Statistical Anomaly Detection of DDoS Attacks Using K-Nearest Neighbour

Thwe Thwe Oo
Faculty of Information and Communication Technology
University of Technology (Yatanarpon Cyber City)
Pyin Oo Lwin, Myanmar
thweoo001@gmail.com

Thandar Phyu
Department of Advanced Science and Technology
Ministry of Science and Technology
Nay Pyi Taw, Myanmar
thandar.phyu@gmail.com

Abstract—Distributed Denial of Service (DDoS) attacks have strong influence on Internet Security because these attacks affect the normal functioning of organizations causing billions of dollars of losses. Although these organizations were well-equipped in security, they were damaged by DDoS attacks. In this paper, the proposed system presents both detecting and classifying sheme of DDoS attack using K-NN. The two proposed algorithms are developed based on various features of attack packets obtained from study the incoming and outgoing network traffic and used K-Nearest Neighbour to analyze these features. The main objectives of this paper are to analyze the DDoS attacks natures and to detect and identify types of DDoS attacks.

Keywords: Internet Security, DDoS attacks, Attack Packets

1. INTRODUCTION

The Internet was designed with functionality, not security. It offers its participants fast, easy and cheap communication mechanism, enforced with various higher level protocols that ensure reliable or timely delivery of messages or a certain level of quality of service. Hence, malicious users exploit this weakness to achieve their purpose. In recent years, the number of network-based threads, especially DDoS attacks, is one of the major type of attack. As Distributed Denial of Service (DDoS) attack is an increasing problem of the current internet, the role of Intrusion Detection System to detect anomalies and attacks in the network, is becoming major role in networking security.

The DDoS attacks usually do not exploit of security vulnerabilities of network-connected system, but instead they aim to disrupt victim services by overwhelming the processing capacity of system or by flooding the bandwidth of the target or by scanning vulnerable hosts, such as SYN Flooding, SYN Scanning, and so on. So it is needed to analyze TCP connections. In this paper, the system approach is to classify attack or normal packet and identify types of DDoS attacks, especially for flooding and scanning attacks. The proposed system is based on statistical approach. The advantages of this approach are:

1. Unknown attacks can be discovered.
2. The proposed system can be analyzed the detail of DDoS attack characteristics based on statistical value.

(3) As DDoS attack is a process changing dynamically and frequently, statistical approach is the most suitable for DDoS detection.

(4) Statistical approach has to address Threshold Setting: Anomalies are detected when the current system state differs from the model by a certain threshold.

This paper is organized as follows: Section 2 presents some previous researches concerning to DDoS attack detection. Section 3 describes the system architecture and finally concludes the paper in Section 5.

2. RELATED WORK

This section describes the previous approaches to detect the attacks.

Through scanning, an attacker obtains information on a target system. By sending scanning packets to the target, it discovers which systems are working and which services are being offered [5]. DoS/DDoS attacks cause a waste of the resources in the host or networks, and make services work improperly. There are two principal classes of DoS/DDoS, logic and flooding attacks [6].

Logic attacks exploit existing software flaws to cause a malfunction in the system. For instance, in a Ping-of-Death attack [7], oversized ICMP ping packets can result in a denial of service. Also, a Land attack [8] may crash the system by sending packets with the source host and port the same as the destination host and port.

Flooding attacks transmit many spurious packets to the target system, and waste CPU, memory, and network resources. In case of TCP SYN flood [9], the victim receives packets that exceed buffer of the data structure limit and cripple its service. Also, ICMP, TCP, and UDP flooding attacks [11] overwhelm bandwidth by sending useless traffic to the victim. Some attacks, such as Smurf [10] and Fraggle [12], amplify traffic by using reflecting services of the third party. There are other examples of a reflecting attack that cause packets to rebound between two hosts using reflecting ports, such as echo.

Karimazad and Faraahi [3] proposed an anomaly based DDoS detection method based on features of attack packets, analyzing them using Radial Basis Function (RBF) neural networks. Vectors with seven features are used to activate an RBF neural network and classify traffic into normal and DDoS attack traffic classes. They evaluated the approach using UCLA Datasets. Their system can be classified either normal.
or attack, but that can’t be classified and identified what types of attacks.

