable, accurate, real-time intrusion detection and prevention.

Table 1: Review and Comparative Analysis of Different Approaches

Sr. no.	Ref	Features	Machine learning classifier	Dataset	Signature-based (Snort)	Anomaly Detection	Real-time Analysis	Threat Intelligence	Deployment Scenario	Prevention	Accuracy	Research Gaps
1.	[17]	KAN + XGBoost	KAN+XGBoost	N-BaIoT	No	Yes	No	No	IoT Environment	No	99%	Not implemented in real-time traffic
2.	[18]	KNN+ XGBoost+ Random Forest	Voting-based hybrid classifier	IoT-23	No	No	No	No	IoT Environment	No	99.9%	Anomaly detection on real-time traffic
3.	[19]	Deep Learning	CNN-LSTM	NSL-KDD	No	Yes	Yes	No	Network	No	95%	Inefficiency in capturing temporal patterns
4.	[20]	Autoencoder -Based IDS	Semi-Supervised AutoIDS	NSL-KDD	No	Yes	No	No	General Network	No	90%	Lack of anomaly detection
5.	[11]	WOA-XGBoost	WOA+ XGBoost	KDD CUP 99	No	Yes	No	No	General Network	No	99%	Lack of real-time implementation
6.	[21]	Virtual Network IDS	Snort with Bayesian, Associative And Decision Tree	NSL KDD, KDD CUP 99	Yes	Yes	Yes	No	Cloud	No	99.3%	Weak IDS in the cloud virtualization layer
7.	[22]	Signature-Based IDS	No	MIT-DARPA	Yes	No	No	No	Network Servers	Yes	98.1%	Need for basic signature detection on servers
8.	[23]	Hybrid Model	XGBoost +DNN	NSL-KDD	No	Yes	Yes	No	General Network	No	97%	Inadequate detection of multiclass classification
9.	[24]	Multi-Model IDS including XGBoost	CWFLAM-VAE	NSL-KDD, CSE-CIC-IDS2018	No	Yes	Yes	No	Network	No	99.1%	The computation time is considerable.
10.	Our Approach	Hybrid Snort + ML + TI + Real-time	WOA + XGBoost + Isolation Forest + Dashboard + External threat Intelligence	Yes	Yes	Yes	Yes	Yes	Cloud environment	Yes	Static dataset-98% Real-time – 99-100% (fluctuates)	Multiclass classification still needs improvement.

3. Proposed Methodology

The proposed methodology constitutes a hybrid approach to IDPS. Initially, Snort [25] is used to constantly monitor the real-time traffic and compare the incoming traffic with predefined rules. However, it covers only known attacks, and its accuracy is limited. Consequently, Snort can be incorporated to enhance the performance of the proposed approach with a machine learning algorithm and external threat intelligence, enabling dynamic and perspective responses to various threats. The WOA-XGBoost model and Isolation forest are trained and tested on dynamic traffic. The working of the proposed approach is as shown in Figure 1.

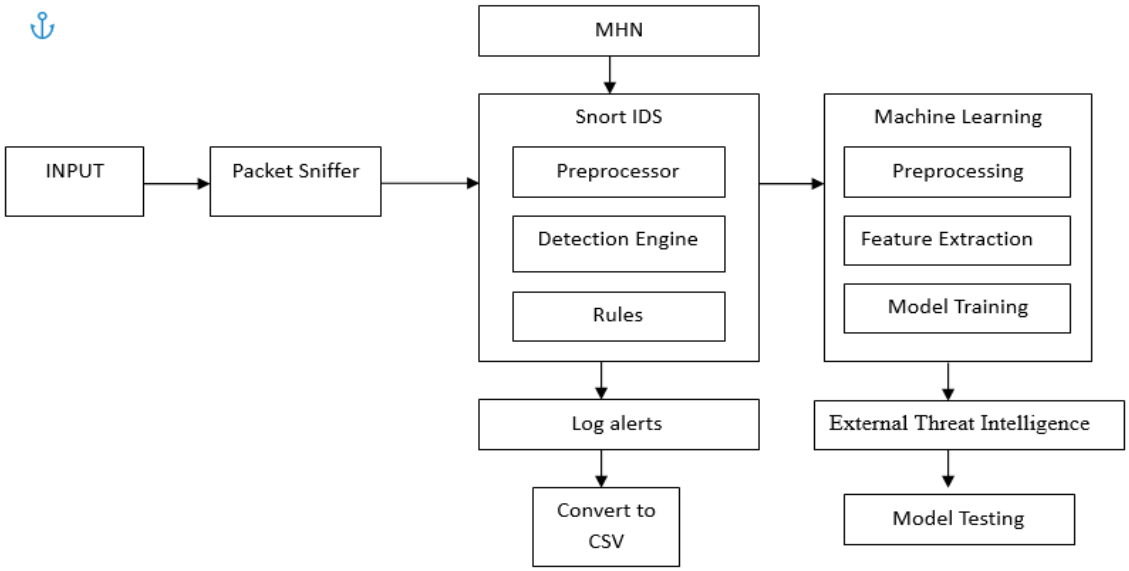


Figure 1: Proposed Methodology

The following are the steps to be involved in the hybrid approach:

1. **Dynamic Traffic Monitoring using Snort and Preprocessing**: In the proposed approach, real-time traffic is analyzed using Snort. It generates alerts stored in the CSV file, and essential features are extracted, such as Protocol, Source IP, Destination IP, Signature ID, etc. These features are forwarded to the pre-trained machine learning model to perform real-time classification to predict malicious and benign traffic.

2. **WOA for Feature Optimization:** This phase is implemented during the training phase to enhance the performance and minimize dimensionality. WOA is a metaheuristic algorithm that is influenced by nature and replicates the bubble-net foraging strategy of humpback whales to identify the most significant features[26]. This improves the computing efficiency and generalization during live prediction and ensures that the model retains only the offensive information.

3. **XGBoost for Classification:** The information derived using WOA can be used to train an XGBoost classifier. It is renowned for its scalability, excellent accuracy in tabular records. This framework helps to capture complex data and provides a strong foundation for the identification of network attacks. After the training phase, the trained model is exported to JSON format and is deployed to classify real-time traffic.

4. **Integrating External Threat Intelligence:** The system is further improved by integrating external threat intelligence, such as AbuseIPDB and MaxMind. These sources help to dynamically find geolocation, IP reputation scores, and the blacklisting status of IP addresses. The incoming traffic flagged by Snort and tested by the ML model is enriched with contextual intelligence to facilitate trust-based filtering and enhance detection confidence.

5. **Secure and Centralized Evaluation Model:** The model aggregates decisions from Snort alerts, ML classifier, and threat intelligence score. This methodology helps prioritize the responses and compute a composite trust score for each incoming IP. It also helps make an efficient decision by blocking malicious IPs, filtering suspicious activities, and sending alerts to the administrator.

6. **Interactive and Dynamic Dashboard:** This phase shows the predictions and decisions on a real-time dashboard interface. It provides information about current network conditions to the administrator about the detected attacks, including the status of the reputation score and the country from which the attack originates.

Figure 2 illustrates the proposed approach's comprehensive architecture.

