

A Host-based Intrusion Detection: Using Signature-based and AI-driven Anomaly Detection for Enhanced Cybersecurity*

Fazalur Rehman*, Farhan Mushtaq†, Hafsa Zaman‡

*School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad, Pakistan)

†Department of Electrical Engineering, University of Engineering and Technology (UET), Taxila, Pakistan

‡School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad, Pakistan

Email: *181219@students.au.edu.pk†22MS-EE-15@students.uettaxila.edu.pk ‡hzaman.msis23seecs@seecs.edu.pk

Abstract—This research introduces a Hybrid Intrusion Detection System (HIDS) that merges signature-based detection, with AI-powered anomaly detection to enhance the accuracy and effectiveness of identifying cyber threats. The proposed HIDS demonstrates an ability to detect uncommon and sophisticated cyber threats with an accuracy rate of 90.37%. By combining Gradient Boosting and K-Nearest Neighbors (KNN) algorithms the system improves detection precision, speeds up response times, and expands coverage across network traffic. This comprehensive approach overcomes the limitations of traditional methods by enabling threat responses while reducing false positive rates. The study highlights the potential of integrating signature-based and AI-driven techniques to strengthen cybersecurity defences which emphasizes the benefits of this approach. When the system detects a potential threat, alerts are sent to the Security Operations Center (SOC) or Network Operations Center (NOC), with details such as nature of the threat, and the affected system. This study establishes a foundation for real-time cyber threat detection and intelligence sharing in cloud environments, with future Hybrid IDS versions operating across multiple hosts and supported by a web service for cross-platform compatibility and centralized alert system.

Index Terms—Host-based Intrusion Detection System (HIDS), Signature-based Detection, AI-driven Anomaly Detection, Cyber Security, Gradient Boosting, k-Nearest Neighbour (KNN), Hybrid IDS, Network Security, Intrusion Detection, Machine Learning.

I. INTRODUCTION

In today's interconnected world, keeping information safe within computer systems is essential. Cyber threats, such, as unauthorized access present dangers to the security and reliability of data. Intrusion detection systems (IDS) play a role in overseeing and pinpointing potential cyber attacks. The main goal of intrusion detection is to recognize suspicious behaviours that suggest access or possible security breaches. This is achieved by monitoring and evaluating system operations. By detecting behaviours before they develop into more severe problems, IDS act as an early warning alert system that enables quick response in order to lessen the impact of cyber attacks.

Intrusion detection systems (IDS) use predefined patterns or signatures to detect known attacks. This method has limitations, especially in spotting emerging threats without estab-

lished signatures [1]. This emphasizes the need for dependable and adaptable intrusion detection system. The threat landscape is constantly changing, encompassing attack methods such as Denial of Service (DoS) attack to more Advanced Persistent Threats (APTs) attack [2]. Moreover, Artificial Intelligence (AI) is being increasingly integrated into intrusion detection systems for enhanced efficiency. In addition, AI can detect behavioral patterns in system operations uncovering unidentified attack techniques. With the help of machine learning algorithms, AI-powered systems can adjust to evolving threats and can enhance detection accuracy [3].

This research aims to create a Host-based Intrusion Detection System (HIDS) that combines the strengths of signature-based detection and AI-driven anomaly detection methods. The main goal is to develop an effective intrusion detection system that can accurately detect and counter cyber threats. By leveraging on the capabilities of these technologies, the proposed system seeks to enhance intrusion detection efficiency by providing a security solution for individuals, for individuals and organizations to safeguard their digital assets against new and complex threats. This study also establishes the foundation for real-time detection, and sharing of cyber threat intelligence in cloud environments. Future versions of the Hybrid HIDS will operate across multiple hosts, supported by a web service for cross-platform compatibility and centralized alert system.

Research Contributions

- 1) This research brings together established security guidelines to outline the requirements, for intrusion detection systems (IDSs) ensuring that all necessary system criteria are considered.
- 2) The research introduces an approach by combining signature-based with AI-driven anomaly detection techniques in an IDS to enhance the identification of cyber threats.
- 3) This study lays the groundwork for Hybrid IDS capable of real-time cyber threat detection, analysis, intelligence sharing, cross-platform compatibility and a centralized

alert system.

