Detection of DDOS Attacks in Network Traffic Using Deep Learning

A. SUNGUR UNAL[1], M. HACIBEYOGLU[1]

[1]Necmettin Erbakan University, Konya/Turkey, aysegulsngr@gmail.com
[1]Necmettin Erbakan University, Konya/Turkey, hacibeyoglu@konya.edu.tr

***Abstract* - In the literature, machine learning algorithms are frequently used in detecting anomalies in network traffic and in building intrusion detection systems. Deep learning is a subfield of machine learning that trains a computer-based system to perform humanitarian tasks, such as disease diagnosis, speech recognition, image recognition, fraud detection, and making predictions. In the experimental study, NSL-KDD dataset was used for evaluating the performance of the proposed deep learning based DDoS detection model. NLS-KDD dataset contains normal network traffic and 23 different DDoS attacks that consists of 41 features. In the experimental study two different experiments are carried out. Firstly, the proposed deep neural network detected the Dos attacks with 0.988 classification accuracy. In the second experiment, the number of features of NSL-KDD is reduced to 24 by examining the previous feature reduction research on NSL-KDD dataset. The proposed deep neural network classified the all cyber-attacks with 0.984 classification accuracy. The 10-fold cross validation is used for all experiments. As a result, the proposed deep learning based DDoS detection achieved good performance.**

***Keywords* - Deep learning, intrusion detection systems, distributed denial of service, classification.**

I. INTRODUCTION

The number of cyber-attacks is increased with the developing technology. Personal and institutional computer systems can face a variety of threats and danger by malicious people, such as information theft, spoofing and denial of service. This can create material and moral damages for both individuals and institutions. For this reason, it is necessary to take security measures.

Security mechanisms have been established to prevent security weaknesses against a variety of threats that may violate information security. Intrusion detection systems are a security mechanism which is created for this purpose.

Catak and Mustacoglu, proposed a model for the detection of malicious network attacks with machine learning and deep learning technologies by using a data set from Cyber Security Laboratory in Australia Cyber Security Center [1].

Kaya, examined performance of artificial neural networks (ANN), K-nearest neighbor algorithm (KNN), support vector machines (SVM) and decision tree by using the datasets KDD CUP99 and NSL-KDD. In determining DOS attacks, KNN, decision trees and ANN have achieved good performances [2].

Yuan et al. have proposed DeepDefense technique to detect DDOS attacks by using the ISCX2012 dataset and found that the error rate is lower than the traditional machine learning methods [3].

Vijayasarathy tried to detect DOS and DDOS attacks using the Naive Bayes method for TCP and UDP with a system designed for real-time and feasibility in his work [4].

Uslu has created two different decision trees using the standard ID3 and the newly proposed method using the KDD CUP99 data set for DOS detection in his work. He compared the success rates of the attacks detected by these two methods and found that the decision tree structure constructed with the new methods is 3% more successful than the ID3 algorithm [5].

Guven uses the SVM, ANN, decision trees, bayesian networks and KNN machine learning algorithms using KDD CUP99 and NSL-KDD datasets to determine the sensitivity, selectivity, precision and F-score of intrusion detection systems [6].

Saied et al. attempted to detect known and unknown DDOS attacks in the real-time environment. In order to detect DDOS attacks, artificial neural networks connected to certain characteristic features that distinguish real traffic from attack traffic are used. [7]. According to this study:

- The detection of DDOS attacks by deep learning methods that are becoming popular nowadays,
- Deep learning methods enable faster and higher success with large data sizes.

II. INTRUSION DETECTION SYSTEM

An intrusion detection system is a warning system that detects an attack if the system's confidentiality, integrity or accessibility is damaged, by monitoring all events that may occur in a computer system or computer network.

Intrusion detection systems are similar to security alarms which are used in our daily lives. Attacks against computer networks can be understood by detecting abnormalities in network traffic. The network traffic is monitored and compared against the database where the attack signatures are already located, and the attack is alerted.

There are 4 basic attack types on the network.

1. Denial of Service Attack (DoS): It is to make the system unserviceable by sending more requests than the system can respond to. The most well-known DOS example attacks can be given as SYN floods, Smurf, UDPstorm, Pingflood, Neptune, Mailbomb attacks [8].

2. User Root Attack (U2R): An attacker with access to a regular user account on the system uses some security vulnerabilities (by sniffing passwords, dictionary attacks or social engineering) to provide root access to the system [9].