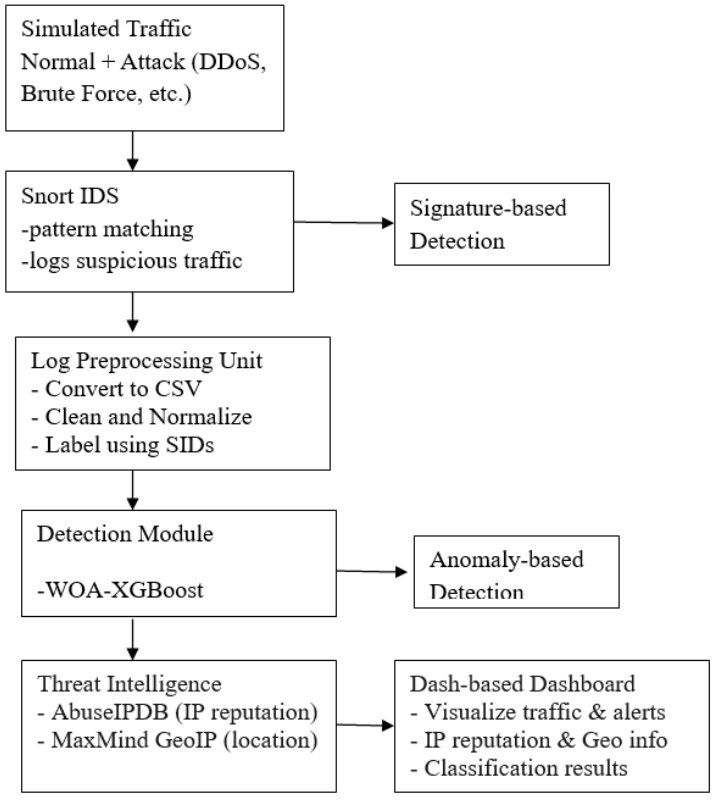


Figure 2: Hybrid Approach

4. Experimental Setup

The proposed IDPS's performance can be evaluated using Oracle VM VirtualBox, and the configuration consists of multiple virtual machines, as shown in Figure 3. Snort is deployed using MHN and acts as a primary intrusion detection engine. The system tests real-time and simulated traffic, including various types of attacks such as DDoS, Port Scanning, and Brute Force that can be generated using tools such as

Hping3, Scapy, Hydra, and Xenmap. Machine learning models such as WOA-XGBoost and Isolation-forest, trained on cleaned Snort logs, and integrated into the real-time detection process. OpenWrt is an open-source Linux-based firmware designed for routers and firewalls. Its purpose is to provide a secure, powerful, and customizable alternative compared to physical routers. Furthermore, AbuseIPDB and MaxMind threat intelligence sources are used to enhance the IDPS interface using the dashboard.

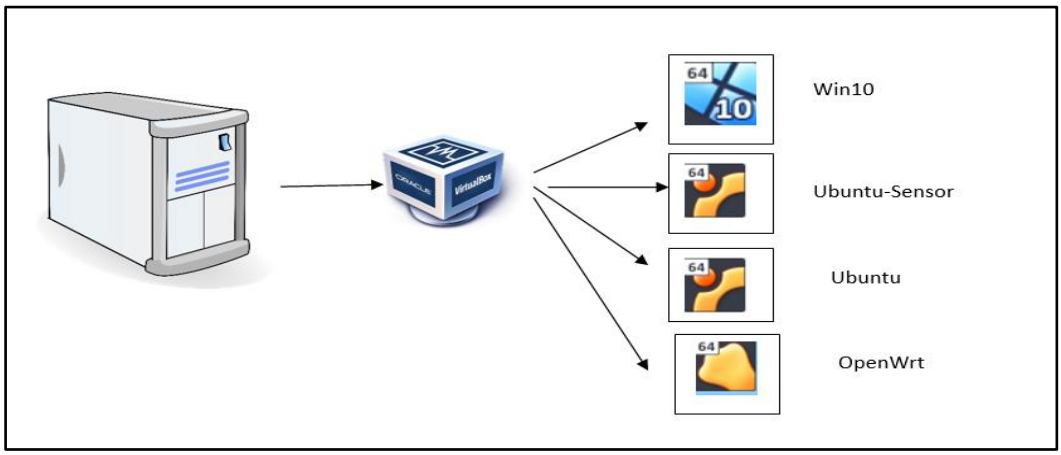


Figure 3: Oracle VM Experimental Setup

4.1. Virtual Environment Configuration

The proposed IDPS can be implemented and evaluated in a controlled virtual network environment. VMs are created to simulate different roles such as router, sensor, monitoring server, and attacker. Thus, real-time attacks are created in a simulated environment. The following are the virtual machines created to set up the simulated environment as given in Table 2.

Sr no.	VM Name	Role/Function	Tools
1.	OpenWrt	Router	-
2.	Ubuntu-Sensor	Snort, Log collection	Snort, Python, Dash, Isolation, XGBoost, MySQL, Wireshark
3.	Ubuntu-MHN	MHN Server, Attack generation	MHN, Hping3, Hydra, Nikto
4.	Window10	Xenmap, Attack generation, Testing	Scapy, Xenmap

Table 2: VM Configuration

4.2. System Architecture

The proposed approach's efficacy can be assessed using a real-time hybrid IDPS, and a virtualized testbed is set up using Virtual machines created on Oracle VM VirtualBox. The architecture is designed to integrate a Signature-based approach with machine learning classification and external threat intelligence for enhancing the security analysis. Table 3 shows the requirements for setting up the proposed approach.

Table 3: Architecture Details

Sr. no.	Requirement	System Configuration
1.	Storage	1TB
2.	RAM	16GB
3.	Processor	Intel(R) Core(TM) i5-8265U CPU @ 1.60GHz 1.80 GHz
4.	Operating system	64-bit operating system, x64-based processor

The various components of architecture and their working are shown in Figure 4.

I. **MHN Server**: It acts as a centralized server for logging and platform visualization.
II. **Snort**: Deployed on MHN to capture and analyse real-time network traffic.
III. **Machine Learning Model**: The Model is trained by extracting features from the Snort log.
IV. **Attack Generator**: Hping3, Scapy, Hydra, Xenmap, etc., are used to simulate attacks.
V. **Dashboard**: An interface for visualization of alerts, model metrics, and classification results.
VI. **External Threat Intelligence:** MaxMind and AbuseIPDB are integrated with Snort and ML to enhance their accuracy.

