DDoS Attack Detection System

Chapter 1 : Introduction

* 1. Problem Statement

Optimizing intrusion detection dataset using navie Bayes algorithm and knn algorithm

* 1. *Overview of the project*

Cybersecurity is the practice of defending computers and servers, mobile devices, networks, and data from malicious attacks. It is also known as information technology security. Cybersecurity relies on cryptographic protocols used to encrypt emails, files. This security protects information that transmits and guards against loss. In the end, user security software scans computers for malicious code and then removes it from the machine.

This project aims at detecting intrusion in the given KDD Cup dataset. This project is comprised of machine learning and deep learning algorithms. Here we will be distributing the KDD Cup dataset as training and test set. In the first step, One-Hot-Encoding is used to transform all categorical features into binary features. The One-Hot-encoding takes a matrix of integers, denoting the values on by categorical features. The output will be a sparse matrix where each column corresponds to one possible value of one feature. Therefore the features first need to be transformed with Label Encoder, to transform every category to a number. In the second step, we will be applying various machine and deep learning algorithms. These algorithms analyze the large datasets and mechanism which show the intrusion in the given KDD Cup dataset with different accuracies. This system is thus very useful to the users and network administrators in particular who is generally responsible for monitoring things on a network.

* 1. *Objective:*

*1.4 – Literature survey*

1. *“DDoS Attack Detection System: Utilizing Classification Algorithms with Apache Spark*”, 9th IFIP International Conference, 2018 [1]
   * They used Fuzzy logic system using parallel computing for classification.
   * accuracy of DT(Gini) 97.9 % and DT (Entropy) 97.6 %.
2. “*Detection DDoS attacks Based on Neural Network Using Apache Spark”,*International Conference, 2016 [2]
   * They used Back-propagation network and implemented in Apache Spark cluster.
   * Used 2000 DARPA LLDOS 1.0 dataset in a real time environment and accuracy obtained was 94%.
3. ”*DeepDefense: Identifying DDoS Attack via Deep Learning***”,** IEEE International Conference, 2017 [3]
   * Recurrent deep neural network approach to identify DDoS attack.
   * UNB ISCX Intrusion Detection Evaluation 2012 Dataset is used.
   * This approach reduced the error rate from 7.517% to 2.103% compared with conventional machine learning method.
4. ”*Machine-Learning-Based Online Distributed Denial-of-Service Attack Detection Using Spark Streaming***”,** IEEE International Conference, 2018 [4]
   * Naive Bayes, Logistic Regression and Decision Tree algorithms are used in a cluster.
   * The system can detect 3 typical DDoS attacks TCP flooding, UDP flooding and ICMP flooding.
   * The classification was made with the Accuracy of 99.3%.
5. ”*DDoS Detection System: Utilizing Gradient Boosting Algorithm and Apache Spark***”,** IEEE Canadian Conference, 2018 [5]
   * Gradient Boosting classification algorithm (GBT) to identify DDoS attack.
   * Apache Spark Processing Engine for processing data.
   * The classification was made with the accuracy of 97%.

Chapter 2: Proposed System

*2.1 – Description of target users*

The target users in our problem are the Cloud Service provider and the clients requesting cloud services.

Cloud Service Provider-Cloud Computing can be defined as a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction. Pay per usage, Virtualization, on demand access, flexibility and reduced hardware and maintenance cost are some of the factors contributing to popularity of cloud computing. Platform as a Service (PaaS), Infrastructure as a Service (IaaS), and Software as a Service (SaaS) are service models of cloud computing. Virtualization plays a major part in cloud computing by making effective and systematic use of the available hardware. It increases the availability of system and also reduces cost and present a superior flexible system. DDoS attack is major trouble to the availability. The attacker can greatly degrade the quality or fully breakdown the victim’s network connectivity. The main intention of a DDoS attack is to make the victim unable to use the resources. Targets could be web servers, CPU, storage, and the other network resources . In cloud environment also DDoS can reduce the performance of cloud services significantly by damaging the virtual servers.

Clients-When a client outsources a server to a cloud vendor, he/she also relinquishes a large degree of control over that infrastructure, resulting in a risk to the client’s details. It is critical that the client negotiate favorable service level agreements that allow onsite audits of the cloud vendor to ensure that the infrastructure is protected, secure and backed up regularly. What cloud computing will not do in the immediate future is eliminate traditional data centers. Many organizations have rigorous regulatory security requirements, and cloud computing is not robust enough at the point to sufficiently satisfy those at the proper level. Protection from DDoS attacks is one of the provision provided to the clients.

*2.2 Block Diagram*

*2.3 – Scope*

Our Project is capable of detecting intrusion in cloud environment using benchmark dataset.

Chapter 3: Software Requirement System

*3.1– Overview of SRS*

The introduction of the software requirement specification (SRS) provides an overview of the entire SRS with purpose, scope, definitions, acronyms, abbreviations, references and overview of the SRS. The aim of this document is to give an in-depth insight about the detection of the DDoS attack in cloud environment by defining the problem statement in detail.

Non-functional requirements

|  |  |
| --- | --- |
| Type | Description |
| Performance | 1.The system should be able to classify anomalies and normal packets with the accuracy of more than 95%.  2.The pre-processing time of the intrusion detection system should be within seconds. |
| Usability | 1.The system should be available all the time. |

Table 3.2.4.1- Non-functional requirements

*Software and hardware requirements*

Software requirements-

* VMware Workstation
* Ubuntu 16.04 LTS/Windows 8 & above
* Anaconda 3.6.0
* Python 3.6.0

Hardware requirements-

* Minimum 8GB RAM
* 512GB Hard disk
* Minimum core i3 processor

*References-*

1. A. Alsirhani, S. Sampalli and P. Bodorik, "DDoS Attack Detection System: Utilizing Classification Algorithms with Apache Spark," *2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, Paris, 2018
2. C. Hsieh and T. Chan, "Detection DDoS attacks based on neural-network using Apache Spark," *2016 International Conference on Applied System Innovation (ICASI)*, Okinawa, 2016
3. X. Yuan, C. Li and X. Li, "DeepDefense: Identifying DDoS Attack via Deep Learning," *2017 IEEE International Conference on Smart Computing (SMARTCOMP)*, Hong Kong, 2017
4. B. Zhou, J. Li, J. Wu, S. Guo, Y. Gu and Z. Li, "Machine-Learning-Based Online Distributed Denial-of-Service Attack Detection Using Spark Streaming," *2018 IEEE International Conference on Communications (ICC)*, Kansas City, MO, 2018
5. A. Alsirhani, S. Sampalli and P. Bodorik, "DDoS Detection System: Utilizing Gradient Boosting Algorithm and Apache Spark," *2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE)*, Quebec City, QC, 2018