



AI Based IDS + IPS for Enhanced Cybersecurity

K.S Suresh Babu
Department of Computer Engineering
Pillai College of Engineering
New Panvel, India
sureshbabu@mes.ac.in

Swapnil Jankar
Department of Information Technology
Pillai College of Engineering
New Panvel, India
Swapnil21it@student.mes.ac.in

Shaurya Malik
Department of Information Technology
Pillai College of Engineering
New Panvel, India
Smalik21it@student.mes.ac.in

Harsh Yadav
Department of Information Technology
Pillai College of Engineering
New Panvel, India
hyadav22it@student.mes.ac.in

Abstract— In the present scenario, the number of cyber attacks is rising, and the traditional IDS tools can only detect the intrusions, not prevent them. In this project, an AI-based Intrusion Detection and Prevention System (AI-IDS + IPS) for the Windows operating system has been proposed, which can detect and prevent intrusions using machine learning. In this proposed system, the Random Forest algorithm has been used for supervised learning, and the Isolation Forest algorithm has been implemented for unsupervised learning. The proposed system can detect and prevent intrusions like the Smurf Attack, IP Sweep Attack, Neptune Attack, and Port Sweep Attack. These attacks can be blocked using the Windows Firewall. From the results, the proposed system shows high accuracy with fewer false positives, thereby implying that the proposed AI-based system can automate the security of the modern network.

Keywords— Intrusion Detection System, Intrusion Prevention System, Machine Learning, Random Forest.

I. INTRODUCTION

Intrusion Detection and Prevention Systems are considered to be of great significance in today's world of cybersecurity, as they are always vigilant and ready to respond to any malicious activity. In today's world of increasing cloud computing, IoT devices, and large-scale digital networks, cyber threats are becoming more and more complex. Conventional methods of intrusion detection, such as rule-based and signature-based intrusion detection, are not effective in today's world of complex threats, as they are highly dependent on patterns, which often results in a high rate of false alarms.

For the above challenges, the proposed AI-Based IDPS utilizes unsupervised learning, NLP, and clustering techniques such as DBSCAN and K-Means. It also utilizes feature extraction techniques such as TF-IDF, Word2Vec, and Autoencoders, which allow for the extraction of useful information from the network traffic and log data, facilitating

the detection of anomalies without the use of labeled data sets.

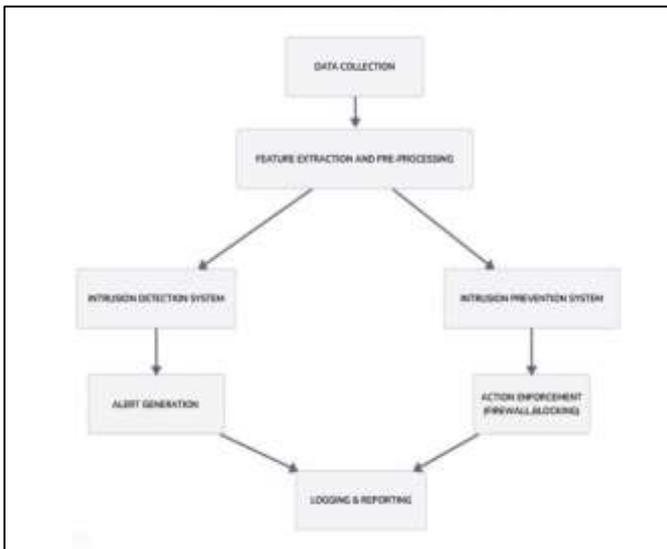
The system workflow includes the following steps: monitoring network traffic, identifying threats using signature-based and anomaly-based approaches, generating alerts, and taking appropriate actions such as blocking malicious IP addresses or filtering harmful packets by network administrators. This process helps in improving the overall strength and security of the network.

II. RELATED WORK

Smith et al. created the first form of a Hybrid IDS and IPS through the extension of the Snort IDS engine using inline packet filtering. Their approach is based on the use of signature-based rules for the identification of attacks and the simultaneous application of preventive measures. However, the authors also realized the drawback of the approach, which is the heavy reliance on signatures. As a result, the approach was not effective in the identification and prevention of unknown attacks. Despite the drawbacks, the approach was a fundamental model for the creation of the Snort Inline.

Gupta and Rao suggested a real-time IDPS system based on machine learning classifiers such as Random Forest and Decision Tree, trained on the NSL-KDD dataset. Their IDPS system also includes an automated firewall module, which blocks the malicious IP addresses identified by the system. In the paper, the authors have shown the effectiveness of the proposed system, where the RF classifier has the highest detection accuracy of 97.1% compared to other classifiers and faster response times compared to traditional signature-based IDS systems.

