Poster Paper: Deep Learning driven Anomaly based Intrusion Detection System for IoT

Yue Guan[1]; Naser Ezzati-Jivan[1]
[1]Brock University

Abstract

Problem:
- The scale of devices connected to an IoT network is increasing, leading to novel vulnerabilities and anomalies.
- Existing methods are not optimally tuned and an end-to end system using machine and deep learning approaches is scarce in existing literature.
- Processing large number of features makes the system compute intensive and require massive memory and training time.

Poster paper solution:
- A hybrid framework of machine learning and deep learning networks for efficient classification of attacks and anomalies is presented.
- The system is optimally tuned, and appropriate feature selection and data balancing approaches are adopted.

Introduction

- Machine learning and deep learning-based approaches have been used to ensure the security and privacy of an IoT network.
- However, both domains are still at their infancy stage for anomaly detection and classification.
- Machine learning based algorithms can distinguish between normal or anomaly data (binary classification) easily, but due to the shallow nature of these algorithms, they are unable to detect or classify numerous types of attack being generated on the IoT network.
- On the other hand, deep learning-based algorithms can distinguish between different types of attacks (multi-classification) but are very compute intensive and require massive memory and training time.
- Moreover, they involve numerous hyperparameters, which needs to be tuned effectively in order to achieve efficient performance and efficacy.

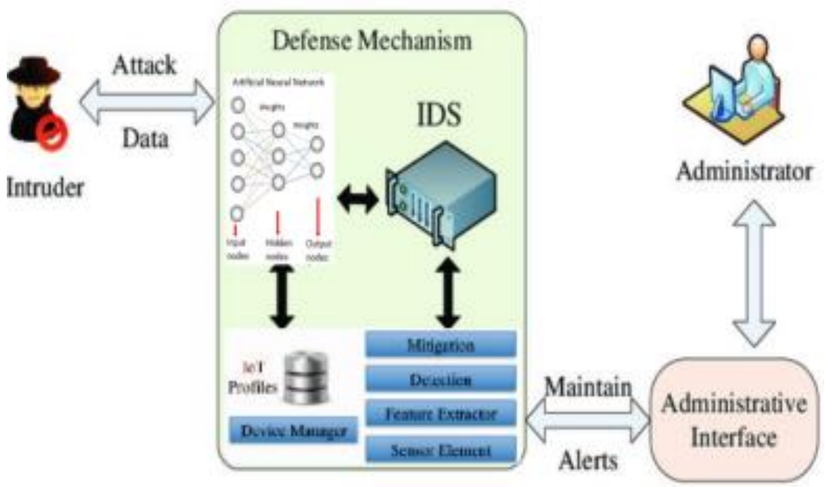

Figure 1. Overview of Proposed Intrusion Detection System.

Methodology

1. We propose a machine learning and deep learning based intrusion detection system (IDS) that analyzes network traffic and packets to classify any intrusion or attack.
   a. In order to check whether a network packet is normal or anomalous, we employ a machine learning based binary classifier. It has less complexity and requires less computation cost and training time[1].
   b. For multi-class classification, we have employed deep learning based recurrent neural network to not only identify whether it is anomaly traffic, but also to distinguish the type of attack being generated on the network.
2. Imbalance data can drastically affect the model's performance and can result in biased model.
   a. we have made the use of Synthetic Minority Oversampling Technique (SMOTE)[3] to balance the data.
   b. It does not generate duplicates, but rather creates synthetic data points that are slightly different from the original data points.
3. Appropriate feature selection helps to discard the irrelevant features, which does not plat any role in the learning of the model and saves execution time.
   a. we used PSO[2] feature selection technique that attempts to find the best subset of the input feature set.
4. We have also used the most optimal hyperparameter tuning of the intrusion detection system, including the loss function, optimizer, batch size, and epochs. It aids and assist the intrusion detection systems to achieve maximum performance and resource utilization.

DataSet

We have used IoTID20 dataset[4] for experiments in which 86 different features have been captured. The dataset contains different types of IoT attacks as well as their families. The dataset is available in CSV format and is free to be used for research.