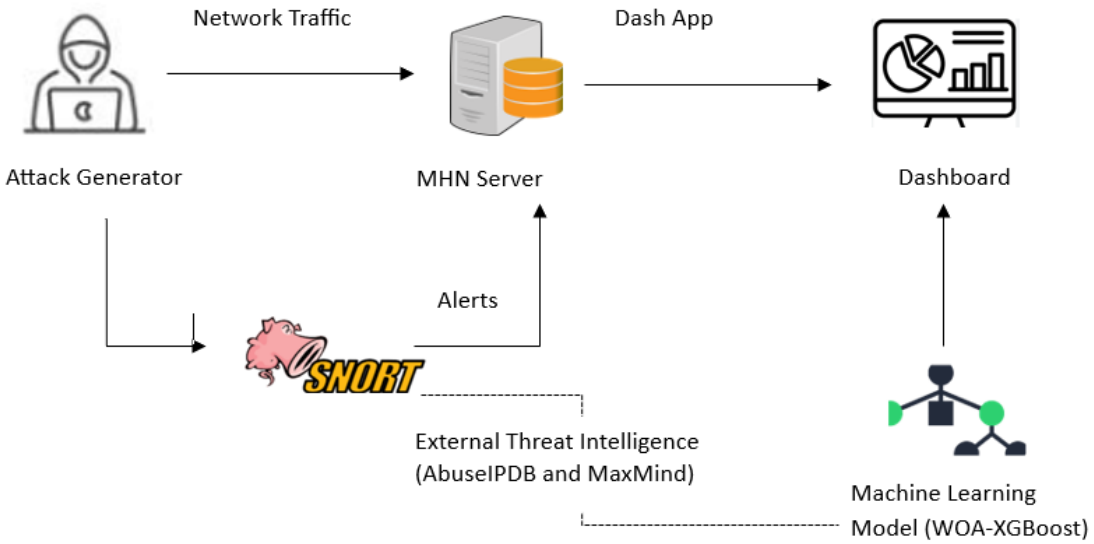


Figure 4: System Architecture

The Virtual machines are configured on the same network to allow seamless communication and traffic monitoring using bridged networking.

4.3. Attack Simulation

Three major attack types are simulated to evaluate the detection process in the proposed IDPS. The goal is to generate both normal and malicious traffic to test the system's ability to detect and classify diverse types of threats.

i. **DDoS Attacks:** using hping3, Scapy, Xenmap, etc.
ii. **Port Scanning:** using hping3, Nmap, etc.
iii. **Brute Force Attacks**: using Hydra, Xenmap.

4.4. Model Integration

The fundamental component of the proposed approach is the integration of an ML model into the Snort pipeline. The objective of this approach is to improve the detection capabilities by classifying real-time traffic as either benign or malicious based on the extracted features from the log. The following steps are to be considered for model integration.

i. **Feature Extraction and Dataset Preparation:** It is the process of extracting features from the existing dataset. The labelled dataset was generated from Snort logs based on predefined SIDs. Data cleaning is applied to the dataset, and the complete dataset is then used to train the model.

This approach is to be implemented using a Python script. Feature normalization techniques are applied when it is required to ensure the performance and stability of the model.

ii. **Model Training:** The proposed approach used two types of models: a supervised learning model, such as WOA-XGBoost, and an unsupervised anomaly detection model using Isolation-forest. The WOA-XGBoost model is trained using traffic generated from Snort logs, as depicted in Figure 5. According to binary classification, that classifies the real-time traffic into benign or specific attack types. The Isolation Forest model is used to detect unknown or zero-day attack patterns that may not match the existing Snort rules. The dataset is used to train both models and is validated using performance parameters.

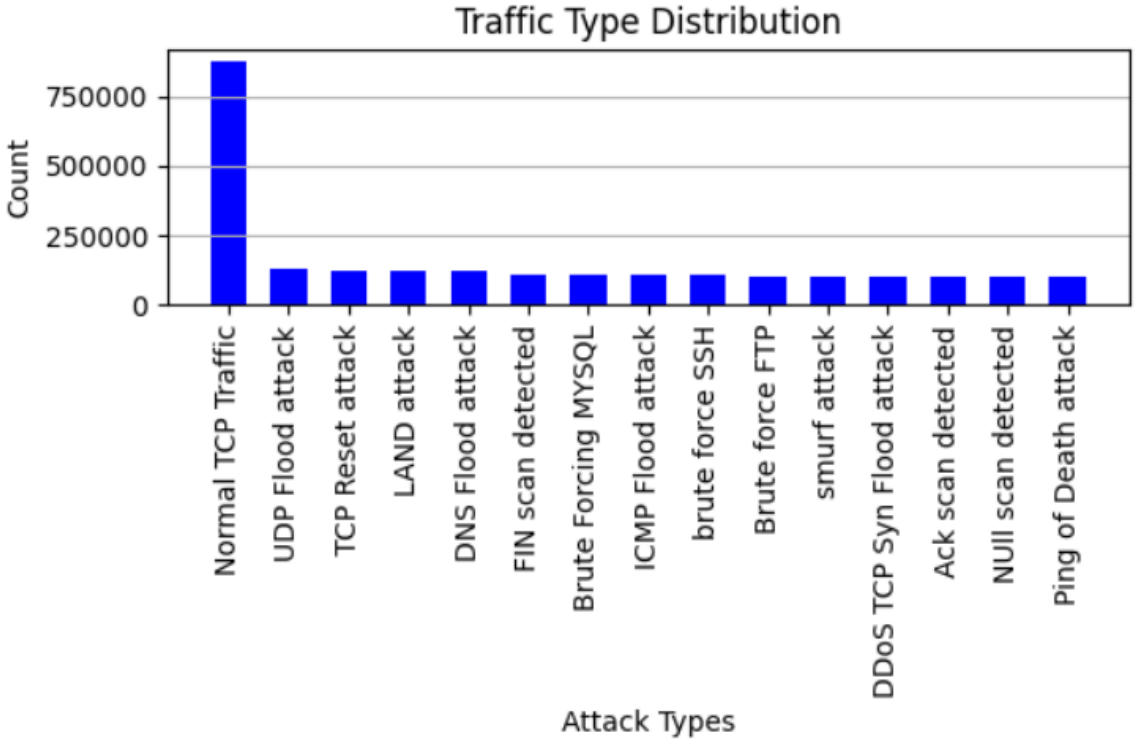
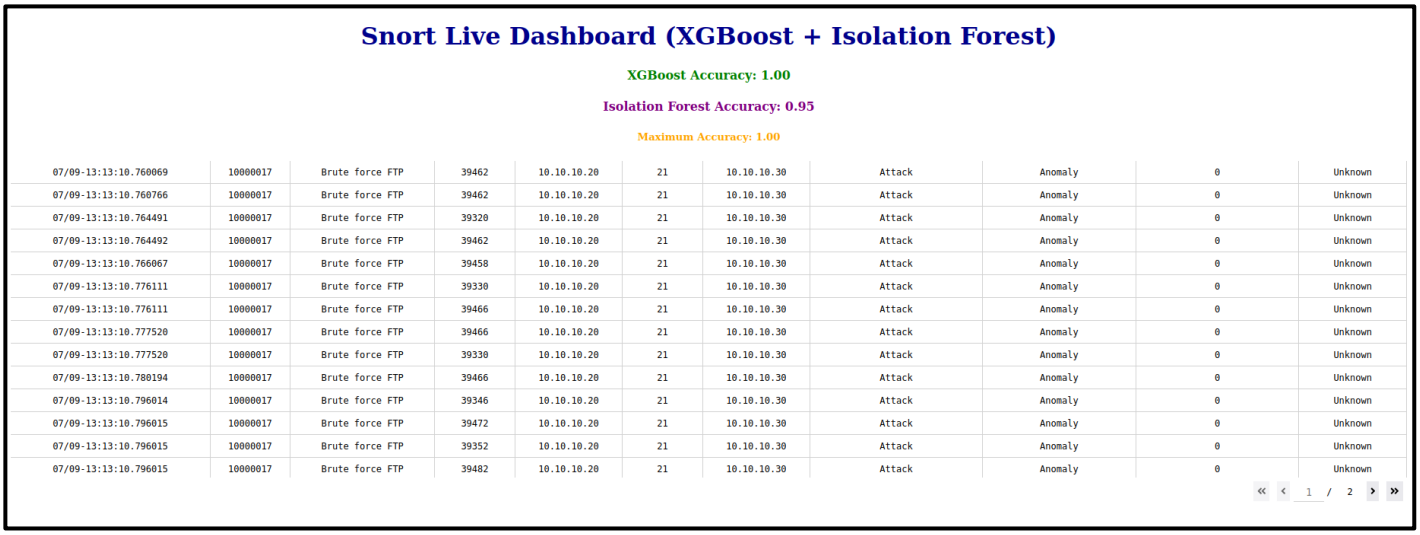