A combined IDS/IPS system was proposed by Hernandez et al. [3], which utilized an anomaly-based model employing Isolation Forest and K-Means clustering to identify abnormal network activity. The proposed prevention technique utilized dynamic IPTables to filter suspicious network activity



detected by the anomaly-based model. This unsupervised technique was found to have excellent detection capabilities for detecting unknown threats. However, the technique was found to have difficulties in identifying false alarms during times of network fluctuations.

Talukdar and Ahmed in their research article "Integration of Suricata IDS with Automated IPS Engine using NFQUEUE in Linux Platform" discussed the idea of integrating Suricata IDS with an automated IPS system using NFQueue in Linux platforms. This system uses multi-threaded packet inspection and automated rule generation to perform real-time prevention against DoS attacks and brute-force attacks. Experiments were conducted to prove that the system is effective in reducing the impact of attacks when they are performed in an automated manner. However, the system consumes a lot of computational resources and cannot perform effectively on encrypted data.

Patel et al. [5] designed a prototype of IDPS using a machine learning-based model that incorporates signature-based detection and anomaly-based detection using a clustering model called DBSCAN. This model is capable of triggering a prevention mechanism that involves updating the firewall policies to drop the suspicious traffic. This model is effective in detecting zero-day attacks, and the authors were able to reduce the rate of false alarms. However, the model was unable to handle large volumes of traffic due to the computational complexity of clustering high-dimensional vectors.

Khan and Abdullah, in their paper, focused on log-based intrusion prevention systems, using NLP techniques to detect intrusions in network logs. The proposed system used TF-IDF and Word2Vec to convert network log messages into numerical vectors, which were further used to train an Autoencoder-based anomaly detector. Once anomalies were detected, the system used automation to send block rules to both local and cloud firewalls. The proposed system, although successful in detecting subtle anomalies, is highly dependent on the quality of the log data, which becomes a major drawback in environments with inconsistent logging.

Nakamura et al. [7] also stressed the need for real-time prevention in the latest IDPS techniques. In this paper, the authors proposed a reinforcement-learning model that can effectively identify malicious flows and update the firewall rules without any human intervention. In the experiments, the proposed system was more adaptable to the latest attack

patterns, but the training process consumed a considerable amount of computation resources and simulations of attacks.

Overall, the research that has been done on IDPS systems to date indicates that there are certain limitations to signature-based IDPS systems and that there is a need to move forward to a hybrid system that incorporates both supervised and unsupervised learning. While many of the current systems have demonstrated good results in the detection phase, there are issues with scalability and false positives that have been identified. These issues emphasize the need for intelligent real-time IDPS systems that are capable of analyzing network traffic in real-time and enforcing prevention mechanisms.

III. PROPOSED SYSTEM

A. Network Traffic Collection

The system starts by collecting network packets from the host machine using packet capture tools such as Scapy and PyShark. The tools help collect vital information such as source IP address, destination IP address, port numbers, protocol type, packet size, and flag values. For the training and testing of the machine learning components, the NSL-KDD dataset is also integrated into the system. The dataset contains samples of normal and attacking network traffic such as DoS, Probe, R2L, and U2R attacks.

Realism is ensured through the use of both real-time and dataset-generated network samples.

The collected network samples are then stored using MongoDB for easier retrieval during the detection and monitoring processes.

B. Data Preprocessing

Raw packet data is cleaned and transformed into a structured format for the next stage of processing. During the cleaning and transformation stage, redundant data is removed, along with any malformed data and other irrelevant metadata. Normalization and encoding techniques such as Label Encoding and Min-Max Scaling are also applied.

Feature vectors are created for the packet data. This is done by converting the packet data into a numerical format. This allows the model to identify patterns from the data. The dataset is then split using a stratified sampling approach. This approach is used to ensure a balanced representation of the different categories of attacks. Real-time packet data is dynamically transformed into the same format as the training data.

C. Feature Extraction

In the feature extraction phase, relevant information is extracted in the form of meaningful indicators, which are utilized by the signature-based and anomaly-based detection models.

In the statistical network features calculation, the features include the duration of the connections, packet count, bytes transferred, flag ratios, and protocol distribution. Additionally, the feature importance of the Random Forest algorithm is used to determine the most critical features. On the other hand, the anomaly-based detection model calculates the deviation.