Organization of the paper

The paper is organized as: section II provides a literature review. Section III provides a details overview of the proposed solution and system architecture. Section IV provides experimental setup and system configuration details. Section V offers in-depth insights into the results obtained and simulations conducted for various attacks. Section VI introduces future research laid down by this research. Section VII is conclusion.

II. RELATED WORK: REVIEW OF THE LITERATURE

In the realm of cybersecurity, Intrusion Detection Systems (IDS) play an important role by reporting any suspicious activities, within digital systems. IDS can be broadly categorized into two types; Signature Based IDS and Anomaly Based IDS. Signature-based IDS, also known as misuse detection rely on established patterns or signatures of known threats to spot malicious activities. While effective at spotting recognized threats, on the other hand, these systems have limitations in detecting threats that do not align with their existing pattern in database [4]. Nevertheless, anomaly-based IDS focus on identifying deviations from system behavior that detects previously unrecognized threats. However, this approach may result in a number of challenges. For instance, harmless activities could be flagged as potential suspicious threat [5]. The drawbacks of these methods emphasize the necessity for an integrated solution that combines the strengths of both strategies to offer adaptable threat detection capabilities.

There has been a growing interest, in blending signature-based and anomaly-based techniques in research [6]. These combined methods aim to make intrusion detection systems more effective, by utilizing the strengths of both approaches. By merging the threat detection of signature-based systems with the features of anomaly-based systems these hybrid models enhance the capability to identify known and unknown threats while also minimizing false alarms [7]. This dual strategy helps overcome the shortcomings of relying on one method, thereby providing a more dependable cybersecurity solution.

Furthermore, [8] engineered a hybrid system, integrating two or more classifiers. Their design performed commendably in identifying diverse cyber threats but struggled with scalability, becoming computationally costly with larger datasets.

Otoum used a machine learning-driven approach to investigate the field of intrusion detection in critical infrastructures monitored by sensor networks [9]. The research most likely used the KDD'99 dataset to test out the proposed approach. While specifics are lacking, it is safe to assume that Otoum's work includes utilising several machine learning algorithms (including supervised, unsupervised, and reinforcement learning approaches) to improve the accuracy of intrusion detection. There is a promise that this study may yield useful insights that can be used to strengthen security measures, but there

might be possibilities of limitations due to issues with dataset representation, algorithm flexibility, and real-time processing constraints.

A logistic regression host-based intrusion detection system (LR-HIDS) was presented by Besharati et al. [10] for use in the cloud. Security issues in cloud computing were investigated, with a focus on the difficulties of detecting malicious assaults of varying forms. The LR-HIDS approach includes picking relevant features with logistic regression, and then improving them with regularisation methods. The attack categorization was carried out by combining the bagging technique with a neural network, a decision tree, and a linear discriminate analysis. The NSL-KDD dataset in Cloudsim was used to measure how well the LR-HIDS framework performed.

Despite the significant contributions from previous studies, voids in the IDS research landscape still exist. The approach to address this issue involves integrating signature-based and AI-driven anomaly detection techniques. The model will utilize a gradient-boosting classifier for signature-based detection in order to recognise known threats using predefined signatures [11]. Additionally, the AI-driven anomaly detection will employ the k-nearest Neighbour (KNN) algorithm to identify deviations from established normal behaviour patterns [12]. In cases of prediction conflict between the two methods, a voting mechanism will be implemented, with the Gradient Boosting classifier's predictions taking precedence. The proposed method aims to provide a comprehensive, efficient, and precise intrusion detection system capable of identifying both known and unknown cyber threats.

In conclusion, the related works emphasise the gaps that this research work aims to address, particularly the need for a holistic IDS that combines high detection accuracy, real-time performance, and reduced false positive rates. The suggested IDS, integrating a distinctive mix of techniques, aspires to bridge these voids in the constantly shifting and perilous digital security terrain.

