Naif Arab University for Security Sciences
Journal of Information Security and Cybercrimes Research
مجلة بحوث أمن المعلومات والجرائم السيبرانية
https://journals.nauss.edu.sa/index.php/JISCR
JISCR

A Security System for Detecting Denial of Service (DDoS) and Masquerade Attacks on Social Networks

Olaniyi Abiodun Ayeni*, O. Owolafe, and O. P. Ogunjobi
Department of Cyber Security, School of Computing, Federal University of Technology, Akure, Ondo State, Nigeria.

Received 1 Dec. 2021; Accepted 08 Apr. 2022; Available Online 01 June. 2022

Abstract

This study on a security system for detecting denial of service (DDoS) and masquerade attacks on social networks specifically describes how a Convolutional Neural Network (CNN) algorithm was employed. The dataset used for this research is the CICIDS2017 dataset, which contains benign data (no attack present) and the most up-to-date, frequent attacks which resemble true, real-world data. The feature extraction method used was recursive feature elimination (RFE), which reduced 77 columns of the dataset to 10 columns. This research was motivated by the limitation of Alguliyev and Abdullayeva 2019, which focused on the prediction of DDoS attack occurrence by getting related texts in social media. It has a limited attack class that focuses solely on DDoS attacks, and it does not perform social media network prediction in general. The objective of this research is to develop a security system for detecting DDoS and masquerade attacks and evaluate the detection model on social media networks. The system was tested on Facebook and Instagram. The result of the training accuracy that we derived from this research is 99.53%, while the testing accuracy is 99.52%. The result of this research is compared with previous studies' results. This study recommends that the model implemented can be enhanced more effectively by comparing the accuracy of alternative deep learning algorithms to that of the CNN utilized in the current prediction model.

I. Introduction

Distributed denial of service attacks and masquerade attacks are types of active attack that will be studied among the various types of active attack that exist. Distributed Denial of Service (DDoS) [14],[18] is a multi-pronged version of the normal denial of service where the attack comes from multiple sources referred to as botnets. These automated systems try to flood the network in question with an unrelenting supply of packets, which in turn prevent the proper usage and functioning of the network [1]. The masquerade attack refers to the usage of a fake personality to gain illegal access to any personal computer [2]. A masquerading attack is one in which one system takes on the identity of another [8]. It is a strategy in which an attacker poses as an authorized individual to get illicit access to confidential information. For millions of internet users, [15] the online network has become a mainstream cultural phenomenon, increasing the integration of our online and offline lives by utiliz-

Keywords: Cybersecurity, CNN, DDoS, Masquerade Attack, Social Media.

Production and hosting by NAUSS

* Corresponding Author: Olaniyi Abiodun Ayeni
Email: oaayeni@futa.edu.ng

doi:10.26735/GMEY8791

ing consumers' real-world social ties. The internet period has advertised unused implies to deliver and share information through substantial social networking sites on the internet [19]. Online social networks allow the delivery and consumption of information by many internet users around the world. They provide a tremendous source of information on a phenomenal scale. Social networking sites like Facebook, LinkedIn, and Twitter are among the most popular web-based online services today [20]. Clearly, social networking sites are crucial applications in terms of user security [21] and privacy [9]. Traditional security threats are frequently used in social media attacks such as malware, worms, spam, and phishing; these attacks, on the other hand, are carried out in a different context, using social media as a new channel to contact victims [3]. The availability of an incredible amount of personal user data that would not have been accessible otherwise has elevated the stakes for privacy protection [4],[10].

This is a concern, in terms of information and communication security on online social network platforms such as Facebook, Twitter and Instagram [11],[22]. The goal of this research is to implement a security framework for detecting DDoS and masquerade attacks and evaluate the detection model on social media networks.

Consequently, the rest of the paper is organized as follows, Section II sheds the light on the previous studies. Section III explains the research methodology. Section IV details the system implementation and the discussion of the results. Section V concludes the paper.