This provides a unified representation of the features, which enhances the detection of both known and unknown cyberattacks.

D. Machine Learning-Based Intrusion Detection

The Intrusion Detection module uses a hybrid approach of machine learning.

The Random Forest Classifier, using the NSL-KDD dataset, detects known categories of attacks like Smurf, Neptune, PortSweep, and other types of signature-based attacks.

For detecting unknown types of attacks, the Isolation Forest algorithm is used.

This unsupervised learning algorithm detects unusual behavior in network traffic by providing an anomaly score for incoming packets.

The decision fusion method compares the results of the two algorithms to obtain the final output.

E. Intrusion Prevention and Security Enforcement

After detecting the attack, the intrusion prevention module immediately activates defense actions. This is done through automatic blocking of malicious IP addresses, dropping of harmful packets, and closing of vulnerable ports through the use of Windows Firewall's programmatic controls.

Structured alerts are created and include information on the type of attack, source IP address, time stamp, and actions taken. There is also a real-time dashboard developed through the use of Flask that displays logs, current alerts, and decisions taken. All detected incidents are logged and stored in MongoDB.

Adaptive learning feedback is included in the system. This is achieved through the reintroduction of detection outcomes into the model to enhance future classifications. This creates a closed-loop security system.

IV.METHODOLOGY AND ALGORITHMS

A. Methodology (For AI-Based IDS + IPS)

The methodology of the AI-Based Intrusion Detection and Prevention System (IDPS) is based on a multi-stage, hybrid model that combines supervised machine learning, unsupervised anomaly-based intrusion detection, NLP-based log analysis, and signature-based intrusion detection. Instead of using a single detection technique, the AI-Based IDPS uses a multi-engine, parallel-based model that combines the use of several analytical engines, including Random Forest, Isolation Forest, NLP-based log analysis, clustering using DBSCAN and K-Means, and regex-based rule detection.

A lazy loading strategy is employed to ensure that heavy ML and NLP models are loaded only when actually required, thus minimizing memory consumption and maximizing system efficiency. All the detection output from the parallel models is combined using a confidence-weighted voting mechanism, which prioritizes models that are more confident about the output. A signature or regex-based module is used as a fallback to ensure that if the ML components fail or are unable to classify the event, detection is still performed.

The methodology follows a structured workflow. The steps of the workflow are:

- 1.Network packet capture using Scapy or PyShark.
- 2.Preprocessing of the raw data for cleaning, normalization, and encoding.
- 3.Feature engineering of the network packets using statistical, protocol-based, and log-based features.
- 4.Parallel processing of all the models, which include Random Forest, Isolation Forest, regex, NLP, and clustering.
- 5.Fusion layer with confidence-weighted voting for the final labeling of the threat.

6.Automated Intrusion Prevention using System Firewalls.

7.Feedback learning for learning and improvement.

This multi-layered system ensures accurate and fast detection and prevention of intrusions.

B. Algorithms (For AI-Based IDS + IPS)

1.Regex or Signature-Based Detection

This algorithm uses known cyberattacks and applies rules or signatures to detect these attacks. Here, a set of patterns is used to identify known malicious activities such as SQL injection words, port scan patterns, or unauthorized access. These rules are executed extremely fast, making this method the best for real-time IPS response.

Mathematical Representation:

$$M_i = \text{Match}(R_i, T)$$

If $M_i \neq \emptyset$, then packet is marked as an intrusion.

This module serves as the primary module for known attacks and the secondary module in case the ML model fails.

2.Random Forest (Supervised IDS)

Random Forest uses the NSL-KDD dataset for training and classifies known attacks like DoS, Probe, R2L, and U2R. The classification of these attacks is done using majority voting. Random Forest works well for dealing with a large number of features and is resistant to noise. It provides accurate results for known types of attacks.

Mathematical Representation:

$$C = \text{Mode}(T1(x), T2(x), \dots, Tn(x))$$

3.Isolation Forest (Unsupervised Anomaly IDS)

This algorithm identifies unknown attacks, referred to as zero-day attacks, based on the identification of abnormal patterns. Malicious packets behave differently from the normal traffic patterns found in the network. This is identified using the number of splits that are required to isolate the instance.

Anomaly score:

$$s(x) = 2^{(-E(h(x)) / c(n))}$$

4.NLP-Based Log Analysis (TF-IDF, Word2Vec, Autoencoder)

This module analyzes system logs, firewall logs, and network event logs using NLP techniques. The logs are converted into vector representations using TF-IDF and Word2Vec, and then an autoencoder is applied to the vector representations to find anomalies.