III. PROPOSED HYBRID IDS METHODOLOGY

The proposed approach is a Hybrid Host-based Intrusion Detection System (HIDS) that combines signature-based and AI-driven anomaly detection methods to create a comprehensive and effective intrusion detection framework. The signature-based component of the HIDS employs a Gradient Boosting Classifier to detect known attack patterns by comparing network traffic or system activity against a database of predefined signatures. Gradient Boosting is an ensemble learning technique that iteratively builds a strong classifier by combining multiple weak classifiers, typically decision trees. For the AI-driven anomaly detection component, the k-Nearest Neighbors (KNN) algorithm is utilized. KNN is an instance-based learning algorithm that classifies data points based on the majority class of their k-nearest neighbors in the feature space. By learning the patterns of normal system behavior during

training, the KNN model can identify significant deviations as potential cyber threats during the testing phase.

The two components (Gradient Boosting and KNN) operate in parallel, and their results are combined using a voting mechanism. In case of a conflict between the predictions of the two models, the Gradient Boosting classifier's predictions are given priority, as they are considered more reliable for known attack patterns.

When a potential threat or suspicious activity is detected the system quickly sends out alerts to the Security Operations Center (SOC) or Network Operations Center (NOC). These centers are in charge of overseeing and handling security incidents to uphold the reliability and availability of network operations. The alerts from this Hybrid Intrusion Detection System provide details such as the nature of the threat, and affected system. This information empowers SOC and NOC staff to respond reducing the impact of identified threats on the organization's infrastructure [13]. Having alert system integrated into IDS frameworks is crucial for maintaining a security stance, and ensuring a response to emerging cyber threats.

The integration of signature-based and AI-driven anomaly detection techniques aims to leverage the strengths of both approaches, resulting in a hybrid system capable of identifying both known and unknown cyber threats effectively. Fig. 1 illustrates the overall architecture of the proposed Hybrid Host-based Intrusion Detection System (HIDS).

IV. EXPERIMENTAL SETUP

A. Brief Overview of the NSL-KDD Dataset Used for Evaluation

The NSL-KDD dataset, a revised edition of the KDD'99 dataset is commonly utilized in studies, of detecting intrusions [14]. It was created to overcome drawbacks of the dataset like redundancies, and irrelevant entries in order to enhance its suitability for intrusion detection purposes. This dataset comprises attributes that depict network traffic data such as attacks. Noteworthy attributes include connection duration, protocol type, service used, source and destination bytes along with indicators of network behavior.

The proposal uses many tools and software to create a hybrid intrusion detection system. Google Colab, a GPU-equipped cloud-based Jupyter environment, was used for resource-intensive machine learning. The project used Python for methodology, assessments, and data management. Scikit-learn, a popular Python machine learning toolkit, helped create Decision Trees, KNN, and Gradient Boosting. Pandas and NumPy helped with data manipulation and preparation, while Scikit-learn's Label Encoder and Standard Scaler converted categorical data, ensured consistent feature scaling, and enabled machine learning model compatibility. CPUs and other hardware resources were efficiently used to train machine-learning models. This complete set of tools and software helped create and test the hybrid intrusion detection system.

TABLE I: Key Features of the NSL-KDD Dataset

Feature Name	Description	Type
Duration	Total time of the connection	Continuous
Protocol Type	Network protocol types such as TCP, UDP, FTP	Categorical
Service	Network service, e.g., HTTP, FTP	Categorical
Flag	Status of the connection, e.g., SF, S0	Categorical
Src_bytes	Bytes sent from source to destination	Continuous
Dst_bytes	Bytes sent from destination to source	Continuous
Land	1 if the connection is from/to the same host/port	Binary
Wrong Fragment	Number of "wrong" fragments in the connection	Continuous
Urgent	Number of urgent packets in the connection	Continuous
Hot	Number of "hot" indicators in the connection	Continuous

B. Data Preprocessing Steps

Prior to utilizing the NSL-KDD dataset for training, and assessment, various preprocessing procedures were carried out to guarantee the quality of data, and its appropriateness for analysis.

These are:

1) Data Cleaning:

- Dealing with data and outliers required examining the dataset to spot patterns and extreme values. Utilizing aids like pie charts and histograms aided in pinpointing these irregularities, which were then handled properly to uphold the integrity of the dataset.