II. Literature Review

Research on the implementation of a security system for detection of DDoS and masquerade attacks has become the interest of many in the last few years [16]. A generous number of methodologies, motivations, objectives, methodologies, contributions to knowledge and restrictions with the purpose of reducing the detrimental consequences of DDoS and masquerade attacks on social network have been analyzed, proposed, and evaluated [13].

Research [5] shows how the deep learning method was used for prediction of DDoS attacks on social media. Detection including an improved CNN [17] and LTSM model is employed by predicting the likelihood of cyber-attack-related phrases by self-learning methods which are then translated in text-type social media discussions. The shortcoming of this research is that it focuses on the prediction of DDoS cyberattacks by locating relevant texts in social media and classifying them with high precision into positive and negative categories. However, it has a limited attack class that focuses solely on DDoS attacks and does not perform social media network prediction in general [7].

A case study on effective masquerade attack detection was presented [6]. A user study was designed for the detection to explore null hypothesis. The experimental hypothesis asserts that if the masquerader's goal is malicious, the null hypothesis states that manipulating the masquerader's intent has no meaningful impact on the masquerader's search behavior. It shows an integrated strategy for detecting masquerade attacks that combines user behavior profiling with a baiting approach that uses carefully produced and strategically placed fake documents to lure attackers in. The shortcoming is that vulnerability exists in the user search behavior profiling sensor.

Research [2] proposes DDoS attack detection and classification via a Convolutional Neural network. The objective of this research is to develop five discrete classification algorithms and implement them to detect and classify DDoS attacks and to build and train by using two different datasets. According to analysis and results, the presented results in this research established that CNN performed better than other classifiers with an accuracy of 99%. The limitation of research is shown in the inadequacy of building a novel model to stop or minimize DDoS attacks based on the output of the CNN classification algorithm developed. A lot of research deals with the issue of DDoS and masquerade attack detection using CNN.

III. Research Methodology

For the purpose of this research, the major approach in building a software framework which can

detect these attacks is implementing a detection model that employs a machine learning approach, a convolutional neural network (CNN) [17]. Each step of the approach is detailed in below:

1. Data collection: The dataset used for this research is the CICIDS2017 dataset, which covers benign data (no attack present) and up-to-date frequent attacks, and which closely reflects real-world data. The implemented attacks in this dataset include Brute Force FTP, Brute Force SSH, DoS, Heartbleed, Web Attack, Infiltration, Botnet and DDoS [12]. However, the dataset attacks were constrained in this work to just 3 attacks: DDoS, infiltration, and botnet. The dataset used was downloaded from (https://www.kaggle.com/cicdataset/cicids2017). The dataset consisted of various features/columns.

2. Data preprocessing: Raw data and images from the real world are frequently imprecise, unreliable, and lacking in specific behaviors or trends. They are also likely to be riddled with errors. As a result, once they have been collected, they are pre-processed into a format that the CNN can utilize to build the model. It is used to check if the dataset has been used before, if it is flawed or it will need reappropriation of sources. The features that constitute a proper dataset will be discussed further. The steps involved are encoding the categorical data, normalization, and checking for missing data.

3. Data training: Splitting the dataset into train set and test set. The training dataset is fed into the machine learning algorithm to train our model; it is the one that is utilized to teach an algorithm how to learn and create results using concepts like neural networks. It contains both the desired outcome and the input data. While the testing dataset is utilized to confirm our model's accuracy, it is not used to train it. It could be referred to as the validation dataset. The test data set is used to see how effectively an algorithm learned from the training data set. The training data set cannot be used in the testing step in projects like this because the algorithm would already know the predicted outcome, which is not our purpose.

4. Feature extraction: Feature is a single column of data. It is an element of an observation and is also called an attribute of a data instance. Recursive feature elimination (RFE) was used to know the features relevant to this research.