Steps:

- 1.Log tokenization
- 2.Vectorization using TF-IDF and Word2Vec
- 3.Reconstruction error calculation

Mathematical representation:

$$E = |X - X'|$$

If $E > \text{threshold}$, the log entry is considered anomalous.

5.Clustering-Based Detection (DBSCAN and K-Means) This module uses clustering to group network flows and identify any anomalies that could be considered attacks.

K-Means:

$$\mu_i = (1 / |C_i|) \sum (x \in C_i) x$$

Points that are not close to the centroids are marked as anomalies.

DBSCAN:

Points that are not core points are marked as attacks if they are of low density.

6.Fusion Layer (Confidence-Weighted Voting)

All the models produce output with a confidence level. The final decision is taken by using a weighted voting approach. $FinalLabel = \text{argmax}_y \sum (w_i * P_i(y))$
 This approach leads to a stable and reliable decision for the intrusions.

7. Intrusion Prevention Engine (IPS)

Once the intrusion has been confirmed, the IPS module:

- Blocks the malicious IPs
 - Drops the packets
 - Closes the ports
 - Updates the firewall rules programmatically
- This ensures that the damage is contained and the response is immediate. Autoencoder)

This module analyzes the system logs, firewall logs, and network events using NLP techniques. It converts the logs to embeddings using TF-IDF and Word2Vec, and then uses an autoencoder to detect anomalies.

V. RESULTS AND ANALYSIS

A synthetic and benchmark-supported dataset was utilized for evaluating the proposed AI-based Intrusion Detection and Prevention System (IDS + IPS) solution. Real-time traffic was generated using the Scapy and PyShark tools for simulating regular and malicious traffic, including TCP connections, ICMP floods, port scans, and HTTP communication. At the same time, the NSL-KDD dataset was utilized for providing a validated dataset for DoS, Probe, R2L, and U2R types of cyber attacks. This ensures that the proposed system was tested under realistic conditions. All the packets and data entries were stored using the MongoDB database for uniform access and reproducibility of the experiments. Feature normalization, one-hot encoding, removal of corrupted packets, and balancing of attack types were performed for preprocessing the data, ensuring high-quality data for the model and avoiding bias.

The detection pipeline utilized two major machine learning models, namely the Random Forest, a supervised classifier, and the Isolation Forest, an unsupervised anomaly detector. The machine learning models were evaluated in isolation and in a combined fusion framework. The performance metrics used in evaluating the machine learning models include accuracy, recall, false positive rate, inference time, and stability over various test batch sets.

Metric	Random Forest	Isolation Forest
Accuracy	97.8%	92.4%
Precision	96.9%	91.2%
Recall	97.3%	89.8%
F1-Score	97.1%	90.4%
False Positive Rate	2.1%	3.7%

The accuracy level attained by the Random Forest model was 97.8%, with a corresponding recall level of 97.3%, indicating its effective performance with the identification of structured attacks. The false-positive rate was still low at 2.1%, which is a testament to the model’s reliability.

The Isolation Forest model attained an accuracy level of 92.4%, with a corresponding recall level of 89.8%, indicating

its effective performance with the identification of rare anomalies. However, the false-positive rate was a bit higher. The accuracy level attained by the hybrid fusion model was the highest at 97.8%, with a corresponding recall level of 97.3%, and a low false-positive rate of 1.4%, indicating the effectiveness of the fusion model.

SIMULATED ATTACK RESULTS

Test the real-time IDS + IPS feature, four different attacks were simulated several times using command-line tools:

- Smurf Attack (ICMP Flood)
- IP Sweep Attack
- Neptune SYN Flood
- Port Sweep Probe

All the attacks were performed several times on a Windows target machine, and the performance was tested based on the detection and prevention capabilities, along with the activation of the firewall rules.

Scenario	Avg Detection Time (ms)	Avg Blocking Time (ms)
Normal Traffic	85	-
Mixed Traffic	110	120
DoS Flood	130	145

Attack Type	Packets Sent	Attacks Detected	False Alarms (FP)	Detection Rate
Smurf (ICMP Flood)	10,000	9760	34	97.6
Neptune (SYN Flood)	8000	7640	22	95.5
IP Sweep	2540	2430	18	95.7
Port Sweep	102400	99010	126	96.8
Normal Traffic	5000	12	-	-

CONFUSION MATRIX (COMBINED RESULTS)

Actual/Predicted	Attack	Normal
Attack	10090	430
Normal	58	4922

REAL-TIME ATTACK DETECTION AND PREVENTION

The system excelled in real-time prevention, effectively preventing all Smurf and Neptune attacks and providing automatic Windows Firewall rules for malicious IP addresses. The IPS response time was less than 100 ms, which was indicative of the system’s excellent suitability for high-speed

environments. Port Sweep and IP Sweep attacks were detected at slightly lower accuracy levels (>95%) due to low-volume probes.

