ML-based intrusion detection system for SDN-based broadband networks

Eya Grami
Department of Networks and Telecommunications.
Institute of Applied Science and Technology (INSAT).
Tunis, Tunisia
eya.grami@insat.ucar.tn

Fatma Bouhari
Department of Networks and Telecommunications.
Institute of Applied Science and Technology (INSAT).
Tunis, Tunisia
fatma.bouhari@insat.ucar.tn

Mohamed Malek Gharbi
Department of Networks and Telecommunications.
Institute of Applied Science and Technology (INSAT).
Tunis, Tunisia
mohamedmalek.gharbi@insat.ucar.tn

Ali Kallel
Department of Networks and Telecommunications.
Institute of Applied Science and Technology (INSAT).
Tunis, Tunisia
ali.kallel@insat.ucar.tn

Mohamed Karrab
Department of Networks and Telecommunications.
Institute of Applied Science and Technology (INSAT).
Tunis, Tunisia
mohamed.karrab@insat.ucar.tn

***Abstract*—The transition from traditional network architectures to Software-Defined Networks (SDNs) has revolutionized network flexibility and management. However, the centralization of control in SDNs introduces significant security challenges, particularly in the realm of intrusion detection. This research proposes a machine learning-based Intrusion Detection System (IDS) designed specifically for SDN environments. Utilizing the KDD CUP 1999 dataset, we develop a hybrid model that integrates Random Forest (RF), Support Vector Machine (SVM), and XGBoost classifiers to classify network traffic into categories such as Normal, DoS, Probe, R2L, and U2R attacks. Our approach enhances detection accuracy, minimizes false-positive rates, and provides a robust solution to safeguard SDN architectures against a diverse range of security threats.**

***Index Terms*—Software-Defined Networking, Intrusion Detection System, Machine Learning, KDD CUP 1999, Network Security, Multi-Class Classification.**

I. Introduction

With the rapid adoption of Software Defined Networks (SDNs) in recent years, the networking paradigm has undergone a revolutionary transformation. By decoupling the control and data planes, SDNs enable centralized management through a single controller, offering unparalleled flexibility, scalability, and efficiency in network operations. However, this architectural shift comes with significant security challenges. The centralized SDN controller, being the brain of the network, represents a single point of failure. Any unauthorized access or malicious activity targeting the controller could have catastrophic repercussions, compromising the integrity and security of the entire network. In the context of SDNs, the security of the central controller is paramount to ensuring the integrity and stability of network operations. As the brain of the network, the controller's vulnerability to unauthorized access or malicious activities poses significant risks, potentially compromising the entire network's functionality and security. To address these challenges, implementing effective Intrusion Detection Systems (IDS) has become a necessity. Traditional intrusion detection approaches often rely on binary classification, categorizing traffic as either normal or malicious. While effective for simpler threats, this approach falls short in addressing the complexity of modern attacks, such as Distributed Denial-of-Service (DDoS) assaults, malware infiltrations, and web-based exploits. The evolving threat landscape demands more sophisticated detection mechanisms capable of identifying and categorizing diverse attack types. In this research, we aim to tackle this issue by focusing on the development of an advanced intrusion detection system using machine learning (ML), specifically tailored for SDN environments. By leveraging the KDD CUP 1999 dataset, we design a hybrid model combining Random Forest (RF), Support Vector Machines (SVM), and XGBoost classifiers, enhanced by a meta-classifier to improve detection accuracy. Our model classifies traffic into five categories: Normal, DoS, Probe, R2L, and U2R. The aim is to move beyond traditional binary detection and implement a multi-class classifier capable of effectively identifying a broader range of attack types, thereby contributing to the protection of SDN-based networks against a wide variety of security threats. This paper makes three key contributions: a detailed analysis of existing IDS techniques, the design of a hybrid machine learning-based multi-class classifier, and an evaluation of its performance using the KDD CUP 1999 dataset. The remainder of the paper is organized as follows: Section II reviews related work ; Section III outlines the methodology and model development; Section IV discusses the experimental results and analysis; and Section V concludes with future research directions.LaTeX.

II. Related Work

A. Anomaly Detection in SDN Environments

Maxime Labonne [1] explored the application of machine learning algorithms for anomaly-based intrusion detection. His research demonstrated their ability to identify deviations from normal network behavior, enhancing security against unconventional threats. Similarly, Zhen Yang et al. [5] provided a comprehensive systematic review of anomaly-based intrusion detection methodologies and datasets. Their work offered valuable insights into state-of-the-art techniques and datasets, serving as a cornerstone for future advancements in anomaly detection.