5. Model implementation.

6. Model evaluation: Dataset evaluation is an integral part of the model development process. It assists in determining the appropriate model to describe our data and determining how well the chosen model will perform in predicting data from the test set.

7. Save the model using Pickle.

8. Simulate the model using social network data captured from Wireshark.

IV. System Implementation

A. Data Collection

The libraries used for implementation were numpy, pandas, matploylib, seaborn, sklearn, time, and warnings. The CICDIS dataset which was used started capturing its data at 9 a.m., Monday, July 3, 2017, and ended at 5 p.m. on Friday, July 7, 2017, a total of 5 days. However, only the data set of Friday afternoon with titles DDoS and PortScan was imported with titles infiltration.

B. Data Concatenation

The data was merged, representing infiltration and portscan as a masquerade attack. We have the following attacks after concatenation, and the new number of rows per attack is as follows:

i. Benign (meaning no attack present) – 513.821.
ii. Masquerade_attack - 158.966.
iii. DDoS - 128.027.

C. Feature Selection

The method used for feature selection was reinforcement feature elimination (RFE). The reason why RFE was used here was because of its simplicity and efficiency of use; it is one of the most often used feature selection algorithms. Its job is to pick the most related features to the target column and eliminate every other one.

The number of rows and columns that were present before feature selection is 800,814 rows and 77 columns, meaning this is the number of fea-

tures that existed before feature extraction. These rows and columns were further reduced (from 76, excluding the label column) to 10 selected features.

D. Training

The dataset was trained using CNN deep learning algorithm where the dataset is split into 20% testing dataset and 80% training dataset. Training data is the information used to train an algorithm or machine learning model to foresee the outcome you want it to predict. Test data is used to evaluate the algorithm you are using to train the machine's performance, such as accuracy or efficiency.

There was an accuracy of 0.9953 on the training data and 0.9952 on the testing data, showing that there was utilization of a good dataset and that the detection model is suitable for use. Fig. 1 shows a pictorial representation of the testing and training accuracy.

E. Model Evaluation with Confusion Matrix

A confusion matrix is a N x N matrix that is used to assess the effectiveness of a classification model, with N denoting the number of target classes. The matrix compares the actual goal values to the predictions of the machine learning model.

True Negative: The model predicted No, and the real or actual value also indicated No.

True Positive: The model has predicted Yes, and the actual value was likewise correct.

False Negative: The model predicted No, but the actual value was Yes. It is also called a Type-II error.

False Positive: Yes was predicted by the model, but the actual result was No. It is also called a Type-I error.

Precision: Precision is the ratio of accurately predicted positive observations (True Positives) to all predicted positive observations, both correct (True Positives) and incorrect (False Positives).

Recall: The ratio of system-generated results that properly predicted positive observations (True Positives) to all observations in the real positives is known as recall.

F1-score: Precision and Recall are weighted in the F1 Score. As a result, to create a compromise

Fig. 1. Accuracy graph over epochs.

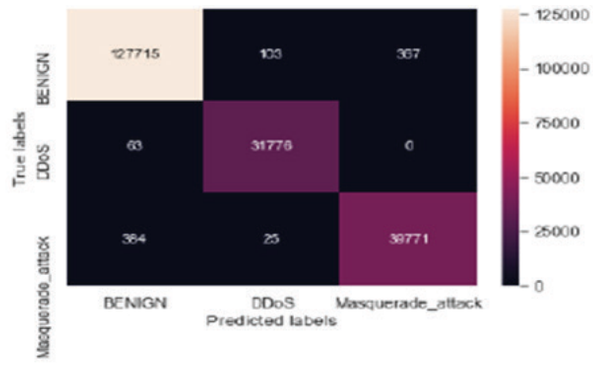

Fig. 2. Confusion matrix graph

```
In [206]: from sklearn.metrics import classification_report
          print(classification_report(y_test, predictions ))
```

	precision	recall	f1-score	support
BENIGN	1.00	1.00	1.00	128185
DDoS	1.00	1.00	1.00	31839
Masquerade_attack	0.99	0.99	0.99	40180
accuracy			1.00	200204
macro avg	0.99	0.99	0.99	200204
weighted avg	1.00	1.00	1.00	200204

Fig. 3. Classification report

between precision and recall, this score considers both False Positives and False Negatives.