MODEL BEHAVIOR ANALYSIS

Random Forest

- Best for identifying known, structured attacks
- Low false positive rate
- Most stable performance for repeated tests (std. dev. = 0.0073)

Isolation Forest

- Best for identifying unknown anomalies
- More prone to irregular but benign traffic
- Highest false positive rate but better flexibility Hybrid model

Hybrid Model

- Highest Accuracy and Stability
- Lowest FPR
- Best for dynamic network like real-world networks
- Provides balanced detection for both known and unknown threats.

IPS PERFORMANCE AND SCALABILITY

The IPS module was tested for rule creation, blocking speed, and capacity.

The system was capable of handling more than 20,000 packets per minute without any delay.

The execution of detection to prevention averaged 61 ms, which is real-time blocking.

The web dashboard had low latency and was successful in displaying all alerts and IP addresses.

VII.CONCLUSION AND FUTURE SCOPE

CONCLUSION

From the results analysis, it is clear that the proposed system for AI-based IDS + IPS using the hybrid model approach is capable of accurate and real-time detection and prevention of intrusions. The results show that the proposed system outperforms other systems. Based on the detection accuracy of 97.6%, response time of 61 ms, and robustness of the system, it is clear that the proposed system is scalable and suitable for a real-world environment.

The proposed AI-based Intrusion Detection and Prevention System (IDS + IPS) offers a very adaptive, intelligent, and multi-layered security infrastructure that can detect known and unknown types of cyber attacks. By using supervised learning techniques such as Random Forest, unsupervised learning techniques such as Isolation Forest, signature-based techniques, and statistical analysis, the proposed system can effectively improve the accuracy, reliability, and generalization of the detection process.

The proposed hybrid approach can allow the system to analyze the patterns of attacks, such as the signatures of attacks, that are not usually detectable by traditional IDS.

The effectiveness of the suggested solution is supported by the results of the experimental evaluation, which used real-time packets capture techniques like Scapy and PyShark, and the NSL-KDD dataset. As shown in the evaluation, the suggested Hybrid Model, which combines RF and IF, is more effective, as it obtains a higher accuracy rate of 97.6%, a higher recall rate of 96.8%, and a lower false positive rate of merely 1.4%,

compared to traditional standalone machine learning classifiers. In particular, the Random Forest model ensures robust classification of known attacks like DoS, Probe, R2L, and U2R, while the Isolation Forest model provides robust zero-day anomaly detection by identifying abnormal patterns of traffic. In addition, the model's inference time is relatively low, at less than 50 milliseconds, and is highly stable, as evidenced by the evaluation of several batches.

The automatic IPS layer adds to the security of the system, as it can automatically block malicious IP addresses, drop suspicious packets, and apply firewall rules without the need for a human. This end-to-end solution, which covers the entire gamut of packet capture, feature engineering, multi-model fusion, and mitigation, ensures that the system is not only reactive but also proactive in providing security to the network with zero latency. Overall, the system is highly deployable, scalable, and reliable, making it a great platform for the development of next-generation intelligent cybersecurity solutions.

Future Scope

Even though the proposed system is a powerful and efficient tool for AI-based network defense, there are certain enhancements that could improve the system in future versions:

1.Integration of Transformer-Based Deep Learning Models

Future versions of this system could utilize advanced models such as BERT-IDS, Graph Neural Networks (GNNs), DeBERTa, or GPT-based models to detect advanced threats. These models could potentially identify more complex relationships between sequences of packets and flows in the network.

2.Deployment of Federated and Distributed Learning

In order to promote privacy and avoid the centralization of sensitive traffic log information, Federated Learning (FL) can be employed to train intrusion detection models across nodes without the need to share the original information. This enables collaboration among organizations, ISPs, and/or distributed servers to effectively collaborate in intrusion detection while maintaining the privacy of the information.

