



Article Title: Deep Intrusion Detection for DOS and DDOS Attacks Using LSTM and Deep Autoencoder Neural Network

Deep Intrusion Detection for DOS and DDOS Attacks Using LSTM and Deep Autoencoder Neural Network

Sujini S. P¹, AnbuShamini G. N², Prija .J. S³

¹PG student, Mar Ephraem College of Engineering & Technology, Elavuvillai,
sujinisp123@gmail.com.

²Assistant Professor, Department of ECE, Mar Ephraem College of Engineering &
Technology, Elavuvillai, shamini.anbu14@gmail.com

³Assistant Professor, Department of ECE, Mar Ephraem College of Engineering &
Technology, Elavuvillai, Jspriya86@gmail.com

ABSTRACT

Early detection of network intrusions is a very important factor in network security. However, most studies of network intrusion detection systems utilize features for full sessions, making it difficult to detect intrusions before a session ends. To solve this problem, the proposed method uses packet data for features to determine if packets are malicious traffic. Such an approach inevitably increases the probability of falsely detecting normal packets as an intrusion or an intrusion as normal traffic for the initial session. As a solution, the proposed method learns the patterns of packets that are unhelpful in order to classify network intrusions and benign sessions. To this end, a new training dataset for Generative Adversarial Network (GAN) is created using misclassified data from an original training dataset by the LSTM-DNN model trained using the original one. The GAN trained with this dataset has ability to determine whether the currently received packet can be accurately classified in the LSTM-DNN. If the GAN determines that the packet cannot be classified correctly, the detection process is canceled and will be tried again when the next packet is received. Meticulously designed classification algorithm based on LSTM-DNN and validation model using GAN enable the proposed algorithm to accurately perform network intrusion detection in real time without session termination or delay time for collecting a certain number of packets. Additionally, a Deep Autoencoder neural network is utilized to automatically extract relevant features from the network traffic. This unsupervised learning approach enables the system to adapt to evolving attack patterns

1 Introduction

Home automation systems provide several ingress points – like smart meters, wireless lamps/bulbs, surveillance equipment, and smart thermostats, to name a few. Such connected devices with attached sensors provide hackers with a tremendous opportunity to exploit the system. While IoT has made the management of various daily tasks simple, it is essential to guarantee that criminals do not enter our homes using a loophole. As systems evolve and become security smart, hackers become smarter. IoT devices gather a huge amount of data



Article Title: Deep Intrusion Detection for DOS and DDOS Attacks Using LSTM and Deep Autoencoder Neural Network

during their lifespan. With the rapid growth of 5G implementation, data communication between devices and networks is projected to increase many folds. It is pertinent to note that if the data captured/generated by these devices is not secured, it remains available for stealing for financial gain or worst, it may put the lives of people at risk across the globe.

The major contributions of the work at hand are as follows:

- Comparison of the UNSW-NB15 and the Bot-IoT datasets to identify common features between them.
- Union of both data-sets using common features which fall in flow and TCP categories.
- Removal of issues pertaining to the imbalance nature of data-sets and biased classification.
- Employing LSTM based un-supervised deep learning model for the detection of DoS and DDoS attacks.

The model is trained on all of the data available in the UNSW-NB15 and BotIoT data-sets to cover the maximum possible types of packets.

2 Related Literature

In the cyber security domain, towards cyber antagonists, an Intrusion Detection System (IDS) serves as a clear defense line, which is very critical for a system. Due to mobile device penetration and the popularity of apps that quickly accomplish various user tasks, we have become dependent on this parallel universe defined by electronic devices. Users need to be made aware of the implication of negligence towards secure practices to protect network infrastructure against intrusion threats. A device is considered protected if it successfully achieves data protection, confidentiality, Integrity and availability (CIA).

To track network dependent system an IDS can be employed in host mode or network mode. A host dependent IDS (HIDS) tracks host affairs by gathering data of computer system activities. In this kind of device, a monitor must be mounted to track hosts and record their actions on the operating system for an investigation trail. It is therefore important that HIDS should be compliant with its devices to track various operating systems. Since investigation trails are used, the process is resource hungry and results in increased load on the host server which is a costly proposition.

3 Proposed Methodology

The proposed Protocol Based Deep Intrusion Detection (PB-DID) architecture as shown in Figure 1. The process involves features comparison to find similar features in the UNSW-NB15 and the Bot-IoT datasets, features selection, data pre-processing and selection and model training using un-supervised LSTM deep learning model.



Article Title: Deep Intrusion Detection for DOS and DDOS Attacks Using LSTM and Deep Autoencoder Neural Network

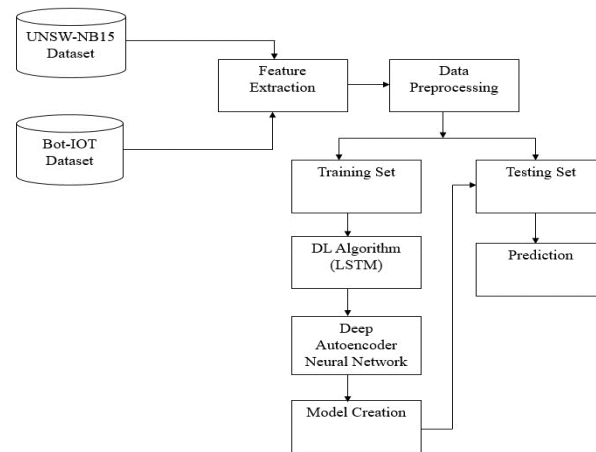


Figure1: Block Diagram

3.1 Data-Sets

We used two well-known raw network packets data-sets namely the UNSW-NB15 and the Bot-IoT for training and validation. Unlike many other studies which used a smaller part of these data-sets, we have used complete data for model training. Brief description of both data-sets is given below.