Fig. 2 shows the results of the confusion matrix, Fig. 3 details the classification report obtained.

Table I shows this research is in line with Alguliyev, R. M., Aliguliyev, R. M., & Abdullayeva, F. J. (2019), Deep Learning Method for Prediction of DDoS Attacks on social media and Wang and Zhang that posited the accuracy results shown.

TABLE I
EVALUATION COMPARING WITH OTHER RESEARCHERS

Name	Recall	Precision	F-1score	Training accuracy	Testing accuracy
This work (2021)	1.0000	1.0000	1.0000	0.9953	0.9952
Wang and Zhang (2017)	0.3469	0.9297	0.5053	0.9925	0.4026
Aliguliyev and Abdullayeva (2019)	0.8455	0.8923	0.8683	0.7761	0.7744

F. Save the Model

Pickle was be used to save the model so the detection model can be done easily without the need to train the dataset again.

G. Prediction of Detection Model on Social Media Data Simulated on Wireshark

In order to build a proper framework of detecting DDoS and masquerade attacks on social media, a prediction has to be made on a social media network dataset which was captured and simulated on Wireshark.

1) Simulating Facebook Data on Wireshark

The first process on Wireshark was to get the IP address which is specific to Facebook using "IP contains Facebook". While Facebook data was used to test, other social media networks can also be used. Wireshark shows the following features: number, time, source, destination, protocol, length, and info. Wireshark captured all packets and TCP/IP packets being transmitted and received over a particular Facebook webpage network attached to the computer.

Definitions of the features captured by Wireshark:

i. Number: The packet's number in the capture file. Even if a display filter is used, this number will not alter.
ii. Time: The starting point for all succeeding packet time computations.
iii. Source: Instagram gateway IP address.
iv. Destination: Address of destination of packet.
v. Protocol: The type of packet, TCP, DNS.
vi. Length: This shows the length of the packet in byte.
vii. Info: Details on the contents of the packet.

a) Data preprocessing for Facebook

The dataset from Wireshark and necessary libraries was imported into the model and encoding of the categorical data to numerical data took place.

b) Prediction of attack for Facebook dataset

This is the last step of the implementation showing a better approximation of how the model will perform in the real world. It is the prediction using the detection model, and results show that DDoS attack was predicted on the Facebook dataset used to test. The result of this prediction shows that there is 100% attack detected on the Facebook dataset as shown in Fig. 4.

2) Simulating Instagram Data on Wireshark

Wireshark was used to get the IP address which is specific to that of Instagram. Wireshark shows the following features: No, time, source, destination, protocol, length, and info. Wireshark captured all packets and TCP/IP packets, being transmitted, and received over a particular Instagram webpage network attached to the computer.

```
In [212]: df4.dtypes
Out[212]: No.            int64
          Time         float64
          Source        object
          Destination   object
          Protocol      object
          Length         int64
          Info          object
          dtype: object

In [213]: dfencode=pd.get_dummies(df4)

          dfencode.shape
Out[213]: (15, 10)

In [214]: dfencode.columns
Out[214]: Index(['No.', 'Time', 'Length', 'Source_192.168.43.204',
                 'Destination_102.132.101.10', 'Destination_102.132.101.15',
                 'Destination_102.132.101.16', 'Destination_102.132.101.35',
                 'Protocol_TLSv1.3', 'Info_Client Hello'],
                dtype='object')

In [215]: model.predict(dfencode)
Out[215]: array(['DDoS', 'DDoS', 'DDoS', 'DDoS', 'DDoS', 'DDoS', 'DDoS', 'DDoS',
                 'DDoS', 'DDoS', 'DDoS', 'DDoS', 'DDoS', 'DDoS', 'DDoS'],
                dtype='<U17')
```