2) Imputation of Missing Values:

- Missing data was dealt with through imputation methods, such as validating input data and predicting replacements. This approach involved substituting missing values with approximations to maintain the dataset's reliability without causing bias.

3) Data Encoding:

- Categorical attributes were changed through a process called one encoding. For example, a categorical attribute, like "Protocol Type" with options such as "TCP," "UDP," and "FTP" would be transformed into columns for each protocol type.

4) Feature Selection:

- The intrusion detection system chooses features by analyzing correlations and mutual information. A heat map was used to show how features were related, helping to find, and keep the ones while getting rid of unnecessary ones.

5) Normalization or Standardization:

- The characteristics were adjusted to maintain significance throughout the learning phase. Normalization rescaled the attributes to a range of 0, to 1 whereas standardization modified them to possess an average of 0 and a standard deviation of 1.

C. Train-Test Split and Evaluation Metrics

The dataset was split into two parts, for training and testing purposes. The training data was utilized to teach the intrusion

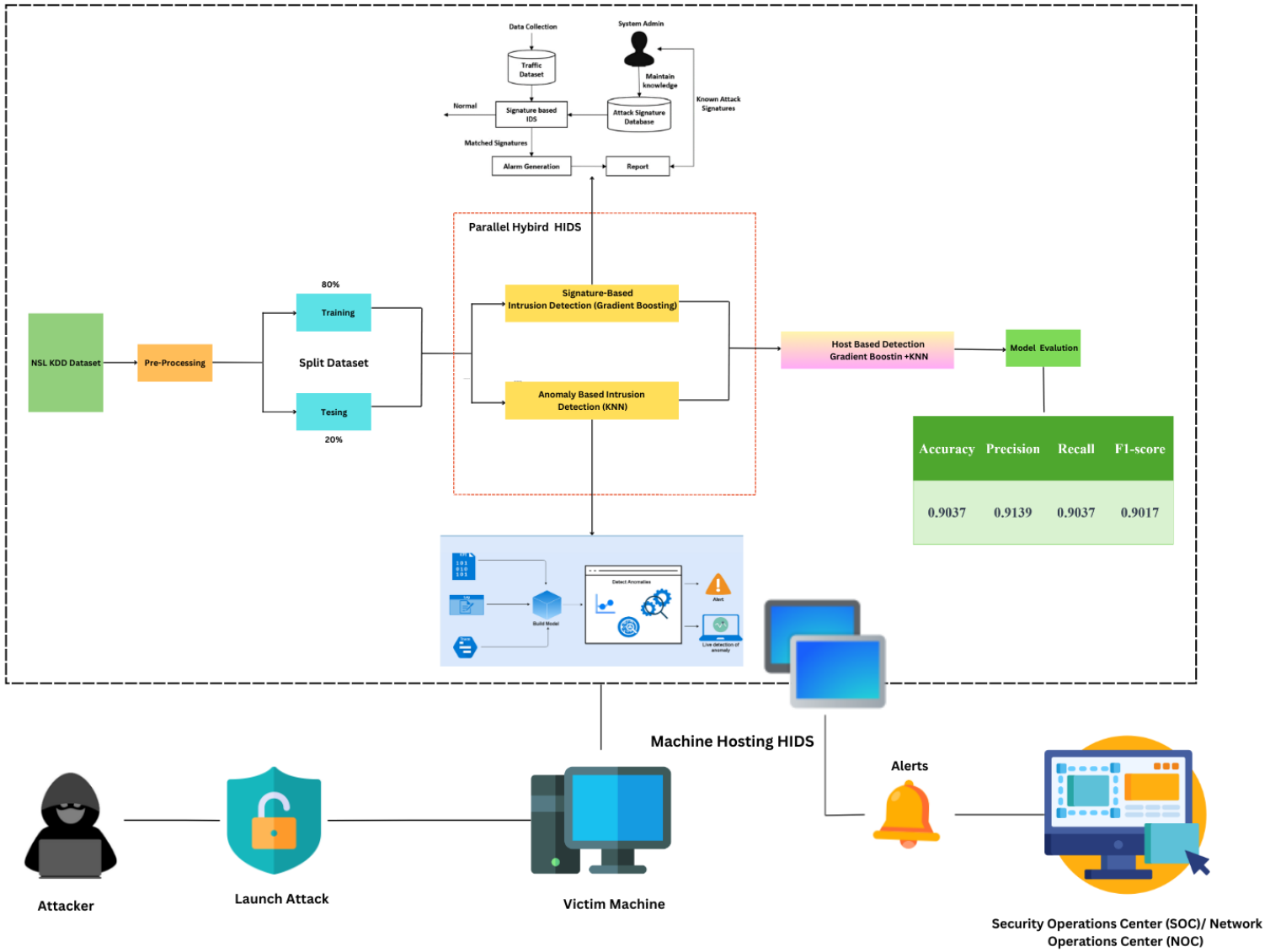


Fig. 1: Provides a comprehensive overview of the architecture.

detection model while the testing data was used to gauge its effectiveness. Various evaluation metrics were employed to assess the model's performance.

- **Accuracy:** The ratio of correctly identified cases, among all cases.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- **Precision:** The ratio of outcomes (true positive), to all the predicted positive outcomes.

$$Precision = \frac{TP}{TP + FP}$$

- **Recall:** The ratio of outcomes (true positive), among all actual positive cases.

$$Recall = \frac{TP}{TP + FN}$$