In [1], this approach is a statistical approach based on several features values. The proposed system only showed features extraction module and saved these features into database to identify normal and attack packets. In [2], the proposed system is a combined data mining approach to detect protocol anomaly against DDoS attack. In this paper, traffic features are extracted from network traffic and then it is clustered into normal and attack traffic by using a data mining classification algorithm. This paper only shows detection phrase using only proposed algorithm. Now, in current paper, the proposed system is statistical detection of attacks using the two proposed algorithms and K-NN. The results show that our system can detect and classify types of DDoS attack with high accuracy.

### 3. SYSTEM ARCHITECTURE

The system architecture is shown in Figure 1. It can be divided into five main modules. They are Collection of packets, Features Extraction, Attack classification by using Packet Classification Algorithm, Estimate using K-NN Classifier and DDoS attack classification.

![Flowchart of the DDoS Detection System](image)

**Figure 1: Overall architecture of the Proposed DDoS Classification System**

#### 3.1. Packet Collection Module

The system collects the incoming and outgoing network packets in every 20 seconds. We evaluate the proposed method using UCLA datasets[4], which is widely used as one of the publicly available data sets for network based anomaly detection system.

#### 3.2. Feature Extraction Module

This section extracts the various features for DDoS detection. These features are very suitable to distinguish abnormal behaviors from variation of normal behaviors. We studied the various traffic parameters and in[ref], Karimazad et al. has mentioned some parameters to detect the DDoS attack packets in traffic data. We choose to use these features as he suggested because the analysis of traffic based on these features can recognize the attack in packets. These features are shown in the following:

1) **Average Packet Size:** DDoS attacks flood victim to consume system resources, then Average Packet Size increases in attack time. We use this feature to identify DDoS attacks.

2) **Number of Packets:** DDoS attacks send a great number of packets to the victim network. Therefore, the number of packets increases in comparison to normal case. We exploited this feature to detect DDoS attacks.

3) **Time Interval Variance:** The experiments show when DDoS attack launches, agents send attack packets in the same time span. Then we can detect this attacks using Time Interval Variance. Whenever packet sending time spans are more similar together, Time Interval Variance will be closer to zero. Variance can be calculated through (1):

\[
\tau_{\sigma} = \sqrt{\frac{\sum (t_n - \bar{t})^2}{n}}
\]

4) **Packet Size Variance:** According to our studies, we found that attack packets sizes are the same. However, normal packets have different sizes even when they belong to the same file. DDoS packets can be identified by using the Packet Size Variance.

5) **Number of Bytes:** Increase in number of bytes demonstrates launching DDoS attacks.

6) **Packet Rate:** This feature shows the packet rate sent from a source address to a destination in a specific time span. Packet rate increases significantly in attack time.

7) **Bitrate:** A very high rate of this feature indicates launching DDoS attack.

Along with the above mentioned features, the system also uses the number of SYN, ACK and FIN flags packets to classify attack types such as SYN Flooding, ACK Scanning and so on.

#### 3.3. Packet Classification Algorithm

After the feature extraction, proposed classification algorithm is applied to classify traffic into normal, other and attack traffic groups. The classification is simple. Our idea is to analyse each flow and classify as attack when thresholds have been exceeded. This threshold is determined by experience after long observation of the traffic spectrum. In this system, we observe only for two types of traffics within the flow TCP and UDP. The classification scheme is shown in Table 1.
Table.1: Classification Scheme

<table>
<thead>
<tr>
<th>No. of pkt s</th>
<th>Avg pkt size</th>
<th>Ti-me Inte-rval size</th>
<th>Pack et size</th>
<th>Pack et rate per sec</th>
<th>No. of Flag packets</th>
<th>Clas-s</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>L</td>
<td>&gt;0</td>
<td>L</td>
<td>&lt;α</td>
<td>L</td>
<td>Normal</td>
</tr>
<tr>
<td>H</td>
<td>H</td>
<td>&lt;0</td>
<td>&lt;0</td>
<td>H</td>
<td>&gt;λ</td>
<td>Attack</td>
</tr>
</tbody>
</table>

where

\[ \alpha = \text{minimum packet rate} \]
\[ \lambda = \text{maximum packet rate} \]
\[ L = \text{LOW} \]
\[ H = \text{HIGH} \]