Fig. 4. Snapshot of data prediction (facebook)

Fig. 5. Snapshot of data prediction (Instagram)

a) Data preprocessing for Instagram

The Instagram dataset instances require necessary libraries, which were imported into the model and after which data encoding also took place.

b) Prediction of attack for Instagram dataset

The prediction using the detection model shows that DDoS attack was predicted on the Instagram dataset used to test. The result of this prediction shows that there is a 100% attack detected on the Instagram dataset, as shown in Fig. 5.

V. Conclusion

Convolutional neural network (CNN) is a highly effective deep learning algorithm, as it can classify, learn from models, make inferences, and make predictions which are suitable for predicting previously unknown data. A prediction model which shows how DDoS and masquerade attacks can be detected was developed using the CICIDS2017 dataset. The dataset comprises very recent attacks. The dataset was analyzed and later used in the prediction of attack on a sample social media network with an accuracy of 99.53%.

Future researchers should focus on specific areas of attacks, especially masquerade attacks, because there is no current dataset depicting data on masquerade attacks. Acquiring a dataset which contains social media network data was a challenging feat in this project.

Funding

This article did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

Authors declare that they have no conflict of interest.

References

[1] Mahadev and V. Kumar, “Behaviour Analysis of DDoS Attack and Its Detection,” *Int. J. Comput. Sci. Eng.*, vol. 6, no. 5, pp. 139-144, May 2018, doi: 10.26438/ijcse/v6i5.139144.

[2] F. Shahzad, M. Pasha, and A. Ahmed, “A Survey on Active Attacks on Wireless Sensor Networks and their Countermeasures,” *Int. J. Comput. Sci. Inf. Secur.*, vol. 14, no. 12, pp. 54-65, Dec. 2016.

[3] D. Irani, M. Balduzzi, D. Balzarotti, E. Kirda, and Calton Pu, “Reverse Social Engineering Attacks in Online Social Networks,” in *8th Int. Conf. Detect. Intrusions Malware Vulnerability Assessment*, in Detection of Intrusions and Malware, and Vulnerability Assessment, T. Holz and H. Bos Eds, in Lecture Notes in Computer Science, vol. 6739, 2011, doi: 10.1007/978-3-642-22424-9_4.

[4] H. Nissenbaum, “A Contextual Approach to Privacy Online,” *Daedalus*, vol. 140, no. 4, pp. 32-48, Oct. 2011, doi: 10.1162/DAED_a_00113.

[5] R. Alguliyev and F. J. Abdullayeva, “Deep Learning Method for Prediction of DDoS Attacks on Social Media,” *Adv. Data Sci. Adapt. Anal.*, vol. 11, no. 01n02, 2019, doi: 10.1142/S2424922X19500025.

[6] M. B. Salem and S. J. Stolfo, “Decoy Document Deployment for Effective Masquerade Attack Detection,” in *8th Int. Conf. Detect. Intrusions Malware Vulnerability Assessment*, in Detection of Intrusions and Malware, and Vulnerability Assessment, T. Holz and H. Bos Eds, in Lecture Notes in Computer Science, vol. 6739, 2011, doi: 10.1007/978-3-642-22424-9_3.

[7] M. Alkasassbeh, G. Al-Naymat, A. B. Hassanat, and M. Almseidin, “Detecting Distributed Denial of Service Attacks Using Data Mining Techniques,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 1, pp. 436-445, 2016, doi: 10.14569/IJACSA.2016.070159.