Figure 5: Training Dataset

iii. **Real-time Prediction:** The real-time prediction dashboard is intended to facilitate the dynamic analysis of network traffic that is captured by Snort. Once Snort generates the alerts based on predefined patterns, the logs are immediately processed using a Python script that extracts the required features for classification. These features are then passed to the pre-trained WOA-XGBoost model for attack classification and to Isolation Forest for anomaly detection. These trained models are evaluated in real-time, categorizing each incoming traffic and presenting the log on the dashboard with forecasts. This enables immediate threat detection with reduced latency between traffic capture and alert generation, as shown in Figure 1.



Snort Live Dashboard (XGBoost + Isolation Forest)

XGBoost Accuracy: 1.00

Isolation Forest Accuracy: 0.95

Maximum Accuracy: 1.00

07/09-13:13:10.760069	10000017	Brute force FTP	39462	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.760766	10000017	Brute force FTP	39462	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.764491	10000017	Brute force FTP	39320	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.764492	10000017	Brute force FTP	39462	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.766067	10000017	Brute force FTP	39458	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.776111	10000017	Brute force FTP	39330	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.776111	10000017	Brute force FTP	39466	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.777520	10000017	Brute force FTP	39466	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.777520	10000017	Brute force FTP	39330	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.780194	10000017	Brute force FTP	39466	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.796014	10000017	Brute force FTP	39346	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.796015	10000017	Brute force FTP	39472	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.796015	10000017	Brute force FTP	39352	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown
07/09-13:13:10.796015	10000017	Brute force FTP	39482	10.10.10.20	21	10.10.10.30	Attack	Anomaly	0	Unknown

« ‹ 1 / 2 › »

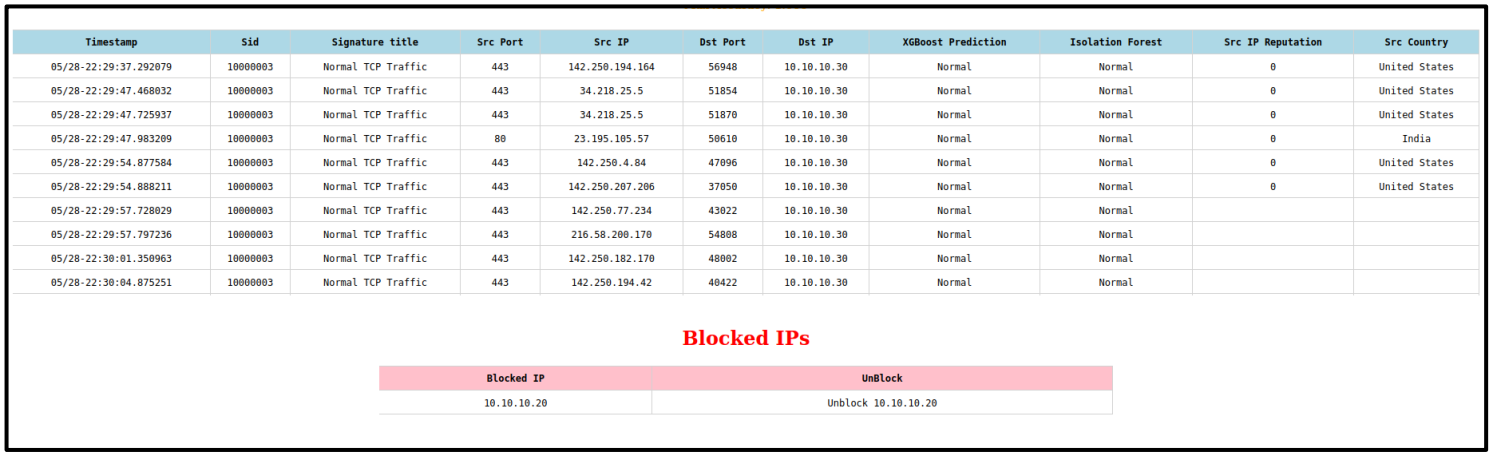
Figure 6: Real-time dashboard displaying predictions and enriched with threat intelligence

iv. **External Threat Intelligence:** The system also integrates external threat intelligence services such as MaxMind and AbuseIPDB, which allows the system to verify suspicious IP addresses and adds its reputation score and its origin of attack generation. All the results on one dashboard help the administrator to monitor and respond expeditiously, as demonstrated in Figure 6. This real-time IDPS significantly enhances the proactive detection and prevention capability.

v. **Prevention of Attacks:** Apart from identifying malicious behavior, the proposed approach includes a preventive mechanism that aims to mitigate threats before they influence the cloud environment. Dynamic blocking, firewall setup, and external threat intelligence integration help to build the preventative framework. It includes the following steps:

   a. **The real-time IP blocking:** Upon detecting a malicious attack through XGBoost or Isolation Forest, or Snort signatures, the proposed approach immediately identifies the source IP address. If the IP is determined to be malicious, it is automatically blocked using Ubuntu firewall tools such as ufw or iptables, as shown in Figure 7.

   b. **Automated Firewall Rule Enforcement:** The firewall rules are automatically updated when the proposed approach detects the malicious traffic patterns, such as DDoS, port scans, and Brute force attacks.



Timestamp	Sid	Signature title	Src Port	Src IP	Dst Port	Dst IP	XGBoost Prediction	Isolation Forest	Src IP Reputation	Src Country
05/28-22:29:37.292079	10000003	Normal TCP Traffic	443	142.250.194.164	56948	10.10.10.30	Normal	Normal	0	United States
05/28-22:29:47.468032	10000003	Normal TCP Traffic	443	34.218.25.5	51854	10.10.10.30	Normal	Normal	0	United States
05/28-22:29:47.725937	10000003	Normal TCP Traffic	443	34.218.25.5	51870	10.10.10.30	Normal	Normal	0	United States
05/28-22:29:47.983209	10000003	Normal TCP Traffic	80	23.195.105.57	50610	10.10.10.30	Normal	Normal	0	India
05/28-22:29:54.877584	10000003	Normal TCP Traffic	443	142.250.4.84	47096	10.10.10.30	Normal	Normal	0	United States
05/28-22:29:54.888211	10000003	Normal TCP Traffic	443	142.250.207.206	37050	10.10.10.30	Normal	Normal	0	United States
05/28-22:29:57.728029	10000003	Normal TCP Traffic	443	142.250.77.234	43022	10.10.10.30	Normal	Normal		
05/28-22:29:57.797236	10000003	Normal TCP Traffic	443	216.58.200.170	54808	10.10.10.30	Normal	Normal		
05/28-22:30:01.350963	10000003	Normal TCP Traffic	443	142.250.182.170	48002	10.10.10.30	Normal	Normal		
05/28-22:30:04.875251	10000003	Normal TCP Traffic	443	142.250.194.42	40422	10.10.10.30	Normal	Normal		

Blocked IPs

Blocked IP	UnBlock
10.10.10.20	Unblock 10.10.10.20

Figure 7: Blocked IPs

vi. **Model Deployment Environment:** The model is deployed on Ubuntu-sensor using Python 3.9 by using supporting libraries such as pandas, numpy, XGBoost, etc. This integration enabled the system

to transcend static rule-based detection, facilitating dynamic intelligent classification and anomaly detection.