3. Remote Local Attack (R2L): An attacker who has the ability to send packets to a machine over a network, but does

not have an account on this machine, gains some security advantage and gains local access as the user of that machine [8].

4. Probing Attack: An attempt to gather information about a computer network to prevent security checks [8].

III. DDOS

A common method for performing Distributed Denial of Service (DDOS) attacks is when an attacker sends a packet flow to a victim; this flow consumes some important resources and makes it unusable for legitimate customers of the victim. Another common approach is that an attacker sends several malformed packets that confuse an application or protocol in the victim machine and force it to freeze or restart [10].

IV. DATA SET

The NSL-KDD data set was used to train and test our system. KDD CUP 99 is the most commonly used data set for anomaly detection. However, the various disadvantages and various statistical analyzes in the KDD CUP 99 dataset affected the accuracy of many IDS determinations modeled by researchers. NSL-KDD, a new data set consisting of the selected records of the full KDD CUP 99 data set was created.

The advantages of NSL-KDD over the original KDD dataset are:

• There are no unnecessary records in the learning set. For this reason, classifiers will not be directed to records more often.

• The number of records selected from each difficulty level group is inversely proportional to the percentage of records in the original KDD data set. As a result, the classification ratios of different machine learning methods vary within a wider range, making it more efficient to correctly evaluate different learning techniques.

• The number of records in the train and test sets is reasonable and this makes it convenient to run the entire set without having to randomly select a small section. As a result, the evaluation results of different research studies will be consistent and comparable [11].

In each record there are 41 properties that reveal different properties of the stream, and each has an attack type or a label that is normally assigned. These types of attacks are grouped as DoS, Probe, R2L and U2R. Attributes are listed on the tables [12].

1. Basic Attribute: This category encompasses all attributes extracted from a TCP / IP connection. Many of these features cause a latent delay in detection.

Table 1. Each Network Connectivity Vectors Basic Features

No.	Feature	Type
1	Duration	Numeric
2	Protocol_type	Nominal
3	Service	Nominal
4	Flag	Nominal
5	Src_bytes	Numeric
6	Dst_bytes	Numeric
7	Land	Binary
8	Wrong_fragment	Numeric
9	Urgent	Numeric

2. Context Properties: Contrary to the majority of DoS and Probing attacks, R2L and U2R attacks do not have any suitable attacks against the frequently encountered sequential die. For this reason, DoS and Probing attacks contain many links with some hosts or hosts in a very short time; but R2L and U2R attacks are embedded in the data partitions of the packages and normally only contain a single link. To detect such attacks, some features are needed to look for suspicious behavior in the data partition (for example, unsuccessful logon attempts). These properties are called content properties [9].

Table 2. Each Network Connectivity Vectors Related Features

No.	Feature	Type
10	Hot	Numeric
11	Num_failed_logins	Numeric
12	Logged_in	Binary
13	Num_compromised	Numeric
14	Root_shell	Binary
15	Su_attempted	Binary
16	Num_root	Numeric
17	Num_file_creations	Numeric
18	Num_shells	Numeric
19	Num_access_files	Numeric
20	Num_putbound_cmds	Numeric
21	Is_hot_login	Binary
22	Is_guest_login	Binary

3. Traffic Properties: This category contains properties calculated according to a window range and is divided into two groups:

a) "Same Host" Features: The same target host with the current connection has features that only review the last 2 seconds of connectivity, and such as protocol behavior or service calculates statistics about.

b) "Same Service" Features: It examines only the last 2 seconds of connections with the same service as the current connection.

The two "traffic" modes mentioned above are called time based. However, there are several slow scan attacks that scan the host machine (or ports) for a time interval greater than 2 seconds, for example, every minute. As a result, these attacks do not create attack patterns with a 2-second time window. To solve this problem, the "same host" and "same service" features are recalculated, but rely on the link window of 100 links instead of a 2 second time interval. These features are called link-based traffic features [11].