[8] V. Bhuse, A. Gupta and A. Al-Fuqaha, "Detection of Masquerade Attacks on Wireless Sensor Networks," in

2007 IEEE Int. Conf. Commun., 2007, pp. 1142-1147, doi: 10.1109/ICC.2007.194.

[9] B. Bowen, M. Ben Salem, S. Hershkop, A. Keromytis and S. Stolfo, "Designing Host and Network Sensors to Mitigate the Insider Threat," *IEEE Secur. Priv.*, vol. 7, no. 6, pp. 22-29, Nov.-Dec. 2009, doi: 10.1109/MSP.2009.109.

[10] I. C. Eian, K. Y. Lim, M. X. L. Yeap, H. Q. Yeo, and F. Zahra, "Wireless Networks: Active and Passive Attack Vulnerabilities and Privacy Challenges," *Preprints*, Oct. 2020, doi: 10.20944/preprints202010.0018.v1.

[11] X. Hu, A. Kim, N. Siwek, and D. Wilder, "The Facebook Paradox: Effects of Facebooking on Individuals' Social Relationships and Psychological Well-Being," *Front. Psychol.*, vol. 8, Article 87, Jan. 2017, doi: 10.3389/fpsyg.2017.00087.

[12] K. Ahmed, S. Verma, N. Kumar, and J. Shekhar, "Classification of Internet Security Attacks," in *Proc. 5th Natl Comput. Nation Dev.*, Delhi, India, Mar. 10-11, 2011.

[13] S. Kumarasamy and R. Asokan, "Distributed Denial of Service (DDOS) Attacks Detection Mechanism," *Int. J. Comput. Sci. Eng. Inf. Technol.*, vol. 1, no. 5, Dec. 2011, doi: 10.5121/ijcseit.2011.1504.

[14] T. Mahjabin, Y. Xiao, G. Sun, and W. Jiang, "A survey of distributed denial-of-service attack, prevention, and mitigation techniques," *Int. J. Distrib. Sens. Netw.*, vol. 13, no. 12, Dec. 2017, doi: 10.1177/1550147717741463.

[15] J. A. Obar and S. Wildman, "Social media definition and the governance challenge: An introduction to the special issue," *Telecommun. Policy*, vol. 39, no. 9, Oct. 2015, pp. 745-750, doi: 10.1016/j.telpol.2015.07.014.

[16] S. S. Sahu, and M. Pandey, "Distributed Denial of Service Attacks: A Review," *Int. J. Mod. Educ. Comput. Sci.*, vol. 6, no. 1, pp. 65-71, Jan. 2014, doi: 10.5815/ijmecs.2014.01.07.

[17] A. R. Shaaban, E. Abd-Elwanis and M. Hussein, "DDoS attack detection and classification via Convolutional Neural Network (CNN)," in *2019 Ninth Int. Conf. Intell. Comput. Inf. Syst. (ICICIS)*, 2019, pp. 233-238, doi: 10.1109/ICICIS46948.2019.9014826.

[18] S. Bravo and D. Mauricio, "Distributed Denial of Service Attack Detection in Application Layer Based on User Behavior," *Webology*, vol. 15, no. 2, pp. 38-53, Dec. 2018.

[19] Nielsen, "State of The Media: The Social Media Report," 2011. [Online]. Available: https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/nielsen-social-media-report.pdf

[20] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in *Proc. 26th Annu. Comput. Secur. Appl. Conf.*, Texas, USA, Dec. 2010, pp. 1-9, doi: 10.1145/1920261.1920263.

[21] J. E. Tapiador and J. A. Clark, "Masquerade mimicry attack detection: A randomised approach," *Comput. Secur.*, vol. 30, no. 5, pp. 297-310, July 2011, doi: 10.1016/j.cose.2011.05.004.

[22] Z. Wang and Y. Zhang, "DDoS Event Forecasting using Twitter Data," in *Proc. 26th Int. Jt. Artif. Intell.*, pp. 4151-4157, doi: 10.24963/ijcai.2017/580.