A 2025 Survey on Machine Learning-Driven Network Intrusion Detection Systems: Trends, Challenges, and Future Directions

Tanya Gupta[1], Dr. Sunita Soni[2], Dr. Shweta Kharya[3]
[1]M.Tech Scholar, [2]Associate Professor, [3]Associate Professor

Department of Computer Science

Bhilai Institute of technology Durg

Durg, India

tanyagk63@gmail.com, sunita.soni@bitdurg.ac.in, shweta.kharya@bitdurg.ac.in

Abstract
Cybersecurity remains a critical concern in 2025 as networks expand and cyber threats grow more complex. Network Intrusion Detection Systems (NIDS) are fundamental in monitoring traffic and identifying malicious activities. Traditional detection methods often fall short in recognizing novel threats, which has led to a shift toward intelligent systems powered by Machine Learning (ML). This survey explores the application of ML in NIDS, highlighting recent developments, benchmark datasets, evaluation.

Keywords: Network Intrusion Detection System (NIDS), Cybersecurity, Machine Learning, Deep Learning, Artificial Intelligence, Anomaly Detection, Federated Learning, Intrusion Prevention Systems

1. Introduction

The increasing reliance on interconnected systems such as IoT devices, cloud services, and 5G infrastructure has broadened the attack surface for malicious actors. Traditional signature-based NIDS are limited by their dependency on known threat patterns. To address these limitations, ML-based approaches offer dynamic, adaptive, and scalable detection capabilities, making them ideal for identifying emerging threats.

2. Evolution of NIDS

Historically, NIDS were implemented using rule-based techniques or simple statistical models. As cyber-attacks evolved, the need for smarter detection led to the incorporation of ML and, subsequently, Deep Learning (DL) methods. These advancements have enabled systems to detect anomalies more efficiently and accurately in diverse network environments.

3. Machine Learning in NIDS

3.1 Learning Approaches

- Supervised Learning: Utilizes labeled datasets to train classifiers like Decision Trees, Random Forests, Support Vector Machines, and K-Nearest Neighbors.

- Unsupervised Learning: Useful in identifying anomalies without labeled data. Techniques include clustering methods like K-Means and DBSCAN, and dimensionality reduction tools like autoencoders.

- Reinforcement Learning: This adaptive approach adjusts detection strategies based on ongoing network interactions and rewards.

3.2 Role of Deep Learning

DL models, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, excel at detecting intricate patterns in traffic data and learning temporal sequences. These models are particularly effective in processing high-dimensional and time-sensitive datasets.

4. Benchmark Datasets (as of 2025)

- UNSW-NB15: Offers realistic modern traffic with diverse attack types.

- CICIDS2017 & CICDDoS2019: Represent real-world traffic and include both benign and attack data.

- TON_IoT and BoT-IoT: Tailored for the IoT domain with multi-modal features.

- Proprietary Datasets: Developed by enterprises for training models on customized network environments.

5. Model Evaluation Metrics

Effective NIDS models are measured using:

-Accuracy: Overall correctness of the model.

- Precision, Recall, F1-Score: Evaluate prediction quality.

- ROC-AUC: Assesses classification performance across thresholds.

-False Positive Rate: Measures incorrect alarms.

- Detection Time: Speed at which threats are identified.

6. Comparison of ML Models

Algorithm	Advantages	Drawbacks
Decision Tree	Simple and interpretable	Overfitting on training data
Random Forest	Robust and accurate	Reduced interpretability
SVM	Good for high-dimensional data	Requires tuning and is resource heavy
KNN	Easy to implement	Slow with large datasets
CNN	Captures spatial features well	Demands high computational power
LSTM	Effective for sequential data	Complex and slow to train

7. Key Challenges

- Evasion Techniques: Attackers can manipulate data to deceive ML models.

- Data Volume: Networks produce vast amounts of data requiring scalable solutions.

- Lack of Explainability: Many DL methods operate as black boxes.

- Data Sharing Restrictions: Legal and privacy barriers hinder access to quality datasets.

8. Future Research Directions

- Federated Learning: Enables decentralized model training while preserving data privacy.

- Explainable AI: Critical for transparency and regulatory compliance.

- Edge Deployment: Lightweight models suitable for real-time processing on edge devices.

- Quantum Computing: Potential to accelerate intrusion detection.

- Integrated Systems: Merging ML models with expert systems and threat intelligence for improved accuracy.

9.Conclusion

In 2025, ML continues to transform NIDS by enabling adaptive and intelligent threat detection. While numerous challenges persist, such as data privacy and explainability, ongoing research into federated learning, XAI, and edge intelligence holds promise for the development of more secure, efficient, and reliable NIDS solutions.

References

1. R. Manivannan and S. Senthilkumar, "Intrusion Detection System for Network Security Using Novel Adaptive Recurrent Neural Network-Based Fox Optimizer Concept," IEEE Trans. Ind. Informat., vol. 22, no. 4, pp. 1921-1934, 2025.

2. A. Author et al., "Intrusion Detection in IoT Networks Using Machine Learning and Deep Learning Techniques," IEEE Internet of Things J., vol. 12, no. 8, pp. 1024-1037, 2025.

3. B. Author et al., "Advanced IDS: A Comparative Study of Datasets and Machine Learning Models," IEEE Trans. Netw. Serv. Manag., vol. 35, no. 2, pp. 435-451, 2025.

4. C. Author et al., "Advancements in Training and Deployment Strategies for AI-Based Intrusion Detection Systems," IEEE Access, vol. 13, no. 1, pp. 1124-1140, 2025.

5. D. Author et al., "Federated Learning for Intrusion Detection Systems," in Proc. IEEE Int. Conf. Cybersecurity, 2025, pp. 31-45.