- **F1-Score:** The harmonic mean of precision, and recall provides a single measure of the model's effectiveness.

$$F1 - Score = \frac{Precision * Recall}{Precision + Recall}$$

V. RESULTS AND DISCUSSION

A. Performance Evaluation of the Hybrid IDS Model

The effectiveness of the combined Intrusion Detection System (IDS) setup, which combines signature-based, with anomaly based detection was tested using the NSL-KDD dataset. Evaluation criteria such as accuracy, precision, recall and F1 score were employed. The integrated model outperformed the models, in all aspects of evaluation.

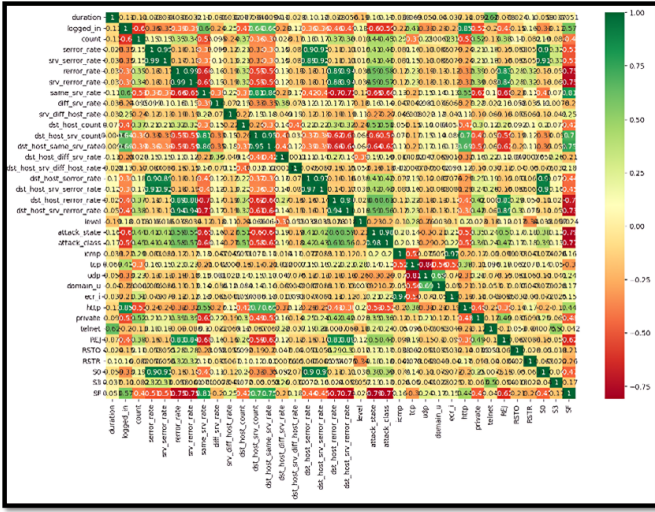


Fig. 2: Heatmap Generated Using Python for Data Visualization

TABLE II: Performance Metrics for Different IDS Models

Model	Accuracy	Precision	Recall	F1-Score
Signature-Based (Gradient Boosting)	0.9037	0.9139	0.9037	0.9017
Anomaly-Based (KNN)	0.8689	0.8933	0.8689	0.8724
Hybrid IDS	0.9037	0.9139	0.9037	0.9017

B. Comparative Analysis with Standalone Signature-Based and Anomaly-Based Models

The analysis comparing models revealed that the combined IDS model performed better than both the signature-based and anomaly-based models. The signature-based model had an accuracy of 90.4%, precision of 91.4%, recall of 90.4% and F1 score of 90.2%. On the other hand, the anomaly-based model achieved an accuracy of 87.0%, precision of 89.3%, recall of 87.0%, and an F1 score of 87.2%. In contrast, the hybrid model attained results with an accuracy of 90.4%, precision of 91.4%, recall of 90.4%, and F1 score of 90.2%.

The exceptional effectiveness of the model stems from its capacity to capitalize on the advantages of both detection methods. By merging the detection accuracy of signature-based techniques with the nature of anomaly-based methods the hybrid model adeptly detects a wide range of attacks, whether known or unknown.