5. Result and Analysis: This section evaluates the performance the efficacy of the proposed hybrid IDPS on real-time traffic using performance metrics.

5.1. Performance Metrics: The proposed approach is assessed using conventional classification metrics. These metrics offer a comprehensive assessment of the model's ability to classify real-time traffic as either benign or malicious. The following are the most commonly used performance parameters for evaluating the model.

i. **True positive (TP):** It describes the total number of attacks that are correctly classified as malicious.
ii. **False Positive (FP):** It shows ordinary packets that are erroneously labelled as attacks.
iii. **True Negative (TN):** It describes the total number of packets that are accurately classified as normal traffic.
iv. **False Negative (FN):** attacks that remain overlooked, such as suspicious traffic that is mistakenly assumed to be legitimate.

The following performance metrics can be computed based on the above-mentioned parameters in a real-time environment.

1. **Accuracy:** It is the portion of accurately classified instances, including both attacks and normal traffic of the total amount of traffic.

$$Accuracy = \frac{(\mathrm{TP} + \mathrm{TN})}{(\mathrm{FP} + \mathrm{TP} + \mathrm{FN} + \mathrm{TN})} * 100$$

2. **Precision:** It is the portion of actual attacks that are genuinely malicious.

$$Precision = \frac{(\mathrm{TP})}{(\mathrm{TP} + \mathrm{FP})} * 100$$

3. **Recall or Sensitivity:** It is the percentage of actual attacks identified by the IDS that indicates how effectively it detects all the threats.

$$Sensitivity\ or\ Recall = \frac{(\mathrm{TP})}{(\mathrm{TP} + \mathrm{FN})} * 100$$

4. **F1 Score:** It balances recall and precision parameters, offering a comprehensive measure of an model's performance.

$$F1\ Score = 2 * \frac{(\mathrm{Precision} * \mathrm{Recall})}{(\mathrm{Precision} + \mathrm{Recall})} * 100$$

Table 4: Comparative analysis of different approaches

Sr. no.	Ref	Year	Approach	Dataset	Accuracy	Precision	Recall	F1 Score
1.	[17]	2024	KAN+XGBoost	N-BaIoT	99.69%	98.10%	98.01%	98.04%
2.	[18]	2025	Voting-based hybrid classifier	IoT-23	99.99%	99.99%	99.99%	99.99%
3.	[19]	2025	CNN-LSTM	NSL-KDD	95%	NA	89%	94%

4.	[20]	2019	Semi-Supervised AutoIDS	NSL-KDD	90.17%	90.80%	92.05%	91.42%
5.	[11]	2022	WOA+ XGBoost	KDD CUP 99	99.06%	NA	99.58%	NA
6.	[21]	2018	Snort with Bayesian, Associative, and Decision Tree	NSL KDD, KDD CUP 99	99.33%	NA	NA	NA
7.	[22]	2022	Signature-Based IDS	MIT-DARPA	98.10%	NA	NA	NA
8.	[23]	2020	XGBoost +DNN	NSL-KDD	97.60%	97%	97%	97%
9.	[24]	2025	CWFLAM-VAE	NSL-KDD, CSE-CIC-IDS2018	99.10%	99.2%	98.40%	97.60%
10.	Our Approach	2025	WOA + XGBoost + Dashboard + External threat Intelligence	Real-time traffic (Online & Offline)	99.3%	98.8%	99.4%	99.14%
11.	Our Approach	2025	Isolation Forest + Dashboard + External threat Intelligence	Real-time traffic (Online)	9%	91.23%	94%	92.6%

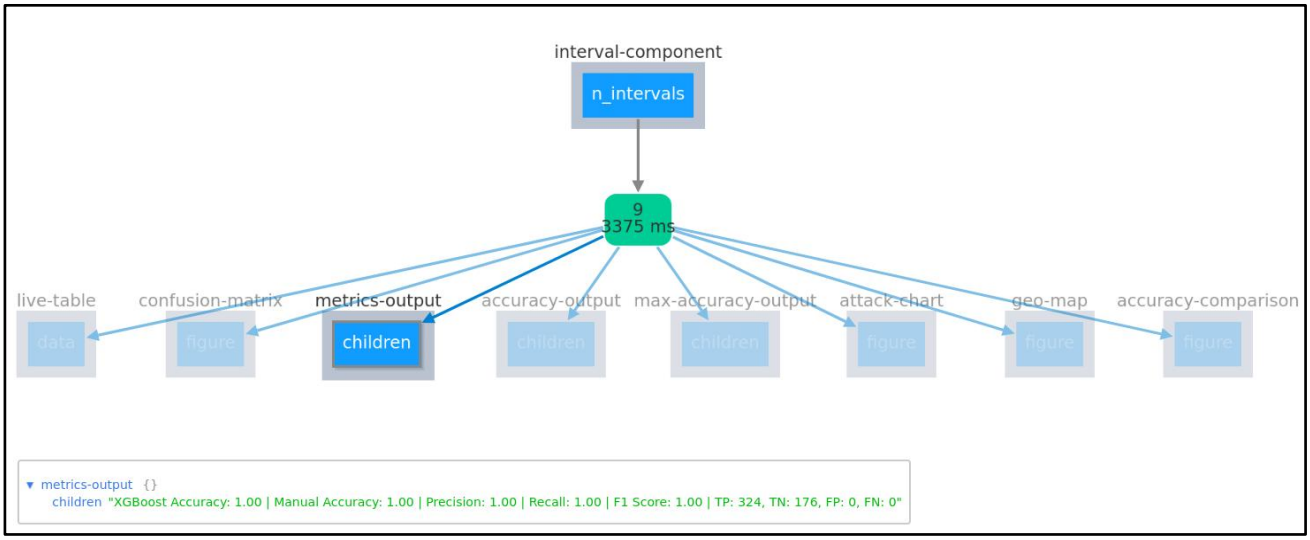


Figure 8: Results of XGBoost

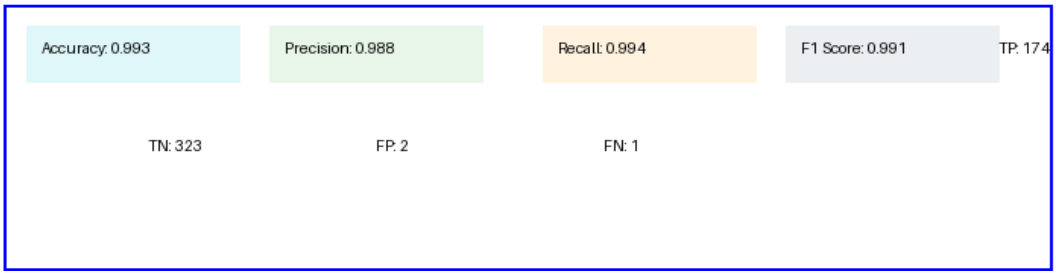


Figure 9: Performance Metrics

The result of these performance metrics, such as accuracy, recall, precision, and F1-score, is evaluated continuously and displayed on the dashboard as shown in Figure 8 based on the live traffic. The dashboard provides instant insight into the detection efficacy of the hybrid approach. The evaluation results, derived

from the most recent 500 entries from real-time traffic that are logged in a CSV file, are visualized in Figure 9. The parameters tend to fluctuate depending on traffic diversity and intensity. The variation occurs due to real-world factors such as dynamic attack strategies, packet drop ratio, and noisy traffic. However, the dashboard enables operators to observe these changes in real-time and permits uninterrupted performance evaluation for dynamic enhancement.