B. Detecting DDoS Attacks with Advanced Techniques

In the context of Distributed Denial-of-Service (DDoS) attacks, Junhong Li [2] proposed an innovative approach combining dense neural networks, autoencoders, and the Pearson correlation coefficient. This method effectively identified DDoS patterns, showcasing how neural network architectures and correlation metrics can significantly improve detection accuracy. Naveen Bindra and Manu Sood [3] further emphasized the importance of feature selection in optimizing machine learning models for DDoS attack detection. Their research highlighted the critical role of preprocessing and feature extraction in enhancing the precision of IDS systems.

C. Hybrid and Flow-Based Intrusion Detection

K. Muthamil Sudar and P. Deepalakshmi [4] introduced a flow-based intrusion detection approach specifically designed for SDNs. By utilizing hybrid machine learning techniques, their study addressed the unique security challenges of SDN architectures. The integration of flow analysis with hybrid ML models proved effective in detecting a wide range of intrusions, underscoring the need for tailored solutions in dynamic SDN environments.

III. RESEARCH METHODOLOGY

A. Dataset

The provided dataset, which is central to our research, contains 4.9 million rows, each characterized by 41 distinct columns. A crucial component of the dataset is its solitary labelled column, which serves as the target variable for classification. The structure of this dataset reflects the intent to use machine learning techniques to predict or classify outcomes based on the interaction of various independent features.

This dataset is notable for its pronounced class disparity, where some attack classes are considerably underrepresented compared to others. This imbalance can present challenges for machine learning algorithms, potentially resulting in biased model performance in which the majority class (normal connections) dominates prediction accuracy. This disparity necessitates the application of specialized techniques to ensure that all classes are treated fairly during the model training process.

Some attack classes within the labelled column exhibit an extremely sparse representation, with only a minimal number of instances, adding to the complexity. This scarcity heightens the need for careful model management, as the limited number of instances of these classes may hinder the algorithm's ability to generalize effectively. Table I contains an overall summary of the dataset.

B. Data Prepossessing

The methodology implements a hierarchical classification system for network attacks. The attacks are categorized into four main classes plus normal traffic:

- **DoS (Denial of Service)**: Includes attacks such as neptune, smurf, and pod
- **R2L (Remote to Local)**: Encompasses httptunnel, ftp_write, and guess_passwd
- **U2R (User to Root)**: Contains buffer_overflow, rootkit, and loadmodule
- **Probe**: Includes ipsweep, nmap, and portsweep
- **Normal**: Represents legitimate network traffic

The preprocessing involves two main types of data transformation:

1) Numerical Data: Connection-related numbers (such as duration and data volume) are standardized to comparable scales. This ensures all numerical information contributes appropriately to the analysis.

2) Categorical Data: Information like connection types and services are converted into a format suitable for analysis. This transformation preserves the essential characteristics of different network activities.

C. Maintaining Data Quality

Throughout the preprocessing, we ensure:

- Consistent treatment of information across datasets
- Preservation of important network patterns
- Reliable transformation of all data types

D. Model Architecture

In this study, a sophisticated ensemble learning architecture is employed to enhance the predictive performance and robustness of the model. The architecture is designed to leverage multiple individual models to improve prediction accuracy and reduce overfitting by combining their strengths. The system consists of three base models: Random Forest (RF), Support Vector Machine (SVM), and XGBoost, each trained on the KDD dataset. These models are tasked with learning different aspects and patterns within the data, enabling them to capture a wider range of features and behaviors, crucial for detecting network intrusions effectively.

Once the base models are trained, their predictions are passed to the meta-model, which in this case is another Random Forest classifier. The role of the meta-model is to integrate the outputs from the three base models (RF, SVM, and XGBoost) and make the final prediction. This step harnesses the diverse strengths of each model, providing a weighted and aggregated prediction based on the individual outputs.

The meta-model is trained on a combination of the original features from the KDD dataset along with the predictions generated by the base models. The Random Forest algorithm, known for its robustness and ability to reduce overfitting, is used here as the meta-model to combine these predictions. By utilizing the majority voting mechanism of Random Forest, the meta-model makes the final decision, drawing on the ensemble's collective insights. This ensemble approach helps to reduce bias, improve generalization, and enhance the accuracy of the final prediction.