C. The Advantages and Improvements Offered by the Hybrid Approach

The hybrid IDS method provides benefits compared to standalone models.

- **Improved Detection Accuracy:** The hybrid approach merges the accuracy of signature based detection, with the adaptability of anomaly based detection resulting in improved precision.
- **Balanced Performance:** The hybrid model maintains a good balance between precision and recall, as indicated

by the high F1-score. This ensures that the model not only identifies a high proportion of attacks but also minimizes false positives.

- **Enhanced Generalization:** The inclusion of anomaly-based detection allows the hybrid model to generalize better to new and unseen attack patterns, addressing one of the main limitations of signature-based methods.
- **Reduced False Positives:** The voting mechanism between the two components helps to reduce the number of false positives, which is a common issue in anomaly-based detection systems.

D. Visualizations to Support the Findings

To provide a clearer understanding of the performance and improvements offered by the hybrid IDS model, several visualizations were created, including confusion matrices and performance charts.

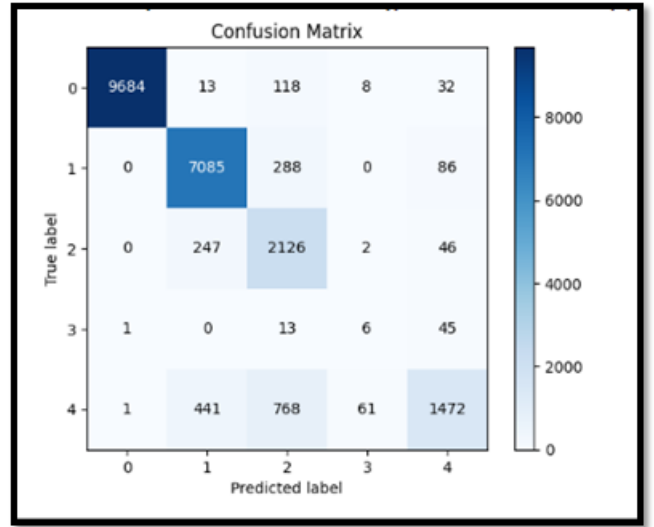


Fig. 3: Confusion Matrix of Hybrid HIDS

Hybrid Host Based Intrusion detection system Evaluation Metrics

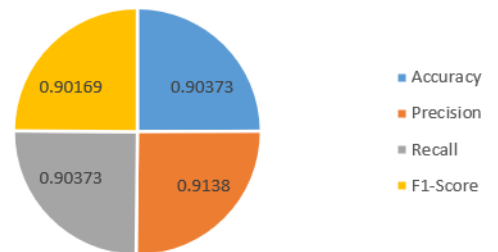


Fig. 4: Performance Metrics of Hybrid HIDS

Fig. 3 shows the confusion matrix for the hybrid IDS model, illustrating its high true positive and true negative rates,

TABLE III: Comparative Analysis with Existing Research

Previous Research	Database Used	Methodology	Achieved Accuracy	Limitations
Liu, Gu, and Wang (2021)	NSL-KDD	Hybrid system integrating two or more classifiers	85.24%	Lack of scalability and high computational cost with larger datasets.
Otoun (2019)	KDD'99	Adaptively Supervised and Clustered Hybrid (ASCH-IDS)	95%	Q-learning is known to have slow convergence, especially in large state and action spaces.
Besharati, Naderan, and Namjoo (2019)	NSL-KDD	Logistic regression host-based intrusion detection system for cloud environments	97.51%	Accuracy lacks in-depth robustness analysis against adversarial attacks or false positives, and complexity due to classifier combination and regularization techniques.
Our Method	NSL-KDD	Host-based Intrusion Detection Using Signature-based (Gradient Boosting) and AI-driven Anomaly Detection Methods (KNN)	90.37%	Accuracy could decrease if training data is insufficient; currently limited to a single host machine.