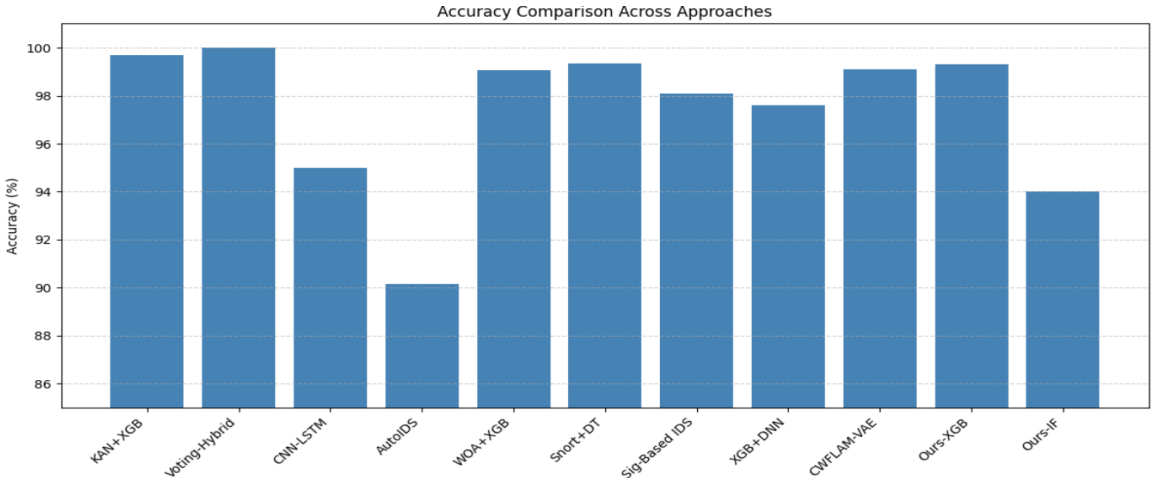


Figure 10: Accuracy Comparison

The comparison of accuracy with different detection approaches, as given in Figure 10 demonstrates that most techniques achieve high accuracy. Our proposed approach achieves an accuracy of 99.3% with the WOA-XGBoost and 94% with the Isolation forest. Thus, from the observation, it is clear that the WOA-XGBoost model outperforms other models.

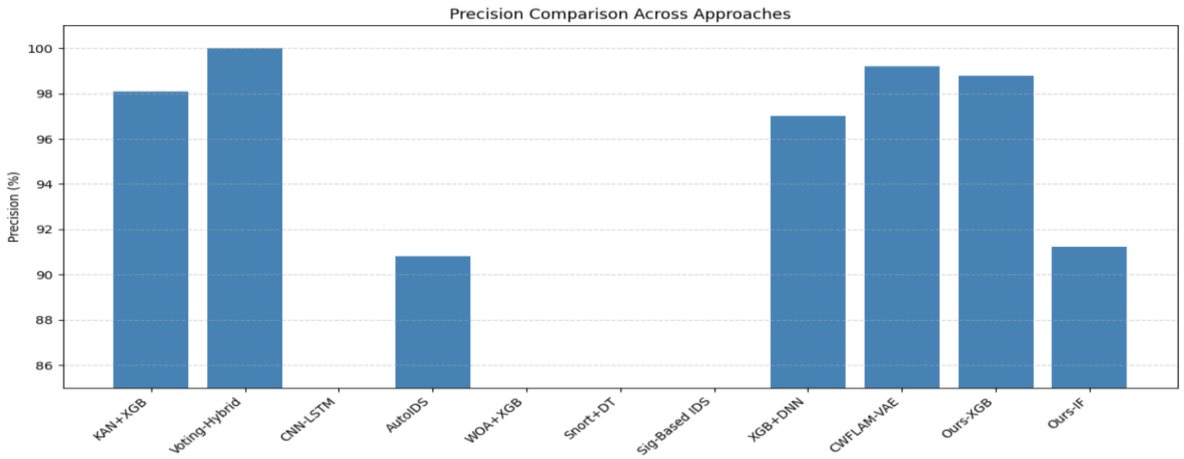


Figure 11: Precision Comparison

The Precision comparison as given in Figure 11, with different approaches, shows that our proposed approach (WOA-XGBoost) achieves a high precision of 98.8%, indicating its effectiveness in minimizing false positives, and our approach effectively detects both attacks and benign traffic.

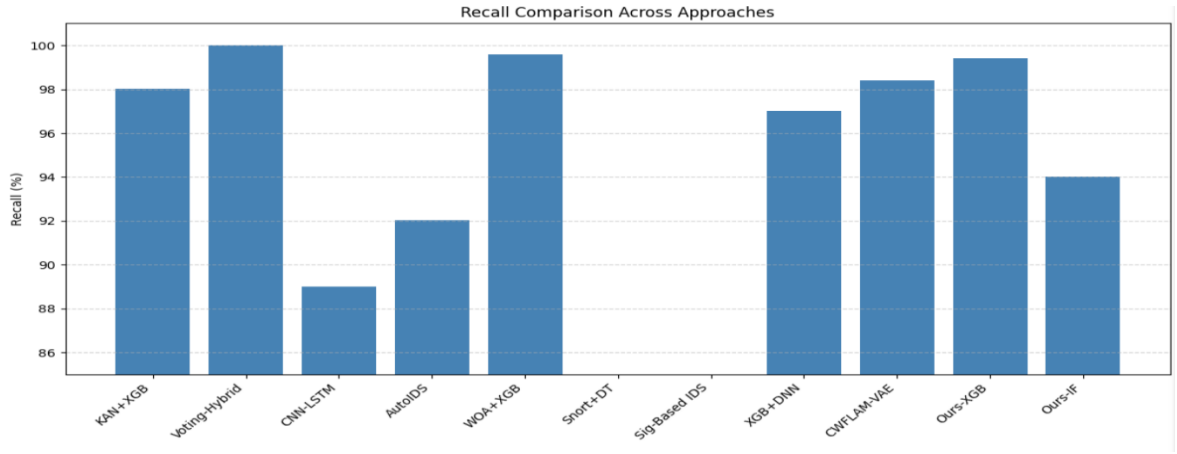


Figure 12: Recall Comparison

The Recall Comparison in Figure 12 shows that our proposed approach (WOA-XGBoost) achieves a high recall of 99.99%. It indicates that the approach effectively detects actual attacks.

