DEEP LEARNING MODEL FOR DISTRIBUTED DENIAL OF SERVICE (DDOS) DETECTION

Chaminda Tennakoon

A dissertation submitted in partial fulfilment of the requirement for the degree of MSc Big Data Analytics

Department of Computing

Informatics Institue of Technology, Sri Lanka
in collaboration with
Robert Gorden University, Soctland

2019

Abstract

Distributed denial of service (DDoS) attacks is one of the serious threats in the domain of cybersecurity where it affects the availability of online services by disrupting the access to the online services to its legitimate users. The consequences of such attacks could be millions of dollars in worth since all of the online services are relying on high availability. The magnitude of DDoS attacks is ever increasing as attackers are smart enough to innovate their attacking strategies to expose vulnerabilities in the intrusion detection models or mitigation mechanisms. The history of DDoS attacks reflects that network and transport layers of the OSI model were the initial target of the attackers, but the recent records from the cybersecurity domain prove that the momentum has shifted toward the application layer. There is a high degree of difficulty distinguishing the attack traffic and benign traffic when it comes to the application-layer DDoS attacks that make the combat against application-layer DDoS attack a sophisticated task. Stride for high accuracy with high DDoS classification recall is key for any DDoS detection mechanism to keep the reliability and trustworthiness of such a system. In this research, a machine learning approach for application-layer DDoS detection is proposed by using Autoencoder to perform the feature selection and Deep neural networks architecture to perform the attack classification. A popular benchmark dataset in the application layer DDoS experiments CIC DoS 2017 is selected for the research by extracting the most appealing features from the packet flows. The model is capable of detecting the application-layer DDoS attacks at a detection rate of 99.84% with a 0.18 false-positive rate and 0.17% false-negative rate. The model's overall false alarm rate is 0.18%. The model has the strength to detect most of the current application layer DDoS attack flavours. Generative Adversarial Networks (GANs) are built using the existing attack traffic pattern to build new application-layer DDoS attack patterns to test the model's capability and performance for the unseen attack traffic patterns that could happen in the future.