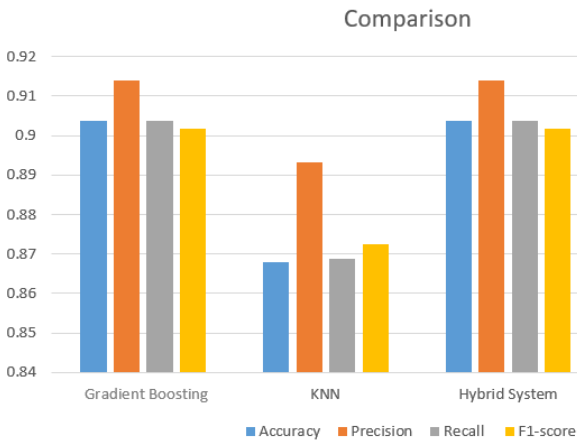


Fig. 5: Performance Metric Comparison: Hybrid HIDS vs. Traditional IDS

which contribute to its overall high accuracy. Fig. 4 shows the performance metrics of our solution; Hybrid Host-based Intrusion Detection System. Fig. 5 compares the performance metrics (accuracy, precision, recall, and F1-score) of the hybrid IDS model with the standalone signature-based and anomaly-based models, clearly demonstrating the improvements offered by the hybrid approach.

These visualizations support the conclusion that the hybrid IDS model provides a more robust and effective solution for intrusion detection compared to traditional standalone models.

VI. FOUNDATION FOR FUTURE HYBRID HIDS RESEARCH

This study lays the groundwork for efforts to create a system, for real-time detection, analysis, and sharing of cyber threat intelligence in cloud environments. Although the current proposed solution is focused on one host machine, upcoming versions of the Hybrid Host-based Intrusion Detection System (HIDS) will be implemented on host machines situated in different geographical locations. The future version will incorporate multiple datasets to avoid reliance on single a dataset. To support this development a web service will be designed to act as a mediator between the Hybrid HIDS

and a central database facilitating communication. This web service will ensure compatibility across platforms enabling the Hybrid HIDS to function on any operating system. As proposed and developed in our cloud security research [13], [15]. Furthermore, alerts by the Hybrid HIDS will be sent to a web portal that serves as a hub where cloud service providers, security entities and the general public can access Indicators of Compromise (IOCs). This integrated strategy aims to boost security by improving threat detection efficiency and promoting intelligence sharing, among platforms and stakeholders. After the success of this research, Fig. 6 illustrates the future research architecture proposed by this study.

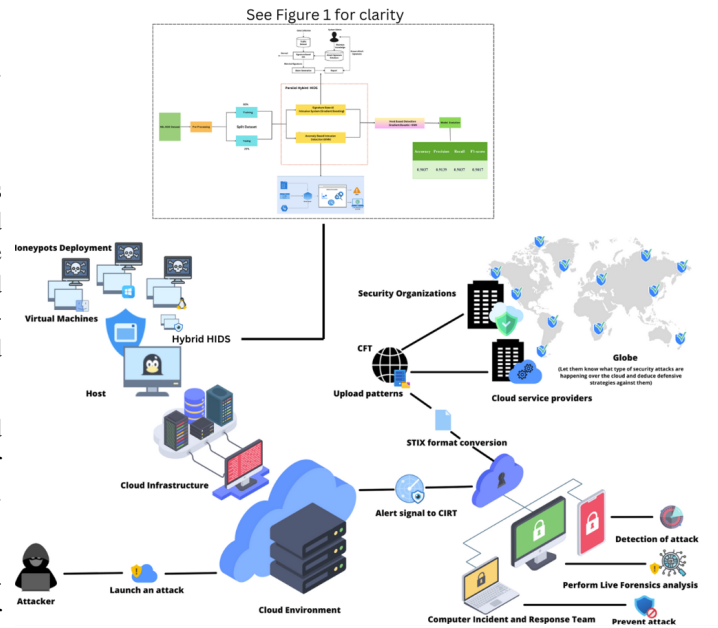


Fig. 6: Overview of the proposed future Hybrid HIDS architecture

VII. CONCLUSION

This research introduces a pioneering hybrid Intrusion Detection System (HIDS) integrating signature-based and AI-driven anomaly detection methods. Key contributions include

the development of a robust HIDS model combining Gradient Boosting and k-Nearest Neighbour algorithms. The model achieved 90% accuracy in identifying known threats and in detecting anomalies. This research highlights the potential of hybrid methodologies in enhancing cybersecurity defences and lays the groundwork for future advancements in intrusion detection systems. They will not be restricted to one system but will expand to cover multiple hosts in different locations. This will allow for compatibility, across platforms and the sharing of real-time threat intelligence. The study demonstrates how combining approaches can strengthen cybersecurity defences and set the stage for improvements, in intrusion detection systems.