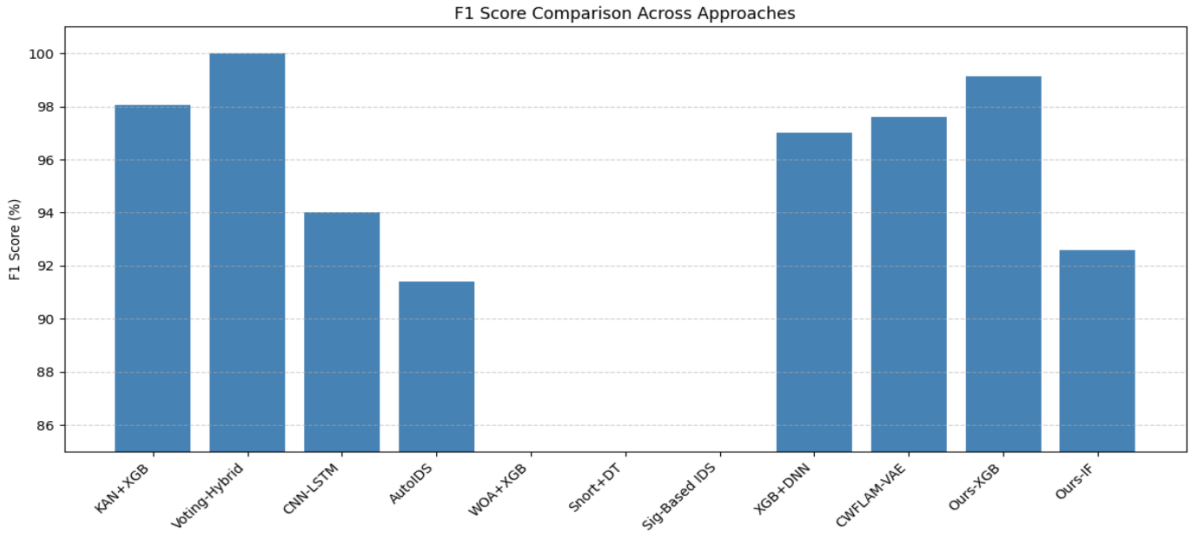


Figure 13: F1-Score Comparison

The comparison of the F1-score in Figure 13 highlights that the proposed approach achieves 99.14%, which reflects a strong balance between recall and precision. It indicates that our approach minimizes both false negatives and false positives compared to other approaches. Thus, from the above analysis, it is observed that the proposed approach (WOA-XGBoost) efficiently detects malicious and benign traffic in the cloud environment.

5.2. Confusion Matrix: It offers a comprehensive breakdown of categorization findings in context of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN)—the performance of the IDPS is shown in Figure 14. Beyond general accuracy, the confusion matrix enables one to evaluate the model's decision-making ability precisely. Based on the top 500 records from the Snort log, this study found a high-accuracy model that effectively identified a notable amount of both attack and normal traffic. While a low FN count indicates a minimal undetectable risk, a high TP score indicates that most real attacks were effectively identified. Likewise, a low FP count reduces false alarms, while a high TN count indicates the dependability of the model in identifying benign traffic. These findings together show how strong and dependable the classification model is in identifying and differentiating between illegal and authorized network traffic.

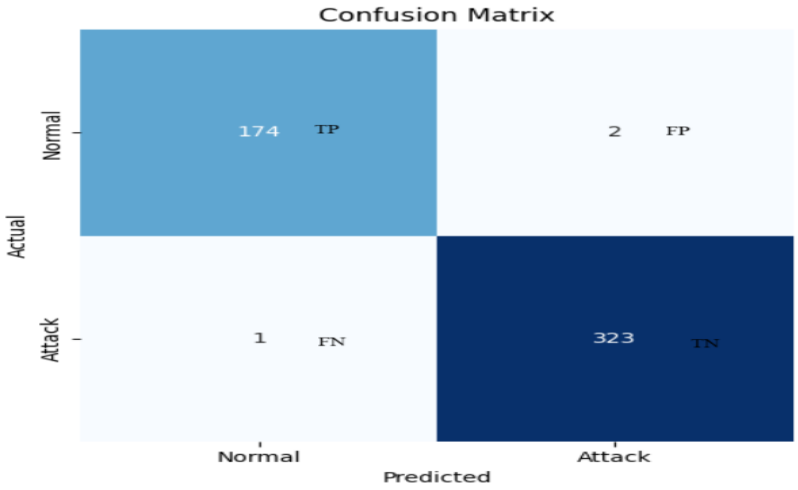


Figure 14: Confusion Matrix

5.3. Map Integration: The system integrates geolocation mapping using MaxMind GeoIP2 DB on a dashboard, as depicted in Figure 15. It also facilitates visualization of the attacker's IP on a world map by highlighting the origin of an intrusion attempt. This tool provides contextual awareness that helps to identify highly-risk region and monitor the worldwide dispersion of attacks.

Figure 15: Geolocation Map using MaxMind GeoIP2

The evaluation of the integrated Snort-ML hybrid approach demonstrated significant enhancement in the intrusion detection and prevention system. The combined approach to WOA-XGBoost and Isolation-forest improved the accuracy and minimized the false positives compared to the traditional algorithms. The inclusion of real-time validation with MaxMind and AbuseIPDB enhanced the threat intelligence.

6. Conclusion

The proposed hybrid approach implemented an IDPS that combines Snort with an ML model, including WOA-XGBoost for classification and Isolation Forest for anomaly detection in a multi-VM environment. The approach significantly enhanced Snort's detection capabilities by reducing false positives and improving the identification of both known and unknown attack patterns. The XGBoost model outperformed the Isolation Forest in the experimental evaluation, achieving an accuracy of 99.3% compared to 94%. This validates the superiority of supervised learning in identifying suspicious traffic within the proposed system. Moreover, real-time validation can be accomplished by the use of external threat intelligence, such as MaxMind and AbuseIPDB, with binary classification. Although the overall performance metrics indicate promising results, the outcome exhibits occasional fluctuations during periods of high traffic. Furthermore, the system exhibited the potential for deployment in modern cloud

or enterprise environments by maintaining real-time classification of benign traffic and attacks at a speed of 250 milliseconds.

Future work includes the implementation of the proposed approach using federated learning by using datasets from multiple machines. In real-time detection systems, there is a critical need to stabilize quick decision-making with accurate detection. While lightweight models offer quick response, they might sacrifice detection depth, necessitating hybrid or interrupted analytical strategies. This research provides a framework for societal implications by supporting scalable and efficient threat detection, and it may additionally be expanded to accommodate multiclass classification for recognizing a wider variety of cyber threats in multifaceted scenarios.

Conflict of Interest- On behalf of all authors, the corresponding author states that there is no conflict of interest.

Competing Interests- NA

Funding Information- NA

Author Contribution- Both authors contributed equally.

Data Availability Statement: The dataset analysed during this research was collected from a Snort using MHN deployed in a controlled environment. Due to cybersecurity and privacy concerns, the dataset is not publicly available.

Ethics Approval: Not required for this study.

Research Involving Human and /or Animals-NA

Informed Consent-NA

References

[1] T. Preethi, P. R. Reddy, L. Likhitha, P. P. Kumar, and A. Kamani, "A Novel Approach for Anomaly Detection using Snort Integrated with Machine Learning," *Proc. 18th INDIAcom; 2024 11th Int. Conf. Comput. Sustain. Glob. Dev. INDIACom 2024*, pp. 796–801, 2024, doi: 10.23919/INDIACom61295.2024.10498401.