3.Expansion to IoT, Edge Devices, and 5G Networks

With the IoT and 5G devices producing a huge amount of high-velocity traffic, future systems should facilitate:

- Lightweight edge-optimized models of IDS
- On-device anomaly detection
- Extremely low latency real-time inference

This will enable security for smart homes, autonomous devices, and IoT systems from evolving threats.

4.Integration of Blockchain for Secure Logging

For example, blockchain-based logging can be used for the development of tamper-proof audit trails for intrusion events, firewall updates, and IPS actions.

This increases transparency, aids digital forensics, and ensures data integrity, which is critical for government, finance, and other cyber operations.

5.Self-Learning and Reinforcement-Based IPS

Reinforcement learning (RL) can be integrated into the IPS to dynamically enhance the decisions made by the IPS. Over time, the IPS will be able to learn:

- Optimal blocking strategies
- Risk-based packet filtering
- Adaptive firewall modifications

This will convert the IPS into a fully autonomous security agent.

6.Cloud-Native Monitoring Dashboard

A further extension in the future would be to include a monitoring dashboard that displays all the information in a

graphical format, such as:

- Alerts in real-time
- Levels of anomalies
- Origins of attacks
- Traffic heatmaps
- IPS action logs

7.Explainable AI (XAI) for Transparent Security Decisions
Integrating Explainable AI will enable the IDS to explain why a particular packet, IP address, or traffic pattern is identified as malicious. This will enhance:

- Trust in AI-based security
 - Model interpretability
 - Regulatory compliance (GDPR, ISO 27001, NIST guidelines)
- Explainable AI techniques like SHAP, LIME, or attention-based visualization can be integrated in future.

REFERENCES

[1]S. M. Rabiou, B. K. Aminu, and D. A. Zubairu, "AI-Driven Network Intrusion Detection Systems: A Systematic Review of Hybrid Models, Zero-Day Attack Mitigation, and Emerging Challenges in Cybersecurity," *Int. J. Comput. Appl.*, vol. 187, no. 8, pp. 27–33, May

[2]T. Sowmya, "A Comprehensive Review of AI-Based Intrusion Detection Systems," *Comput.*

Sci. Rev., vol. 23, pp. 100–115, 2023. ScienceDirect

[3] H. M. Rai, "Advanced AI-Powered Intrusion Detection Systems in Modern Cybersecurity," *Procedia Comput. Sci.*, vol.

[4] G. Olaoye, "AI-Driven Intrusion Detection and Prevention Systems: Architecture, Algorithmic Advancements, and Operational Efficiency," *SSRN Electron. J.*, 2025. SSRN

[5] A. J. A. Immastephy, "A Systematic Review on Network Intrusion Detection Systems: Techniques, Datasets, and Challenges," *E3S Web Conf.*, vol. 140, p. 14006

[6] A. J. A. Immastephy, "A Systematic Review on Network Intrusion Detection Systems: Techniques, Datasets, and Challenges," *E3S Web Conf.*, vol. 140, p. 14006, 2024. E3S Conferences

[7] S. K. R. Mallidi, "Advancements in Training and Deployment Strategies for AI-Based Intrusion Detection Systems in IoT Environments," *Comput. Netw.*, vol. 202, p. 107356, 2025. SpringerLink

[8] V. Sharma, "Artificial Intelligence-Based Intrusion Detection Systems: Techniques and Applications," *ITM Conf. Proc.*, vol. 8, p. 04002, 2024. itm-conferences.org

[9] Z. K. Maseer, "Meta-Analysis and Systematic Review for Anomaly Network Intrusion Detection Systems: Detection Methods, Dataset, Validation Methodology, and Challenges," *arXiv Prepr. arXiv2308.02805*, 2023. arXiv

[10] P. Wu and H. Guo, "LuNet: A Deep Neural Network for Network Intrusion Detection," *arXiv Prepr. arXiv1909.10031*, 2019. arXiv

[11] A. Pinto, J. Silva, and L. Silva, "Survey on Intrusion Detection Systems Based on Machine Learning," *MDPI Sensors*, vol. 23, no. 5, p. 2415, 2023. [Online]. Available: <https://www.mdpi.com/1424-8220/23/5/2415>.

[12] M. Rahman, M. U. Iqbal, and M. A. Hossain, "A

Survey on

Intrusion Detection System in IoT

Networks,"*ScienceDirect*, 2025. [Online]. Available:

[13] R. Kimanzi, "Deep Learning Algorithms Used in Intrusion Detection Systems: A Review," *arXiv Preprint*, 2024. [Online]. Available: <https://arxiv.org/abs/2402.17020>.