REFERENCES

- [1] A. Thakkar and R. Lohiya, "A review on challenges and future research directions for machine learning-based intrusion detection system," *Archives of Computational Methods in Engineering*, vol. 30, no. 7, pp. 4245–4269, 2023.
- [2] F. Rehman, J. Hashmi, M. Abdullah, and H. Zaman, "Guarding voices, protecting homes: A comprehensive case study on voice assistant security in smart living," in *2024 12th International Symposium on Digital Forensics and Security (ISDFS)*, pp. 1–6, IEEE, 2024.
- [3] C. Pham-Quoc, T. H. Q. Bao, and T. N. Thanh, "Fpga/ai-powered architecture for anomaly network intrusion detection systems," *Electronics*, vol. 12, no. 3, p. 668, 2023.
- [4] J. Goh, S. Adepu, M. Tan, and Z. S. Lee, "Anomaly detection in cyber physical systems using recurrent neural networks," in *2017 IEEE 18th International Symposium on High Assurance Systems Engineering (HASE)*, pp. 140–145, IEEE, 2017.
- [5] O. Y. Al-Jarrah, P. D. Yoo, S. Muhaidat, G. K. Karagiannidis, and K. Taha, "Efficient machine learning for big data: A review," *Big Data Research*, vol. 2, no. 3, pp. 87–93, 2015.
- [6] A. Alrawais, A. Alhothaily, C. Hu, and X. Cheng, "Fog computing for the internet of things: Security and privacy issues," *IEEE Internet Computing*, vol. 21, no. 2, pp. 34–42, 2017.
- [7] A. N. Cahyo, A. K. Sari, and M. Riasetiawan, "Comparison of hybrid intrusion detection system," in *2020 12th international conference on information technology and electrical engineering (ICITEE)*, pp. 92–97, IEEE, 2020.
- [8] C. Liu, Z. Gu, and J. Wang, "A hybrid intrusion detection system based on scalable k-means+ random forest and deep learning," *Ieee Access*, vol. 9, pp. 75729–75740, 2021.
- [9] S. Otoum, *Machine Learning-driven Intrusion Detection Techniques in Critical Infrastructures Monitored by Sensor Networks*. PhD thesis, Université d'Ottawa/University of Ottawa, 2019.
- [10] E. Besharati, M. Naderan, and E. Namjoo, "Lr-hids: logistic regression host-based intrusion detection system for cloud environments," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, pp. 3669–3692, 2019.
- [11] P. Verma, S. Anwar, S. Khan, and S. B. Mane, "Network intrusion detection using clustering and gradient boosting," in *2018 9th International conference on computing, communication and networking technologies (ICCCNT)*, pp. 1–7, IEEE, 2018.
- [12] M. Markevych and M. Dawson, "A review of enhancing intrusion detection systems for cybersecurity using artificial intelligence (ai)," in *International conference Knowledge-based Organization*, vol. 29, pp. 30–37, 2023.
- [13] F. Rehman, Z. Muhammad, S. Asif, and H. Rahman, "The next generation of cloud security through hypervisor-based virtual machine introspection," in *2023 3rd International Conference on Artificial Intelligence (ICAI)*, pp. 116–121, IEEE, 2023.
- [14] L. Dhanabal and S. Shantharajah, "A study on nsl-kdd dataset for intrusion detection system based on classification algorithms," *International journal of advanced research in computer and communication engineering*, vol. 4, no. 6, pp. 446–452, 2015.
- [15] F. Rehman and S. Hashmi, "Enhancing cloud security: A comprehensive framework for real-time detection analysis and cyber threat intelligence sharing," *Advances in Science, Technology and Engineering Systems Journal*, vol. 8, no. 6, pp. 107–119, 2023.