[2] X. Sun, D. Zhang, H. Qin, and J. Tang, "Bridging the Last-Mile Gap in Network Security via Generating Intrusion-Specific Detection Patterns through Machine Learning," *Secur. Commun. Networks*, vol. 2022, no. Ml, 2022, doi: 10.1155/2022/3990386.

[3] M. Sajid *et al.*, "Enhancing intrusion detection: a hybrid machine and deep learning approach," *J. Cloud Comput.*, vol. 13, no. 1, 2024, doi: 10.1186/s13677-024-00685-x.

[4] A. Hajjouz and E. Y. Avksent'eva, "An approach to configuring CatBoost for advanced detection of DoS and DDoS attacks in network traffic," *Vestn. Astrakhan State Tech. Univ. Ser. Manag. Comput. Sci. informatics*, vol. 2024, no. 3, pp. 64–74, 2024, doi: 10.24143/2072-9502-2024-3-65-74.

[5] C. C. Lo, C. C. Huang, and J. Ku, "A cooperative intrusion detection system framework for cloud computing networks," *Proc. Int. Conf. Parallel Process. Work.*, pp. 280–284, 2010, doi: 10.1109/ICPPW.2010.46.

[6] S. A. Chamkar, M. Zaydi, Y. Maleh, and N. Gherabi, "Improving Threat Detection in Wazuh Using Machine Learning Techniques," *J. Cybersecurity Priv.*, vol. 5, no. 2, pp. 1–25, 2025, doi: 10.3390/jcp5020034.

[7] Microsoft Threat Intelligence, "Microsoft Digital Defense Report-Building and improving cyber resilience," *Microsoft Threat Intell.*, no. October, pp. 1–131, 2023.

[8] T. Davies, M. H. Eiza, N. Shone, and R. Lyon, "A Collaborative Intrusion Detection System Using Snort IDS Nodes," 2025, https://doi.org/10.48550/arXiv.2504.16550.

[9] Vinay Kumar Singh, "MHID: Malware Detection Using Hybrid Honeypot and Intrusion Detection System," *Commun. Appl. Nonlinear Anal.*, vol. 31, no. 2s, pp. 145–162, 2024, doi: 10.52783/cana.v31.617.

[10] V. Dhawan, S. Rani, and L. K. Joshi, “An appraisal for the identification of a novel approach to IDPS in cloud environment,” *AIP Conf. Proc.*, vol. 2916, no. 1, p. 50005, 2023, doi: 10.1063/5.0177661.

[11] Y. Song, H. Li, P. Xu, D. Liu, and S. Cheng, “A Method of Intrusion Detection Based on WOA-XGBoost Algorithm,” *Discret. Dyn. Nat. Soc.*, vol. 2022, 2022, doi: 10.1155/2022/5245622.

[12] M. Zivkovic, M. Tair, K. Venkatachalam, N. Bacanin, Š. Hubálovský, and P. Trojovský, “Novel hybrid firefly algorithm: An application to enhance XGBoost tuning for intrusion detection classification,” *PeerJ Comput. Sci.*, vol. 8, pp. 1–38, 2022, doi: 10.7717/peerj-cs.956.

[13] L. Mukhija, R. Sachdeva, and M. Singh, “Analytical Study of Virtual Machine Migration Techniques in Cloud Computing,” in *Soft Computing for Intelligent Systems*, 2021, pp. 425–433, https://doi.org/10.1007/978-981-16-1048-6_34.

[14] S. Bajpai, K. Sharma, and B. K. Chaurasia, “A Hybrid Meta-heuristics Algorithm: XGBoost-Based Approach for IDS in IoT,” *SN Comput. Sci.*, vol. 5, no. 5, p. 537, 2024, doi: 10.1007/s42979-024-02913-2.

[15] J. L. Lewis, G. F. Tambaliuc, H. S. Narman, and W.-S. Yoo, “IP Reputation Analysis of Public Databases and Machine Learning Techniques,” in *2020 International Conference on Computing, Networking and Communications (ICNC)*, 2020, pp. 181–186, doi: 10.1109/ICNC47757.2020.9049760.

[16] A. Y. Nur, “Accuracy and Coverage Analysis of IP Geolocation Databases,” *2023 Int. Balk. Conf. Commun. Networking, Balk. 2023*, no. July, 2023, doi: 10.1109/BalkanCom58402.2023.10167899.

[17] A. Amouri, “Enhancing Intrusion Detection in IoT Environments : An Advanced Ensemble Approach Using Kolmogorov-Arnold Networks.”2024 International Symposium on Networks, Computers and Communications (ISNCC), Washington DC, DC, USA, 2024, pp. 1-6, doi: 10.1109/ISNCC62547.2024.10758956.

[18] R. I. Dataset, “Hybrid Machine Learning Models for Intrusion Detection in IoT :,” pp. 1–9, 2025.2024 International Symposium on Networks, Computers and Communications (ISNCC), Washington DC, DC, USA, 2024, pp. 1-6, doi: 10.1109/ISNCC62547.2024.10758956.

[19] S. S. Bamber, A. V. R. Katkuri, S. Sharma, and M. Angurala, “A hybrid CNN-LSTM approach for intelligent cyber intrusion detection system,” *Comput. Secur.*, vol. 148, p. 104146, 2025, doi: https://doi.org/10.1016/j.cose.2024.104146.

[20] M. Gharib, B. Mohammadi, S. H. Dastgerdi, and M. Sabokrou, “AutoIDS: Auto-encoder Based Method for Intrusion Detection System,” pp. 1–9, 2019, [Online]. Available: http://arxiv.org/abs/1911.03306.

[21] C. Modi and D. Patel, “A feasible approach to intrusion detection in virtual network layer of Cloud computing,” *Sadhana - Acad. Proc. Eng. Sci.*, vol. 43, no. 7, 2018, doi: 10.1007/s12046-018-0910-2.

[22] P. B. Pramudya, “Implementation of signature-based intrusion detection system using SNORT to prevent threats in network servers,” *J. Soft Comput. Explor.*, vol. 3, no. 2, pp. 93–98, 2022, doi: 10.52465/joscex.v3i2.80.

[23] P. Devan and N. Khare, “An efficient XGBoost–DNN-based classification model for network intrusion detection system,” *Neural Comput. Appl.*, vol. 32, no. 16, pp. 12499–12514, 2020, doi: 10.1007/s00521-020-04708-x.

[24] O. H. Abdulganiyu, T. Ait Tchakoucht, A. E. H. Alaoui, and Y. K. Saheed, “Attention-driven multi-model architecture for unbalanced network traffic intrusion detection via extreme gradient boosting,” *Intell. Syst. with Appl.*, vol. 26, no. March, p. 200519, 2025, doi: 10.1016/j.iswa.2025.200519.

[25] E. Jaw and X. Wang, “A novel hybrid-based approach of snort automatic rule generator and security event correlation (SARG-SEC),” *PeerJ Comput. Sci.*, vol. 8, pp. 1–37, 2022, doi:

10.7717/PEERJ-CS.900.

[26] A. T. Siahmarzkooh and M. Alimardani, “A Novel Anomaly-based Intrusion Detection System using Whale Optimization Algorithm WOA-Based Intrusion Detection System,” *... J. Web Res.*, no. December 2021, doi: 10.22133/IJWR.2022.305467.1106.