3.4. K-Nearest neighbors

K-nearest neighbor is a supervised learning algorithm where the result of new instance query is classified based on majority of K-nearest neighbor category. The purpose of this algorithm is to classify a new object (document) based on attributes (words) and training samples. It works by using minimum distance from the query instance to the training samples to determine the K-nearest neighbors [17],[ref pxc38811105]

3.5. DDoS Attack Classification

Finally, the proposed attack classification algorithm classifies and identifies what types of attacks.

- IF (no-packets==HIGH) AND (avg-packet-size with same Destination IP==HIGH) AND (protocol==UDP)
  THEN UDP Flooding.
- IF (no-packets==HIGH) AND (avg-packet-size with same Destination IP==HIGH) AND (protocol==TCP)
  Begin
    IF(no-ACK with same destination port==HIGH)
    THEN TCP ACK Flooding
    IF (no-SYN with same destination port==HIGH)
    THEN TCP SYN Flooding
  End
- IF (no-packets==HIGH) AND (no-destination-IP==HIGH) AND (no-destination-port==LOW)
  THEN Network Scanning Attack

4. EXPERIMENTAL RESULTS

This section describes the performance evaluation results of the proposed DDoS attack classification and identification system. We evaluate our system with traffic traces of DDoS attacks using UCLA data sets.

4.1. Intra-Cluster Similarity Analysis of K-NN

In order to analyze the performance of K-NN clustering algorithms, K-NN algorithm is implemented on the UCLA training data set and is tested with various numbers of clusters ranging from two to six. Root Mean Square Error criterion which described in the following equation is used for the analysis of KNN algorithm in terms of intra-cluster similarity. The lower RMSE can give the better intra-cluster quality.

\[ E = \frac{1}{N} \sum_{i=1}^{N} (\sum_{p \in C_i} | m_i - p |^2) \]

\[ E = \text{the average of the sum of the square error for all samples in the data set} \]
\[ P = \text{the point in space representing a given sample} \]
\[ m_i = \text{the mean of cluster} \]
\[ N = \text{the number of samples in cluster}. \text{RMSE} \text{is the measure of the within cluster variability} \]

Predicted Class | Actual Class
--- | ---
Yes | No
Yes | TP
No | FN

<table>
<thead>
<tr>
<th>Predicted Class</th>
<th>Actual Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
<td>TP</td>
</tr>
<tr>
<td>No</td>
<td>FN</td>
</tr>
<tr>
<td>No</td>
<td>TN</td>
</tr>
</tbody>
</table>
The clustering accuracy for measuring the clustering results was defined as:

\[ r = \frac{\sum_{i=1}^{c} a_i}{n} \]  

\( n \) = number of samples in the dataset, 
\( a_i \) = number of data samples occurring in both cluster \( i \) and its corresponding class, which had the maximal value, 
\( c \) = number of cluster 
\( r \) = accuracy.

**Figure 2:** Intra-Cluster Similarity Analysis of K-NN

**4.2. Evaluation of Attack Detection**

The performance evaluation of proposed system using UCLA Dataset is evaluated according to classification accuracy rate and recall rate. Its measurement using F-measure with matrix confusion standard are shown in the Table 2.

**Table 2: Evaluation result classification using KNN**

<table>
<thead>
<tr>
<th>Category</th>
<th>k</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attack</td>
<td>2</td>
<td>0.6311</td>
<td>0.6476</td>
<td>0.6392</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.6528</td>
<td>0.6467</td>
<td>0.6497</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.6420</td>
<td>0.6387</td>
<td>0.6403</td>
</tr>
<tr>
<td>Normal</td>
<td>2</td>
<td>0.6211</td>
<td>0.6174</td>
<td>0.6192</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.5976</td>
<td>0.6062</td>
<td>0.6019</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.4560</td>
<td>0.4437</td>
<td>0.4498</td>
</tr>
</tbody>
</table>

