A Novel Approach for Intrusion Detection in 5G Networks using Deep Learning

Ira A. Fulton Schools of Engineering
Arizona State University

Amit Noel Thokala, Vishwanath Reddy Yasa, Harsha Vardhan Yallavula, Hari Priya Gottam, Yashwanth Reddy Kikkuri, Sai Rithvik Vaikuntam

Abstract

This research proposes a novel Intrusion Detection System (IDS) for 5G mobile networks to address the escalating cyber-attack threats from IoT devices. The study emphasizes the importance of robust IDS to protect network integrity, ensure data confidentiality, and mitigate security breaches. Key challenges such as data preprocessing, feature engineering, model selection, and scalability are identified and tackled through a multi-pronged approach including data preprocessing techniques, machine learning algorithms, and ensemble learning models. Evaluation on the 5G-NIDD dataset demonstrates high accuracy in detecting various attack types. This research contributes to cybersecurity by offering a comprehensive IDS solution for 5G networks.

Introduction

The advent of 5G networks and the widespread adoption of Internet of Things (IoT) devices have introduced new challenges and heightened the importance of effective intrusion detection and prevention systems. The escalating threat landscape of cyber-attacks underscores the critical need for robust security measures, particularly in the context of 5G networks. Researchers and practitioners are turning to machine learning and deep learning techniques to develop advanced intrusion detection systems capable of effectively identifying and mitigating cyber threats in 5G networks. By leveraging data-driven analytics and predictive modeling, these systems aim to enhance network security, preemptively detect anomalous behavior, and mitigate the impact of cyber-attacks.The effectiveness of our approach is demonstrated through evaluation on the 5G-NIDD dataset, showcasing high accuracy in both binary and multi-class classification tasks for detecting various attack types.

Methodology

Data Acquisition: Utilized the 5G-NIDD dataset for training and testing our intrusion detection models. The dataset includes network traffic categorized as benign or malicious, with malicious instances further classified into specific attack types (e.g., HTTP Flood, ICMP Flood).

Model Development:

Binary Classification: Developed models to distinguish between benign and malicious traffic. Applied various machine learning algorithms including CNNs, RNNs, and ensemble methods such as Voting Classifier and Boosting Techniques.

Multi-Class Classification: Extended the classification to identify specific types of network attacks. Employed both hard and soft voting ensemble models to improve reliability and predictive capabilities.

Evaluation Metrics: Assessed model performance using accuracy, precision, recall, and F1-score metrics. Conducted thorough validations to ensure robustness in diverse network scenarios.

Implementation of Ensemble Techniques:

Voting Classifier: Integrated multiple models (AdaBoost, CatBoost, SVM, LightGBM, and XGBoost) to leverage their collective strengths in a hard voting scheme.

Ensemble Model: Combined different classifiers (Decision Tree, Random Forest, Logistic Regression, KNN, MLP, and GBM) to enhance detection accuracy and minimize false positives.

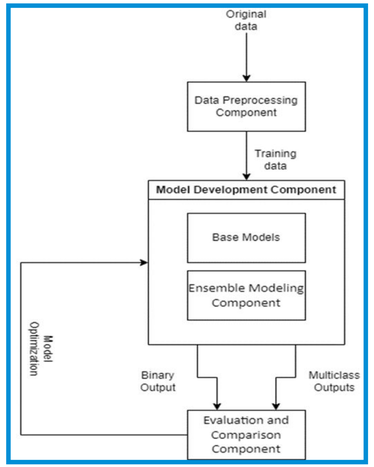

Fig 1. High Level Architecture

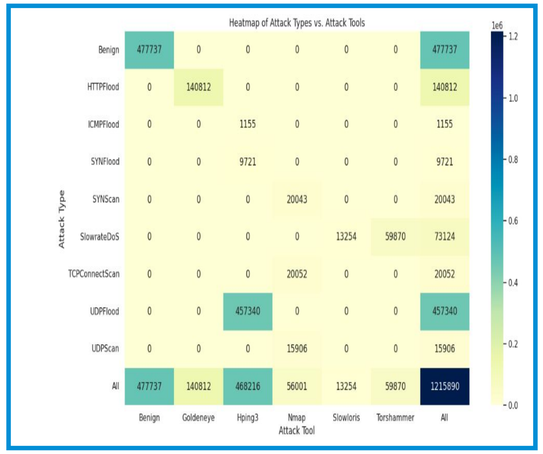

Fig 2. Heat Map for Attack Tool vs Attack Type

Results

Binary Classification Performance:

Voting Classifier: Achieved the highest accuracy of 99.96% with precision and recall both exceeding 99.9%, demonstrating exceptional efficacy in detecting network intrusions.