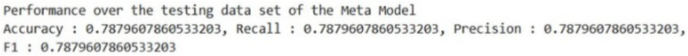
The ensemble approach offers several key advantages over individual models. By using multiple models, the system leverages diversity, allowing it to better handle complex relationships within the data. The independent learning process of each base model ensures that various facets of the dataset are explored, making the system more adept at identifying patterns in network intrusion data. Additionally, by combining bagging (Random Forest) and boosting (XGBoost) techniques, the system can effectively address issues like overfitting and underfitting, ensuring that it performs well on both the training data and unseen validation data.

Furthermore, the use of a Random Forest meta-model for combining base model predictions ensures that the final output is more robust. Random Forest reduces the risk of any individual model dominating the decision-making process by using random feature selection and majority voting, which ultimately enhances the stability and reliability of the predictions.

After the meta-model has made the final prediction, the results are evaluated using the validation data to assess the model's effectiveness in real-world scenarios. The model's performance is quantified using various metrics such as accuracy, precision, recall, and F1-score, which are essential for understanding how well the system identifies both attack and normal network connections. The validation phase helps fine-tune the model, ensuring that it is both accurate and capable of generalizing well to new, unseen data.

IV. Result Analysis

In our study, we evaluated the performance of our ensemble model in detecting different types of network traffic and security threats, classified into categories such as Normal, DoS, Probe, R2L, and U2R. The model was tested on the KDD99 dataset, using separate training and testing data to assess its effectiveness in both learning and generalizing to unseen data. The ensemble model, which integrates Random Forest (RF), Support Vector Machines (SVM), and XGBoost, achieved promising results. The F1-score on the testing dataset was 0.7879, indicating strong performance in correctly identifying most types of network traffic and intrusions. This was achieved by leveraging a meta-classifier, which aggregated predictions from the base models, thereby enhancing the overall accuracy and robustness.



```
Performance over the testing data set of the Meta Model
Accuracy : 0.7879607860533203, Recall : 0.7879607860533203, Precision : 0.7879607860533203,
F1 : 0.7879607860533203
```

Fig. 1. Performance Metrics of the Meta-Model.

A. Model Classification Results

The ensemble model effectively classified the traffic into the predefined categories: Normal Traffic DoS Attacks Probe Attacks R2L Attacks U2R Attacks

B. High Classification Accuracy

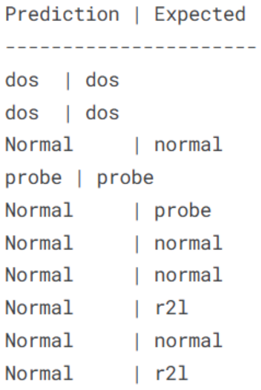
The ensemble model performed well in detecting common attack types such as DoS and Probe, with high precision and recall scores. The meta-classifier's ability to combine predictions effectively reduced misclassifications, demonstrating the advantage of ensemble learning

C. Class Imbalance Challenges

While the ensemble model addressed the class imbalance issue using data augmentation, rare attack types like R2L and U2R presented a challenge due to the limited number of samples in the dataset. This affected the model's ability to classify these categories with the same accuracy as more frequent ones.

D. Complex Attack Patterns

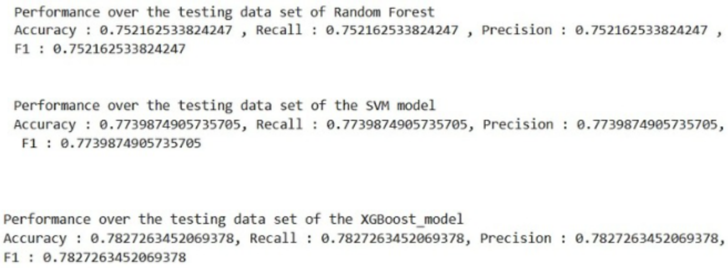
The model struggled to achieve high precision in detecting attacks with complex patterns, such as web-based exploits or intricate R2L behaviors. These challenges suggest a need for additional data or feature engineering to capture nuanced attack characteristics.



```
Prediction | Expected
----------------------
dos  | dos
dos  | dos
Normal     | normal
probe | probe
Normal     | probe
Normal     | normal
Normal     | normal
Normal     | r2l
Normal     | normal
Normal     | r2l
```

Fig. 2. Prediction vs. Expected Classification Results of the Meta-Model .