Table 3. Traffic Characteristics for Each Networking Vectors Time

No.	Feature	Type
23	Count	Numeric
24	Srv_count	Numeric
25	Serror_rate	Numeric
26	Srv_serror_rate	Numeric
27	Rerror_rate	Numeric
28	Srv_rerror_rate	Numeric
29	Same_srv_rate	Numeric
30	Diff_srv_rate	Numeric
31	Srv_diff_host_ rate	Numeric

Table 4. Host Based Traffic Features in Network Connection Vectors

No.	Feature	Type
32	Dst_host_count	Numeric
33	Dst_host_srv_count	Numeric
34	Dst_host_same_srv_rate	Numeric
35	Dst_host_diff_srv_rate	Numeric
36	Dst_host_same_src_port_rate	Numeric
37	Dst_host_srv_diff_host_rate	Numeric
38	Dst_host_serror_rate	Numeric
39	Dst_host_srv_serror_rate	Numeric
40	Dst_host_rerror_rate	Numeric
41	Dst_host_srv_rerror_rate	Numeric

V. DEEP LEARNING

Deep learning is one of the machine learning techniques and machine learning is done in one layer at the same time in many layers. A group of machine learning algorithms are used at the same time to produce results in a single operation. Deep learning is the use of multi-level "deep" neural networks with advanced technology to create feature detection systems with large numbers of untagged training data. Figure 1 shows the input layer, multiple hidden layers and output layers.

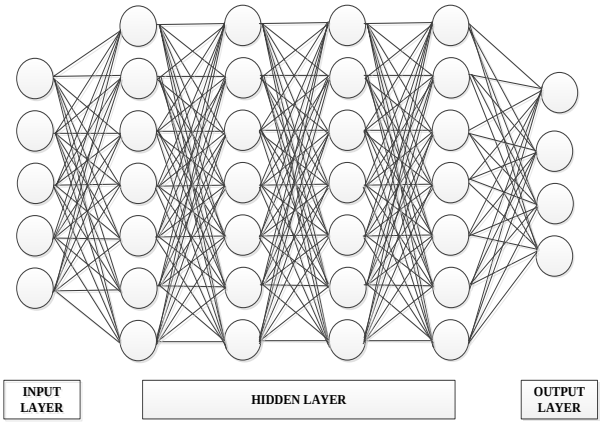

Figure 1. Deep Learning Neural Network

Deep learning methods learn the characteristics that distinguish them from a large number of data inputs. This feature should be adequately trained so that the learning can be done successfully. Feature learning consists of layers. The sublayers have less distinctive features while the uppermost layers have more distinctive features. The lower level features are the basis for building the upper level features. This type of learning is different from learning a machine. This is because the machine is trained with the characteristics determined by the people at the learning stage. In other words, while machine learning algorithms are human dependent, deep learning is human independent. This approach is an important influence on the success of deep learning.

Deep learning has many effects on our life. These;

1. Increasing the amount of data: Over the years the widespread use of the internet has caused it to be produced and stored digitally in very large sizes. The use of these great data has been achieved through deep learning systems [13].

2. Increase GPUs and processing power: Powerful and efficient parallel calculations can be done by GPU (Graphics Processing Unit) calculation. GPUs are being used to train deep learning algorithms using much larger training sets in a much shorter time and with much less data center infrastructure [13].

3. Increased depth: Increased processing power enables deep models to be used in practice. Deep learning models are also models with multi-layered structure. We can connect with the visual system of the human brain in order to understand the depth. The signals coming to the brain through the nerves in the eyes are processed by hierarchical evaluation in a layered structure. In the first layer the signal follows, the more local and basic features of the image, such as the edges and the corners, are recognized. In the next layer these edges and corners are brought together to form the mouth and nose, then faces on the next layer. On the next layer, the characteristics of the whole of the view like the settlement of the person and objects can be recognized. Many deep learning systems work on this principle [13].

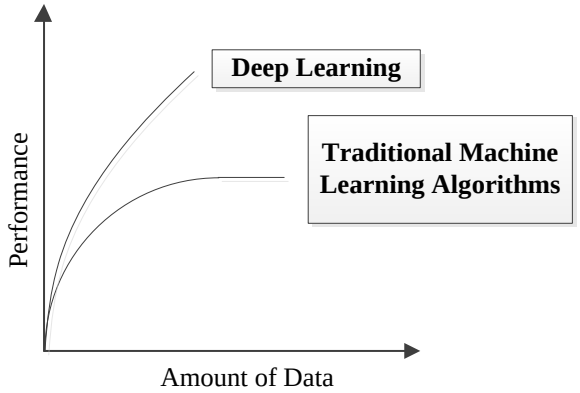

Figure 2. Performance Graph of Deep Learning and Machine Learning Algorithms

Increasing number of layers brings a deeper decision mechanism. Deep learning can work with more layers than traditional artificial neural networks. In the classical ANN