Ensemble Model: Showed nearly comparable performance, with an accuracy of 99.93% and consistently high precision and recall scores.

Individual Models: CNN and RNN models also demonstrated robust performances with accuracies of 99.95% and 99.81% respectively.

Multi-Class Classification Performance:

Precision and Recall: Across attack types, models displayed high precision and recall, particularly in categories like ICMPFlood and SYNFlood, where precision reached 100% in several models.

F1-Scores: Indicative of a balanced model, F1-scores were generally high, reflecting effective harmony between precision and recall across all models. Notably, the ensemble models and CNN achieved impressive scores across multiple attack types.

Models	Accuracy	Precision	Recall	F1 score
Boosting LSTM	0.9973	0.9975	0.9972	0.9973
Hard voting classifier	0.9992	0.9977	0.998	0.9978
Soft voting classifier	0.9992	0.9981	0.9989	0.9985
CNN	0.9984	0.9984	0.9983	0.9984
RNN	0.9981	0.9981	0.9981	0.9981

Table 1. Accuracies for Multi Classification Models

Models	Accuracy	Precision	Recall	F1-score
Boosting Classifier (AdaBoost, CatBoost, LightGBM, XGBoost)	0.9996	0.9993	0.9998	0.9996
Hard voting Model (Decision Tree, Random Forest, Logistic Regression, KNN, MLP, GBM)	0.9992	0.9991	0.9990	0.9991
Soft voting Model (Logistic Regression, MLP, GBM)	0.9991	0.9993	0.9984	0.9992
CNN	0.9994	0.9996	0.9993	0.9995
RNN	0.9981	0.9981	0.9981	0.9981

Table 2. Accuracies for Multi Binary Classification Models

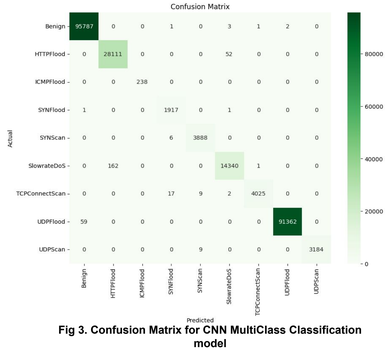

Fig 3. Confusion Matrix for CNN MultiClass Classification model

Overall Model Efficacy:

Highest Performing Models: The CNN and ensemble models showed superior accuracy and F1-scores across both benign and attack classes, confirming their robustness and reliability in a diverse set of network intrusion scenarios.

Special Mention: The Boosting Classifier and Hard Voting Ensemble models consistently delivered high performance, underpinning the utility of ensemble strategies in complex classification tasks.

Conclusion

The study conclusively demonstrates that ensemble and deep learning models, especially the Voting Classifier and CNN, exhibit superior performance in the detection and classification of network intrusions within 5G networks. These models achieved nearly perfect accuracy rates in both binary and multi-class classifications, underlining their potential to serve as robust solutions for real-time security threats in advanced telecommunications systems. The use of ensemble methods, which integrate multiple learning algorithms, proved particularly effective, enhancing the predictive accuracy and reliability necessary for dealing with sophisticated cyber threats. These results highlight the critical role of advanced machine learning techniques in the development of next-generation intrusion detection systems that are capable of maintaining the integrity and security of 5G infrastructure.

Acknowledgements

We extend our gratitude to Dr. Abdallah Moubayed, our mentor throughout this project conducted under the course SER 517: Software Capstone. His guidance was invaluable in navigating the complexities of our research and achieving our objectives effectively