IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CyberShield -Sentinel Intelligence for Secure IoT Environments

Nithiya G
Department of Computer Science and Engineering
Anjalai Ammal Mahalingam Engineering College Kovilvenni, Thiruvarur,
Tamil Nadu, India

Sri Vardhini D
Department of Computer Science and Engineering
Anjalai Ammal Mahalingam Engineering College Kovilvenni, Thiruvarur,
Tamil Nadu, India

Mrs. Karthika K
Assistant Professor, Department of Computer Science and Engineering
Anjalai Ammal Mahalingam Engineering College Kovilvenni, Thiruvarur,
Tamil Nadu, India

Abstract

The rapid growth of Internet of Things (IoT) devices in domains such as smart homes, healthcare, and industrial systems has introduced significant cybersecurity challenges. Many IoT devices have limited computational resources and weak security mechanisms, making them vulnerable to cyber threats such as malware, denial-of-service attacks, and unauthorised access. Traditional intrusion detection systems (IDS) that rely on rule-based or signature-based methods often fail to detect zero-day attacks and evolving intrusion patterns. To address these limitations, this paper proposes CyberShield – Sentinel Intelligence for Secure IoT Environments, an AI-powered intrusion detection framework. The system uses the NSL-KDD dataset and employs a hybrid machine-learning approach integrating Isolation Forest for anomaly detection and Random Forest for threat classification. In addition, Small Batch Processing is implemented to improve scalability and reduce computational overhead. The proposed framework enhances detection accuracy and provides an efficient solution for securing IoT networks.

Keywords: IoT Security, Intrusion Detection System (IDS), NSL-KDD Dataset, Hybrid Machine Learning, Isolation Forest, Random Forest, Anomaly Detection, Small Batch Processing.

I. Introduction

The Internet of Things (IoT) has emerged as a critical component of modern digital infrastructure, enabling seamless connectivity among heterogeneous devices across applications such as smart homes, healthcare systems, industrial automation, transportation, and smart cities. These interconnected systems facilitate real-time monitoring, automated control, and efficient data exchange, enhancing operational efficiency and decision-making. However, the rapid proliferation of IoT devices introduces substantial cybersecurity challenges. Many devices operate with limited computational power, memory, and built-in security, rendering them susceptible to attacks, including Distributed Denial-of-Service (DDoS), device spoofing, unauthorised access, malware injection, and data breaches. Such attacks can disrupt network operations and compromise sensitive information.

Conventional signature-based Intrusion Detection Systems (IDS) detect known threats by matching network traffic against predefined signatures, but are often ineffective against zero- day attacks and

require continuous maintenance. In contrast, machine learning-based IDS analyse network behaviour to identify anomalies, offering improved adaptability, scalability, and suitability for dynamic IoT environments.

To address these limitations, this study proposes CyberShield – Sentinel Intelligence for Secure IoT Environments, an AI-driven intrusion detection framework. The framework integrates Isolation Forest for anomaly detection and Random Forest for threat classification, combined with small batch processing to efficiently handle large volumes of network traffic. The goal is to provide accurate, scalable, and real-time threat detection for modern IoT ecosystems.

II. Literature Survey

Recent studies have explored the use of intelligent techniques for improving intrusion detection in modern network environments. Usmani, Hussain, and Hasan [1] proposed the Sentinel AI framework, demonstrating how artificial intelligence can enhance cybersecurity through adaptive intrusion detection systems. Their work highlighted the importance of AI-driven approaches in identifying complex attack patterns and improving detection accuracy. Kiran et al. [2] investigated the application of machine learning algorithms in intrusion detection systems, showing that ML-based models can effectively identify abnormal network behaviour when compared with traditional rule-based approaches. Similarly, Azam, Islam, and Huda [3] conducted a comparative analysis of decision tree-based intrusion detection models, reporting improved classification performance in detecting cyber threats. Saqib, Malhotra, Ali, and Tariq
[4] examined the role of IoT sensor networks and big data analytics in large-scale infrastructures, emphasising the need for efficient data processing and secure communication mechanisms in distributed IoT systems. Nuthalapati [5] discussed the importance of optimised data architectures and preprocessing techniques for managing large-scale data environments, highlighting the role of efficient data management in improving analytical performance. Sanjrani, Saqib, and Tariq [6] provided a comprehensive survey of security challenges in IoT and edge computing environments, identifying vulnerabilities associated with distributed network architectures. Smith and Doe [7] proposed a behaviour-based anomaly detection model for resource-constrained IoT devices, demonstrating improved adaptability in dynamic network conditions. Recent research also highlights the benefits of hybrid machine learning approaches for improving intrusion detection accuracy. Wang [8] conducted a systematic review of hybrid machine learning models for intrusion detection, concluding that combining multiple algorithms can significantly enhance detection performance and reduce false positives. Gupta [9] analysed the effectiveness of Small Batch Processing in handling high- velocity network traffic, demonstrating its advantages in reducing computational overhead and improving processing efficiency. In addition, Lee [10] studied the application of Isolation Forest for anomaly detection, showing its capability to identify unknown and zero-day attacks in network traffic datasets.