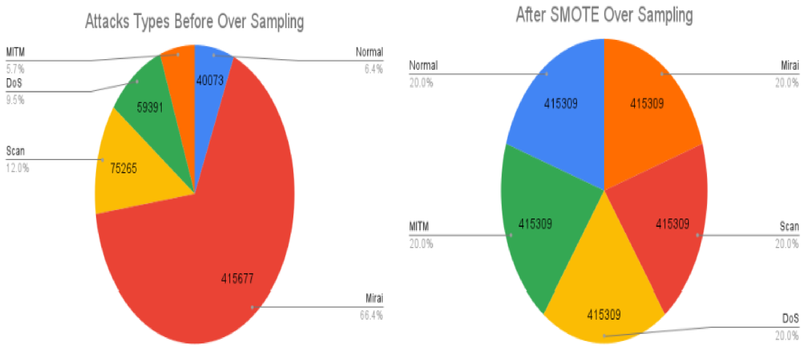

Figure 2. A Visualization of Data Imbalance

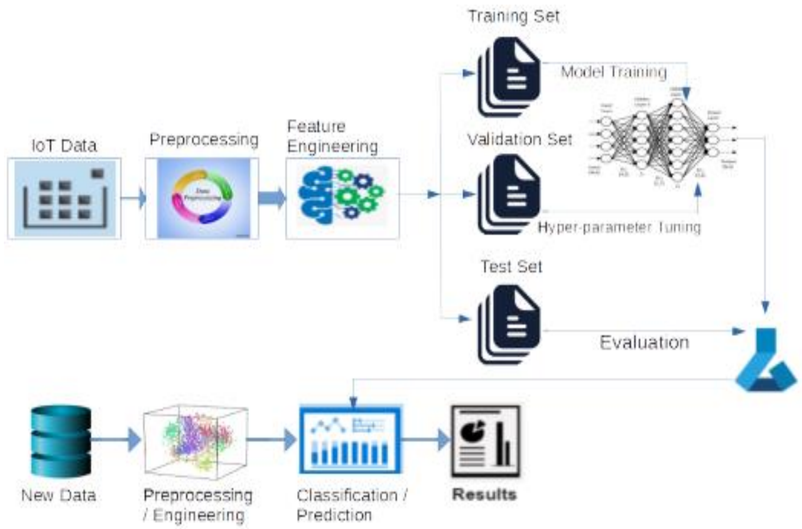

Figure 3. Intrusion Detection System Model Pipeline

Conclusions

- A hybrid framework to perform anomaly detection and classification is presented.
- The proposed framework is optimally tuned by selecting the most appropriate hyperparameter values including loss function, optimizer, batch size and epochs.
- We showed that feature selection and data balancing approaches such as SMOTE and SMO can improve performance and execution time of the overall system.
- In future we would like to extend this work by working on more complex datasets incorporating more real-life attacks, anomalies and challenges.

Future Work

- In future we would like to extend this work by working on more complex datasets incorporating more real-life attacks, anomalies and challenges.
- We would also like to work on other machine learning and deep learning-based models for anomaly classification and intrusion prevention.
- Another important aspect is to use the hardware accelerators such as GPU based cloud infrastructure for the efficient performance of the system.
- We will develop a high-performance platform which will be equipped with modern GPUs and will help to accelerate our proposed anomaly classification pipeline for intrusion detection and prevention.

Contact

Yue Guan
Brock University
1812 Sir Isaac Brock Way
rv18jq@brocku.ca

Naser Ezzati-Jivan
Brock University
1812 Sir Isaac Brock Way
nezzatijivan@brocku.ca

References

1. Ahmad, Zeeshan, Adnan Shahid Khan, Cheah Wai Shiang, Johari Abdullah, and Farhan Ahmad. "Network intrusion detection system: A systematic study of machine learning and deep learning approaches." *Transactions on Emerging Telecommunications Technologies*, 2021.
2. Zhao, Shanhui, Wanjun Xu, and Linghai Chen. "The modeling and products prediction for biomass oxidative pyrolysis based on PSO-ANN method: an artificial intelligence algorithm approach." *Fuel* 312, 2022.
3. Fernández, Alberto, Salvador Garcia, Francisco Herrera, and Nitesh V. Chawla. "SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary." *Journal of artificial intelligence research*, 2018.
4. Ullah, Imtiaz, and Qusay H. Mahmoud. "A scheme for generating a dataset for anomalous activity detection in iot networks." In *Canadian Conference on Artificial Intelligence*, pp. 508-520. Springer, Cham, 2020.