structure, each cell is connected to all the cells in the previous and subsequent layers. There are mathematical operations that need to be computed for each connection. As the number of layers and the number of cells increase, these processes require a high amount of CPU (Central Processing Unit) power. The CPU power in personal computers is insufficient to create a deep network structure. With the development of the GPU, the development of deep learning has also been accelerated. The difference between the GPU and the CPU is how a process is processed. CPUs contain several cores. Each core has a high processing capacity and processes are processed in a serial manner distributed to these cores. GPUs have hundreds of cores. The processing capacity of each kernel is less than CPU. However, it has high parallel processing power and can perform many simultaneous operations. This processing capacity of the GPU is an important point for implementing deep learning. In this way, the performance of training with a very large amount of data has increased. This is a point that distinguishes deep learning from traditional machine learning algorithms. Because in traditional machine learning algorithms, a high amount of data causes the success to increase some and then the success to remain constant. This success is shown in the graphic in Figure 2 [14].

VI. EXPERIMENTAL RESULTS

In order to evaluate the performance of the deep learning techniques for DDoS detection problem the proposed models are implemented using Python Keras library [15]. All experiments are carried out using a computer with an Intel Core i7 3840QM@2.80 Ghz processor with 16 GBs of memory operated on Microsoft Windows 8 system. In all experiments, the well-known evaluation measures classification accuracy (CA), recall and F-measure are obtained using a 10-fold cross validation for the proposed deep learning model. These evaluation measures are calculated in terms of true positive (TP), true negative (TN), false negative (FN) and false positive (FP) that are given in Eq. (1).

$$CA = (TP + TN)/(TP + TN + FP + FN)$$
$$Precison = (TP)/(TP + FP) \quad (1)$$
$$Recall = (TP)/(TP + FN)$$

The NSL-KDD dataset contains four main attack types and 22 different attacks classes which are shown in Table 5.

Table 5. Attack types of NSL-KDD

4 Main Attack Types	22 different attacks
Denial of Service (DoS)	back, land, neptune, pod, smurf, teardrop
Remote to User (R2L)	ftp_write, guess_passwd, imap, multihop, phf,spy, warezclient, warezmaster
User to Root (U2R)	buffer_overflow, perl, loadmodule, rootkit
Probing	ipsweep, nmap, portsweep, satan

Two different experiments were carried out in the experimental study. In the first experiments, DoS attacks are labeled as "True" and the other attacks are labeled as "False" and the deep learning model shown in Figure 3 is used.

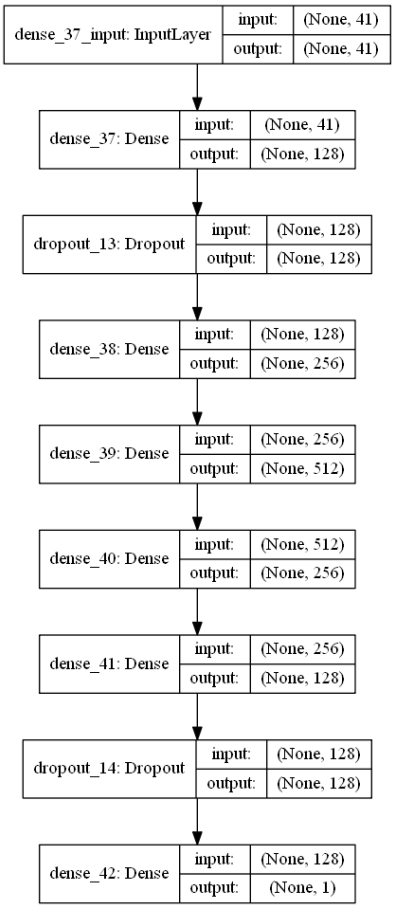

Figure 3. Proposed model for first experiment

For the first experiment, the number of epoch is set to 20 in the training process, binary_crossentropy is used as loss function and adam is used as optimizer. The evaluation measure results of the first experiment are shown in Table 6.

Table 6. Evaluation measures for DoS classification

CA	Precision	Recall
0.988	0.982	0.984

In the second experiments, deep learning is used to classify all cyber-attacks in NSL-KDD dataset. Prior the classification, most relevant features of DoS attacks are selected as a data pre-processing according to the previous literature [16]. The selected features are shown in Table 7.