These studies indicate that combining behaviour-based anomaly detection, hybrid machine learning techniques, and efficient data processing mechanisms can significantly improve intrusion detection performance. Based on these findings, the proposed CyberShield framework integrates Isolation Forest, Random Forest, and Small Batch Processing to provide an efficient and scalable intrusion detection solution for IoT environments.

III. Proposed CyberShield Framework

CyberShield – Sentinel Intelligence for Secure IoT Environments is an AI-driven intrusion detection framework for IoT networks. Using the NSL-KDD dataset, it employs a hybrid approach combining Isolation Forest for anomaly detection and Random Forest for threat classification.

Network traffic from IoT devices undergoes data preprocessing, including cleaning, normalization, and feature selection (packet size, frequency, port number, communication duration). To handle large datasets efficiently, the framework implements small batch processing, reducing memory usage and computation time.

Detection begins with Isolation Forest to identify abnormal traffic. Detected anomalies are classified by Random Forest into normal, suspicious, or malicious categories. The system then triggers corresponding actions: monitoring normal traffic, alerting for suspicious activity, or blocking malicious traffic. The overall architecture is shown in **Figure 1.**

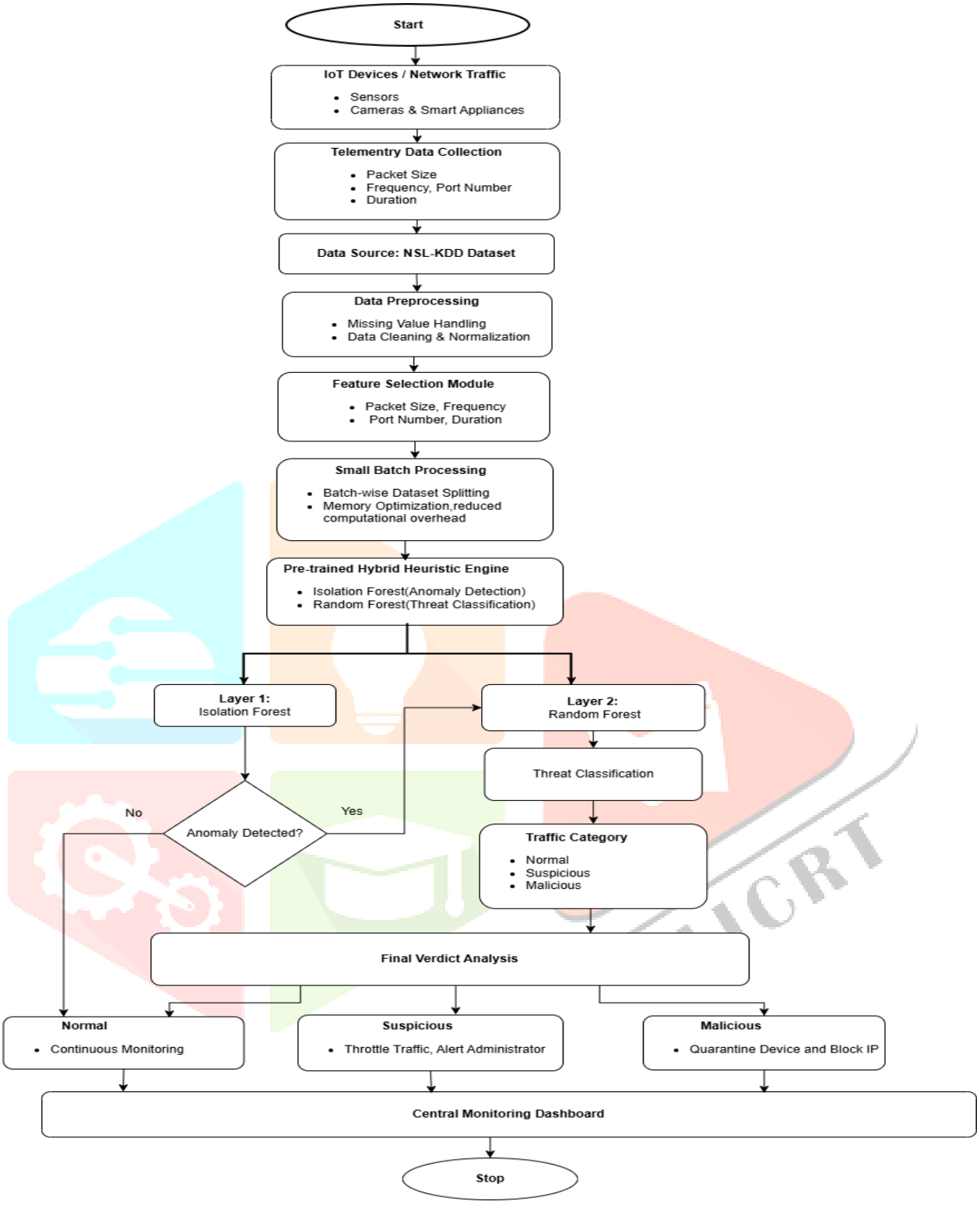

Figure 1: Proposed CyberShield Architecture for Secure IoT Environments

Methodology:

A. Data Collection

The first step in the proposed framework is collecting network traffic data for cyber threat detection. In this study, the **NSL-KDD dataset** is used as the primary data source. The dataset contains both normal and malicious network traffic records representing various types of cyberattacks. It is widely used in intrusion detection research for training and evaluating machine learning models.

B. Data Preprocessing

After collecting the dataset, preprocessing is performed to improve data quality. This process includes removing missing values, eliminating noisy data, and normalising feature values to ensure consistent input for machine learning algorithms. Data preprocessing helps improve model performance and ensures reliable analysis of network traffic data.

C. Feature Selection