- i. **UNSW-NB15:** The data-set was published by Moustafa et al., in 2015. This data-set is simulated with over 2.5 million network packets. This data-set consists of nine types of attacks (Exploits, Reconnaissance, DoS, Generic, Shellcode, Fuzzers, Backdoors, Worms and Analysis)) with non-anomalous packets as well. More than 87% packets are of non-anomalous type which makes the data-set highly imbalanced.
- ii. **BOT-IoT DATA-SET:** This data-set is the latest in the field. Koroniotis et al., [9] published the data-set in 2018. It consists of more than 72 million records with the mixture of simulated and real time scenarios. It has four categories of attacks but major portion of data-set has DoS and DDoS type of packets. This data-set is imbalanced just like the UNSW-NB15 data-set.

3.2 Features Comparison

In the UNSW-NB15, there are 49 features including 48th as a multi-class label and 49th as a binary label. In the BotIoT data-set, there are 46 features and the last three are label features. In proposed PB-DID, the features in both data-set are compared and we found that 29 features in the Bot-IoT are similar or can be evaluated in the UNSW-NB15 data-set as well.



Article Title: Deep Intrusion Detection for DOS and DDOS Attacks Using LSTM and Deep Autoencoder Neural Network

4 Experiments And Results

The results of experiments performed using the proposed PB-DID architecture. All experiments are performed on google colab using tensorflow as back-end and the language used is Python3.

4.1 Classification Results

Binary classification was performed by taking two classes at a time from a total of three classes which gives us three different configurations of experiments. Figure 2 shows the heat map of fourteen confusion matrices for each chunk of DDoS

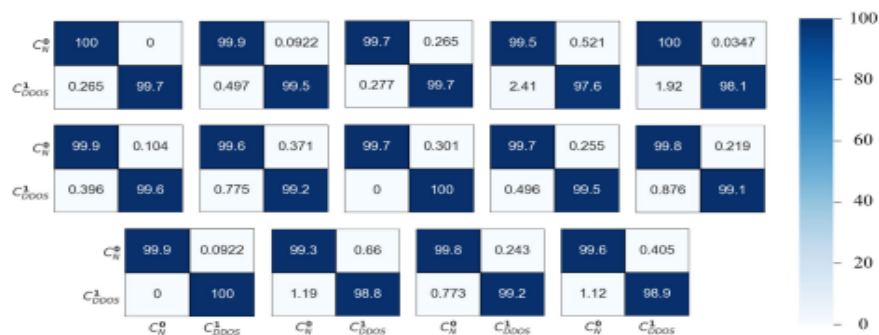


Figure 2: Confusion matrices of all 14 data chunks, C_{ij} , $i = 0$ or 1 is the label and $j =$ non-anomalous (N) when $i = 0$ and DDoS (DD) when $i = 1$ is the class

4.2 Parameters Tuning

We performed parameter tuning by taking a bracket of values and performing the experiments. The optimal values are chosen, where we achieved the best accuracy.

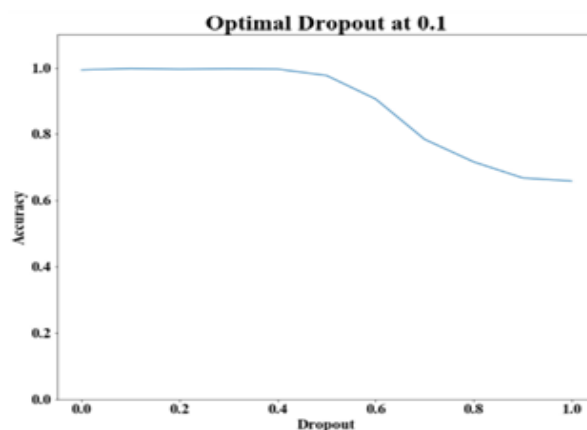


Figure 3: Optimal recurrent dropout graph over $[0, 1]$ and best accuracy 99.885% at 0.3 is achieved



Article Title: Deep Intrusion Detection for DOS and DDOS Attacks Using LSTM and Deep Autoencoder Neural Network

Five parameters are involved in parameter tuning: dropout, recurrent dropout, activation function, recurrent activation function, and number of epochs. Dropout and recurrent dropout are tested over a range of 0 to 1 figure 3 activation and recurrent activation are tested over RELU sigmoid, and tanh functions

References

1. O. Amir, Year: 2019, "Cyber Threats to IoT in 2020".
2. LIFARS, Year: 2020, "Impact of 5G Network on IoT Security".
3. D. Ciso, Year: 2019, "50% of Top 12 Global Exploits Targeted IoT Devices: Fortinet Threat Landscape Report".
4. Z. Doffman, Year: 2019, "Cyberattacks on IoT Devices Surge 300% in 2019", Measured in Billions.
5. Year: 2016, SR Department, "Internet of Things (IoT) Connected Devices Installed Base Worldwide From 2015 to 2025".
6. T. Alsop, Year: 2020, "Internet of Things Security Spending Worldwide From 2016 to 2021". [Online].
7. Nour Moustafa; Jill Slay, Year: 2015, "UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)", in Proc. IEEE Military Commun. Inf. Syst. Conf. (MilCIS), pp. 1–6.