International Conference on Advances in Modern Technology of Research in Engineering Field (AIMTREF) April, 2026
Organized By: SKN Sinhgad Institute of Technology& Science, Kusgaon (Bk), Lonavala, Pune

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

NetSentinel: A Machine Learning-Based Intrusion Detection Framework Using Random Forest

Associate Professor C.P. Lachake , Maitray Rangari, Sonawane Gaurav Vishwas, Ritik Palve, Shailesh Bhise

Department of Computer Engineering, SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra

Abstract- This paper presents NetSentinel, a machine learning-based intrusion detection framework designed to identify malicious network traffic with high accuracy. The system utilizes a Random Forest classifier trained on the NSL-KDD dataset using an 80:20 train-test split. The proposed architecture integrates a Python-based machine learning pipeline with a Node.js backend and MongoDB database for efficient data handling and user interaction. Performance evaluation using metrics such as accuracy, precision, recall, and F1-score demonstrates strong detection capability. Additionally, confusion matrix and ROC curve analysis validate the robustness of the model. The system aims to provide a scalable, efficient, and reliable cybersecurity solution. Future enhancements include real-time traffic capture, streaming integration, and deployment in live network environments.

Keywords- Intrusion Detection System (IDS), Random Forest, NSL-KDD Dataset, Machine Learning, Cybersecurity, Network Security, Anomaly Detection

I. INTRODUCTION

With the rapid growth of digital infrastructure, cybersecurity threats have become increasingly sophisticated and frequent. Intrusion Detection Systems (IDS) are essential for monitoring network traffic and detecting unauthorized access or malicious activities. Traditional IDS solutions rely heavily on signature-based detection methods, which are limited in identifying new or unknown attacks.

1. Limitations of Traditional IDS

- Signature-based systems fail to detect zero-day attacks
- High dependency on predefined attack patterns
- Poor adaptability to evolving threats
- High false positive and false negative rates
- Lack of real-time intelligent decision-making

2. Proposed Machine Learning-Based Solution

- The NetSentinel leverages Random Forest, an ensemble learning method
- Uses NSL-KDD dataset for improved training reliability
- Capable of detecting both known and unknown attacks
- Integrates backend (Node.js) + database (MongoDB)
- Provides scalable and efficient intrusion detection

This approach enhances detection accuracy and reduces dependency on manual rule creation.

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

II. PROBLEM STATEMENT

Modern network environments face continuous threats from malicious attacks such as DoS, Probe, R2L, and U2R. Traditional intrusion detection systems are insufficient due to their inability to adapt to evolving attack patterns. Additionally, many existing systems suffer from high false alarm rates and require significant computational resources.

Therefore, there is a need to design a machine learning-based intrusion detection system that:
- Accurately classifies network traffic
- Reduces false positives and false negatives
- Works efficiently on benchmark datasets
- Can be extended to real-time environments

III. METHODOLOGY

1. Dataset Used
- NSL-KDD dataset (improved version of KDD Cup 99)
- Eliminates redundant records
- Provides balanced and reliable training data

2. Data Preprocessing
- Data cleaning and normalization
- Feature selection and encoding
- Implementation using Pandas and NumPy

3. Model Implementation
- Algorithm: Random Forest Classifier
- Library: Scikit-learn
- Dataset split: 80% training, 20% testing

4. System Architecture
- Machine learning model built in Python
- Backend developed using Node.js
- Database: MongoDB
- Modular pipeline for scalability

IV. RESULTS AND ANALYSIS

The proposed system achieved strong classification performance on the NSL-KDD dataset.
- High accuracy, precision, recall, and F1-score
- Effective classification of normal vs malicious traffic.

As shown in the Confusion Matrix
- True Positives and True Negatives are significantly high
- Very low misclassification (False Positives/Negatives).

The ROC Curve demonstrates
- Strong separation between classes
- High Area Under Curve (AUC), indicating excellent model performance.

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

Feature Importance Graph highlights
- Key features contributing most to intrusion detection
- Helps in dimensionality reduction and optimization.

V. DISCUSSION

The results clearly indicate that Random Forest is highly effective for intrusion detection tasks. Compared to traditional approaches, the model provides:
- Better generalization
- Reduced overfitting
- Improved detection of unseen attacks

However, limitations include:
- Dependency on dataset quality
- Lack of real-time implementation (currently offline).

VI. CONCLUSION

NetSentinel successfully demonstrates an efficient IDS using machine learning
- Random Forest provides high detection accuracy and robustness
- The system reduces false alarms and improves reliability
- Integration with backend and database enhances usability

Overall, NetSentinel offers a scalable, cost-effective, and intelligent cybersecurity solution.

REFERENCES

1. D. Denning, 'An Intrusion Detection Model,' IEEE Transactions on Software Engineering, 1987.
2. S. Axelsson, 'Intrusion Detection Systems: A Survey and Taxonomy,' Technical Report, 2000.
3. L. Breiman, 'Random Forests,' Machine Learning Journal, 2001.
4. T. Tavallaee et al., 'A Detailed Analysis of the KDD Cup 99 Dataset,' IEEE CISDA, 2009.
5. M. Roesch, 'Snort: Lightweight Intrusion Detection,' USENIX, 1999.
6. K. Kendall, 'A Database of Computer Attacks,' MIT Lincoln Lab, 1999.
7. J. McHugh, 'Testing Intrusion Detection Systems,' ACM TISSEC, 2000.
8. S. Mukkamala, 'Intrusion Detection using Neural Networks,' IEEE, 2002.
9. R. Sommer, 'Outside the Closed World: Machine Learning for IDS,' IEEE S&P;, 2010.
10. G. Creech, 'Generation of a New IDS Dataset,' Military Communications Conference, 2013.
11. M. Ring, 'A Survey of Network-based Intrusion Detection Data Sets,' Computers & Security, 2019.
12. Javaid, 'Deep Learning for Network Intrusion Detection,' IEEE MILCOM, 2016.
13. W. Lee, 'Data Mining Approaches for IDS,' USENIX Security, 1998.
14. P. Laskov, 'Learning Intrusion Detection,' Springer, 2005.
15. H. Debar, 'Intrusion Detection Systems,' IEEE Security & Privacy, 2007.
16. S. Garfinkel, 'Network Intrusion Detection,' Addison-Wesley, 2002.
17. J. Brownlee, 'Machine Learning Mastery,' 2016.
18. Goodfellow, 'Deep Learning,' MIT Press, 2016.
19. R. Bace, 'Intrusion Detection,' Macmillan, 2000.
20. E. Eskin, 'Anomaly Detection over Noisy Data,' ICML, 2000.
21. Patcha, 'Overview of Anomaly Detection Techniques,' Computer Networks, 2007.
22. N. Hubballi, 'Network Intrusion Detection,' IEEE Communications Surveys, 2014.

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

23. S. Stolfo, 'Data Mining-based IDS,' Columbia University.