Feature selection is applied to identify the most relevant attributes required for cyber threat detection. Important features such as packet size, protocol type, connection duration, and port number are selected to represent network communication behaviour. Selecting the most relevant features reduces data dimensionality and improves the efficiency of the detection system.

D. Hybrid Detection Model

The main component of the proposed framework is the **Hybrid Detection Model**, which combines **Isolation Forest** and **Random Forest** algorithms. The **Isolation Forest** algorithm is used for anomaly detection to identify unusual patterns in network traffic data. It isolates abnormal data points from normal data by randomly partitioning the dataset. Data points that require fewer splits to isolate are considered anomalies.

After anomaly detection, the detected data is passed to the **Random Forest** classifier. Random Forest is an ensemble learning algorithm that uses multiple decision trees to classify network traffic. Each decision tree predicts the class label of the input data, and the final classification is determined through majority voting. The classifier categorises the traffic into normal or malicious activity.

E. Small Batch Processing

To improve computational efficiency, the system processes the dataset using **small batch processing**. When the dataset is uploaded, it is divided into smaller batches before being analysed by the hybrid detection model. Each batch of data is processed sequentially for anomaly detection and classification. This approach reduces memory usage and allows faster analysis of large network traffic datasets, making the system more efficient for cyber threat detection in IoT environments.

V. Results and Discussion

The performance of the proposed cyber threat detection framework was evaluated using the NSL-KDD dataset. The dataset was processed using the hybrid detection model that integrates Isolation Forest and Random Forest algorithms. The evaluation focuses on analysing the detection accuracy and classification performance of the proposed system compared with individual machine learning algorithms.

The accuracy comparison of the evaluated machine learning algorithms is illustrated in Figure 2.

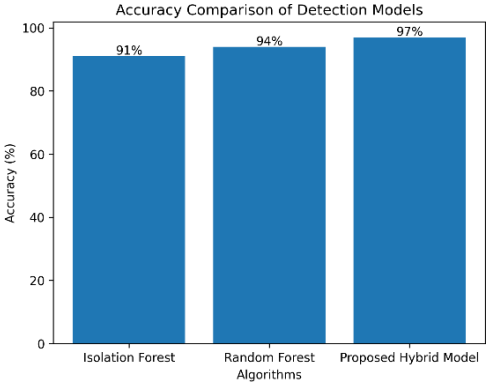

Figure 2: Accuracy Comparison

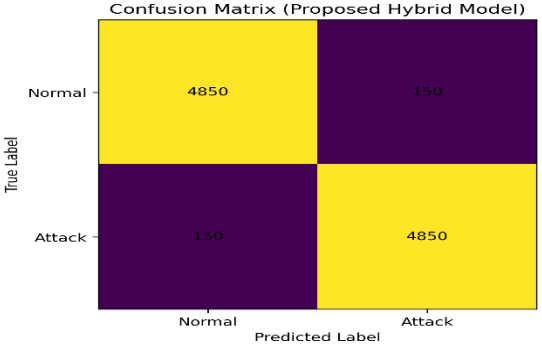

Figure 3: Confusion Matrix

The accuracy comparison shown in Figure 2 indicates the performance difference between the evaluated algorithms. Isolation Forest performs anomaly detection by identifying unusual patterns in

