INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

University of Westminster, Coat of Arms

PhoeniXhield

A Hybrid Deep Learning Approach for DDoS Attack Classification using Network Traffic Data Analysis

A Dissertation by

Mr. Oshadha Goonathilaka

w1762649 / 2018402

Supervised by

Mr. Iresh Bandara

Submitted in partial fulfilment of the requirements for the BEng (Hons) in Software Engineering degree at the University of Westminster.

April 2023

Abstract

A distributed denial-of-service (DDoS) attack is a malicious operation that seeks to disrupt the legitimate flow of traffic to a server, service, or network by flooding it with overwhelming network traffic, rendering the target inoperable. This has been a great challenge due to frequently changing attack patterns, the rapid development of cyber offense tools, and the open availability of cyber offense materials in the dark web. DDoS attacks can cause severe disruption to online services, resulting in lost revenue, reputation damage, and decreased customer trust. DDoS attacks can also cause damage to infrastructure, such as overheating or power outages. By considering all these, this research focuses on classifying DDoS attacks by analyzing network traffic data using a hybrid deep learning approach to automate the manual process, which can ultimately save time and effort and reduce human errors in the detection process. This project was completed by combining the strengths of the RNN (LSTM) autoencoder and RNN (LSTM) multi-class classifier to train the "DDoS Evaluation – CICDDoS2019" dataset. The LSTM autoencoder is trained in an unsupervised environment to reconstruct encoded data to increase anomaly detection accuracy. The LSTM multi-class classifier is trained in a supervised environment to classify the network traffic data into DDoS attacks. The experiments included hyperparameter tuning, architecture layer modifications, and several preprocessing techniques. Finally, an accuracy of 0.94, an F1-score of 0.94, a precision of 0.95, and a recall of 0.94 was obtained for the proposed model. To validate and compare the performance of the proposed model, the author selected a baseline model for benchmarking where the proposed deep learning model achieved a 5% increment in accuracy, precision, and recall and a 6% increment in F1 Score. According to the evaluation results and most of the comments obtained from the evaluators, it can be concluded that the proposed system is useful, and this approach can be used as a complete product if the approach is further tuned. During the study, the author identified that the project could be enhanced by analyzing live network traffic data, PCAP file extraction instead of using CSV files, removing highly correlated data before model training, novelty detection, and classifying more than five classes.

Keywords: *Hybrid Deep Learning, Recurrent Neural Network, Multiclass Classification, Distributed Denial of Service, Supervised learning*