Table 7. Attack types of NSL-KDD

Class Label	Relevant Features
Land	7
Smurf	2,3,5,23,24,27,28,36,40,41
Neptune	4,25,26,29,30,33,34,35,38,39

Teardrop	8
Back	10,13

For the second experiment, the number of epoch is set to 20 in the training process, sparse_categorical_crossentropy is used as loss function and adam is used as optimizer. The number of input is increased to 25 and the number of output is increased to 23. The evaluation measure results of the second experiment are shown in Table 8.

Table 8. Evaluation measures for DoS classification

CA	**Precision**	**Recall**
0.948	0.945	0.948

VII. CONCLUSION

In this paper, deep learning approach is proposed to detect DDoS cyber-attacks and NSL-KDD dataset is used for the experiments. A deep learning model is proposed for two different experiments. In the first experiment, our model achieved 0.988 CA for DoS attacks. In the second experiment, proposed model achieved 0.948 CA for all types of cyber-attacks. For future work, deep learning approach will be used with other machine learning approaches in a hybrid way and the parameters of deep learning will be tuned for improving the deep neural.

REFERENCES

[1] Çatak, F. Ö., Mustaçoğlu, A. F., (2017). Derin Öğrenme Teknolojileri Kullanarak Dağıtık Hizmet Dışı Bırakma Saldırılarının Tespit Edilmesi,2017. The 5th High Performance Computing Conference.

[2] Kaya, Ç. (2016). Saldırı Tespitinde Makine Öğrenmesi Tekniklerinin Performans Analizi.

[3] Yuan, X., Li, C., & Li, X. (2017, May). DeepDefense: Identifying DDoS Attack via Deep Learning. In Smart Computing (SMARTCOMP), 2017 IEEE International Conference on (pp. 1-8). IEEE.

[4] Vijayasarathy, R. (2012). A system approach to network modeling for DDoS detection using a Naïve Bayesian classifier.

[5] Uslu, N. Celal. (2009). Veri Madenciliği ile Bilgisayar Ağlarında Yeni Bir Saldırı Tespit Algoritması

[6] Güven, E. N.(2007). Zeki Saldırı Tespit Sistemi Tasarımı ve Gerçekleştirilmesi.

[7] A. Saied, R. E. Overill ve T. Radzik, "Detection of known and unknown DDoS attacks using Artificia Neural Networks", Neurocomputing, 172, pp.385-393, 2016.

[8] Mukkamala, S., Janoski, G., & Sung, A. (2002). Intrusion detection using neural networks and support vector machines. In Neural Networks, 2002. IJCNN'02. Proceedings of the 2002 International Joint Conference on (Vol. 2, pp. 1702-1707). IEEE.

[9] Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009, July). A detailed analysis of the KDD CUP 99 data set. In Computational Intelligence for Security and Defense Applications, 2009. CISDA 2009. IEEE Symposium on (pp. 1-6). IEEE.

[10] Mirkovic, J., & Reiher, P. (2004). A taxonomy of DDoS attack and DDoS defense mechanisms. ACM SIGCOMM Computer Communication Review, 34(2), 39-53.

[11] Datti, R., & Verma, B. (2010). B.: Feature Reduction for Intrusion Detection Using Linear Discriminant Analysis. In International Journal on Engineering Science and Technology.

[12] Dhanabal, L., & Shantharajah, S. P. (2015). A study on NSL-KDD dataset for intrusion detection system based on classification algorithms. International Journal of Advanced Research in Computer and Communication Engineering, 4(6), 446-452.

[13] Genç Ö., 2016, Keras ile Derin Öğrenmeye Giriş [online], https://medium.com/turkce/keras-ile-derin-%C3%B6%C4%9Frenmeye-giri%C5%9F-40e13c249ea8, [Available Date: 11 Mart 2018]

[14] Buber, E., 2017, Derin Öğrenme Nedir? [online], Cyber Security with Machine Learning, https://cybrml.com/2017/06/06/derin-ogrenme-uygulamalari/ [Available Date: 06 Şubat 2018]

[15] Keras, https://keras.io/

[16] Noureldien, N. A., & Yousif, I. M. (2016). Accuracy of machine learning algorithms in detecting DoS attacks types. Science and Technology, 6(4), 89-92