network traffic, while Random Forest performs classification based on learned patterns from the dataset. The experimental results show that the Isolation Forest algorithm achieved an accuracy of **91%**, while the Random Forest algorithm achieved **94%** accuracy. In contrast, the proposed hybrid detection model achieved the highest accuracy of **97%**. This improvement demonstrates that combining anomaly detection and classification techniques can significantly enhance cyber threat detection performance in IoT networks.

The confusion matrix of the proposed hybrid detection model is presented in Figure 3.

The confusion matrix provides a detailed representation of the classification results. It includes the number of correctly and incorrectly classified instances such as true positives, true negatives, false positives, and false negatives.

From the confusion matrix results, it can be observed that the proposed hybrid model correctly classifies most of the normal and malicious network traffic samples with minimal misclassification. This indicates that the model is effective in detecting cyber threats and maintaining high classification accuracy.

In addition, the implementation of small batch processing improves computational efficiency by enabling the system to process large datasets in smaller segments. This approach reduces memory usage and improves processing speed during cyber threat detection.

Overall, the experimental results confirm that the proposed hybrid detection model provides improved accuracy and reliable performance for detecting cyber threats in IoT networks compared to individual machine learning algorithms.

VI. Conclusion

In this paper, a hybrid cyber threat detection framework for IoT networks was proposed using Isolation Forest and Random Forest algorithms. The objective of the proposed system is to improve the accuracy and reliability of detecting cyber threats in IoT environments. Isolation Forest is used for anomaly detection to identify unusual patterns in network traffic, while Random Forest performs classification to determine whether the detected traffic is normal or malicious.

The experimental evaluation was carried out using the NSL-KDD dataset. The results show that the proposed hybrid model achieves higher accuracy compared to the individual machine learning algorithms. The model achieved an accuracy of **97%**, outperforming Isolation Forest and Random Forest. The confusion matrix analysis also confirms that the model effectively classifies most of the network traffic samples with minimal misclassification.

Overall, the proposed hybrid approach provides an efficient and reliable solution for detecting cyber threats in IoT networks. Future work may focus on integrating deep learning techniques and evaluating the model with more recent cybersecurity datasets.

REFERENCES

[1] T. N. Usmani, M. Z. Hussain, and M. Z. Hasan, “Sentinel AI: Revolutionising Cybersecurity with Intelligent Intrusion Detection,” *Dialogue Social Science Review (DSSR)*, vol. 3, no. 1, pp. 1445–1462, 2025.

[2] A. Kiran et al., “Intrusion detection system using machine learning,” in *2023 International Conference on Computer Communication and Informatics (ICCCI)*, 2023, pp. 1–4.

[3] Z. Azam, M. M. Islam, and M. N. Huda, “Comparative analysis of intrusion detection systems and machine learning-based model analysis through decision tree,” *IEEE Access*, vol. 11, pp. 80348–80391, 2023.

[4] M. Saqib, S. Malhotra, R. Ali, and H. Tariq, “Harnessing Big Data Analytics for Large- Scale Farms: Insights from IoT Sensor Networks,” *International Journal of Advanced Research, Ideas and*

Innovations in Technology, vol. 11, no. 1, 2025.

[5] A. Nuthalapati, "Architecting Data Lake-Houses in the Cloud: Best Practices and Future Directions," *Int. J. Sci. Res. Arch.*, vol. 12, no. 2, pp. 1902–1909, 2024.

[6] A. Sanjrani, M. Saqib, and H. Tariq, "Security challenges in IoT and edge computing: A survey," *IEEE Access*, vol. 13, pp. 7555–7573, Jan. 2025.

[7] J. Smith and K. Doe, "Behaviour-based anomaly detection in resource-constrained IoT devices," *Journal of Network Security*, vol. 15, no. 3, 2024.

[8] L. Wang, "Hybrid machine learning for intrusion detection: A systematic review," *Artificial Intelligence Review*, vol. 22, no. 4, 2023.

[9] R. Gupta, "Efficiency of Small Batch Processing in high-velocity network traffic," *International Journal of IoT Studies*, vol. 8, no. 1, 2024.

[10] F. Lee, "Isolation Forest and its role in modern anomaly detection systems," *Cybersecurity Research*, vol. 14, no. 3, 2024.