The meta-classifier provided an aggregated prediction, significantly enhancing the overall accuracy compared to individual models. For instance: The Random Forest achieved an accuracy of 75.2% , while The SVM reached 77.4% , and The XGBoost delivered 78.3% . When combined, the ensemble model improved upon these individual performances, achieving an overall accuracy of 78.8 with balanced precision, recall, and F1-score.



```
Performance over the testing data set of Random Forest
Accuracy : 0.752162533824247 , Recall : 0.752162533824247 , Precision : 0.752162533824247 ,
F1 : 0.752162533824247

Performance over the testing data set of the SVM model
Accuracy : 0.7739874905735705, Recall : 0.7739874905735705, Precision : 0.7739874905735705,
 F1 : 0.7739874905735705

Performance over the testing data set of the XGBoost_model
Accuracy : 0.7827263452069378, Recall : 0.7827263452069378, Precision : 0.7827263452069378,
F1 : 0.7827263452069378
```

Fig. 3. Performance Metrics of Base Models.

V. CONCLUSION

In terms of future work, several approaches can be applied to further enhance the performance and robustness of our ensemble model. First, addressing the handling of rare attack classes, such as U2R attacks (e.g., ”Buffer Overflow” and ”Loadmodule”) and R2L attacks (e.g., ”Guess Password” and ”FTP Write”), is essential. These attack types are significantly underrepresented in the KDD99 dataset, and focusing on methods to accurately detect these rare instances will improve the model’s ability to identify a broader range of attack scenarios.

Second, further exploration of feature engineering is critical. By identifying more discriminative features from the network traffic data, we can improve the model’s ability to capture complex patterns and enhance its overall performance. Additionally, techniques such as feature selection and dimensionality reduction could be investigated to remove irrelevant or redundant data, thereby improving training efficiency and model accuracy.

Third, addressing the class imbalance issue in the dataset is a crucial area of improvement. Many of the attack types, particularly Probe attacks (e.g., ”Portscan” and ”Satan”), are underrepresented in comparison to the normal connections class. Implementing techniques like oversampling, undersampling, or cost-sensitive learning could enable the model to learn more effectively from underrepresented classes, ensuring more equitable performance across all classes.

The optimization of threshold values and hyperparameters also represents a key area for future research. Fine-tuning these settings through comprehensive experiments could help achieve a better balance between precision and recall, particularly for rare attack categories. Exploring cross-validation techniques and optimizing model parameters through grid search or random search could further enhance performance.

Moreover, additional refinement of the ensemble model architecture could be considered by incorporating a variety of classifiers, such as neural networks or hybrid models. The integration of techniques like class-specific weighting or experimenting with different ensemble configurations could help address performance disparities among specific attack categories.

Lastly, the possibility of training separate models for each class or group of attack types (such as DoS, R2L, and U2R) could be explored. For example, DoS attacks (e.g., ”Smurf” and ”Teardrop”) could benefit from different modeling strategies than rare attack types. This tailored approach could help optimize the prediction accuracy for each class, further improving the ensemble model’s robustness.

In conclusion, while the current model provides promising results, these future directions hold significant potential for improving its detection capabilities, particularly for less common attacks in the KDD99 dataset.

REFERENCES

[1] Maxime Labonne. Anomaly-based network intrusion detection using machine learning. Cryptography and Security [cs.CR], Institut Polytechnique de Paris, 2020. English. NNT: 2020IPPAS011. Tel-02988296.

[2] Junhong Li. Detection of DDoS Attacks Based on Dense Neural Networks, Autoencoders, and Pearson Correlation Coefficient. Dalhousie University, April 2020.

[3] Naveen Bindra, Manu Sood. Evaluating the Impact of Feature Selection Methods on the Performance of the Machine Learning Models in Detecting DDoS Attacks. Romanian Journal of Information Science and Technology, Volume 23, Number 3, 2020, pp. 250–261.

[4] K. Muthamil Sudar, P. Deepalakshmi. Flow-Based Intrusion Detection System for Software Defined Networking Using Hybrid Machine Learning Techniques. International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume 9, Issue 2S2, December 2019.

[5] Zhen Yang et al : Recognized for conducting a comprehensive systematic review of anomaly-based intrusion detection methodologies and datasets. Their work serves as an essential reference in the field, offering an in-depth exploration of state-of-the-art techniques and datasets that advance the understanding and development of intrusion detection systems.