**Table 3 and Figure 4 show the classification accuracy using KNN only.**

**Table 3: Evaluation result classification using KNN**

<table>
<thead>
<tr>
<th>Category</th>
<th>k</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attack</td>
<td>2</td>
<td>0.4322</td>
<td>0.5422</td>
<td>0.4810</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.5431</td>
<td>0.5611</td>
<td>0.5520</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.6210</td>
<td>0.6387</td>
<td>0.6297</td>
</tr>
<tr>
<td>Normal</td>
<td>2</td>
<td>0.5310</td>
<td>0.5254</td>
<td>0.5282</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.6211</td>
<td>0.6174</td>
<td>0.6192</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.6190</td>
<td>0.6082</td>
<td>0.6136</td>
</tr>
</tbody>
</table>

**Table 4 and Figure 5 show the classification accuracy using KNN only and Packet Classification Algorithm.**

**Table 4: Evaluation result classification using KNN**

<table>
<thead>
<tr>
<th>Category</th>
<th>k</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attack</td>
<td>2</td>
<td>0.4322</td>
<td>0.5422</td>
<td>0.4810</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.5431</td>
<td>0.5611</td>
<td>0.5520</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.6210</td>
<td>0.6387</td>
<td>0.6297</td>
</tr>
<tr>
<td>Normal</td>
<td>2</td>
<td>0.5310</td>
<td>0.5254</td>
<td>0.5282</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.6211</td>
<td>0.6174</td>
<td>0.6192</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.6190</td>
<td>0.6082</td>
<td>0.6136</td>
</tr>
</tbody>
</table>
Attack traffic detection has some error rates because some traffic in DDoS attack are seemingly legitimate traffic. In the case of open services, such as web servers, the DDoS attacker only needs to send large quantities of useless service requests. Thus, there might be no specific features of DDoS attack traffic that the classifier or proposed algorithm can be instructed to detect as attack. With such malicious but legitimate traffic, DDoS attackers are able to relatively easily bypass most means of DDoS defense. Our detection system often misdetect these kinds of traffic as normal but the aggregation of these incoming flows or outgoing flows can exhibit that each of seeming legitimate traffic are attack. So that, our proposed attack classification algorithm in next phrase can detect and identify correctly these traffic because the proposed approach is based on the flow information.

### 4.3. Evaluation of Attack Type Classification

This section describes the classification of types of DDoS attacks by using attack classification algorithm. The performance of proposed system using UCLA Dataset is evaluated according to

1. Sensitivity
2. Specificity and
3. Accuracy.

Sensitivity is the probability of a positive test among traffic flows.

\[
sensitivity = \frac{t_{pos}}{pos}
\]

Specificity is the probability of a negative test among traffic flows.

\[
specificity = \frac{t_{neg}}{neg}
\]

\[t_{pos}\] is the number of true positives, is the number of positive samples,

\[t_{neg}\] is the number of true negatives, is the number of negatives samples.

It can be shown that accuracy is a function of sensitivity and specificity:

\[accuracy = sensitivity \times \frac{pos}{(pos + neg)} + specificity \times \frac{t_{neg}}{neg}\]

### 5. CONCLUSION

This paper implements a system that analyzes network traffic and classify the network packets as normal and DDoS attack packets. In particular, we propose a combined approach to detect traffic anomaly against DDoS attacks and to classify attack traffic according to attack types. Proposed system has two main phrase; detection phrase and classification phrase. First seven packet features that exhibit the traffic behaviors apparent to DDoS attack are extracted by analyzing packets. Then these features are analyzed to distinguish abnormal behavior from variations of normal behavior. The system apply KNN classifier that is trained to learn the difference between normal and abnormal traffic features and proposed packet classification algorithm to cluster the traffic as normal and attack. Then attack traffic is created based on the output cluster set. Second, the DDoS attack traffics are clustered as attack types; especially, UDP and TCP flooding attacks and scanning attacks using the proposed flow based attack classification algorithm.

The proposed system is evaluated using UCLA Dataset, which is widely used as one of the few publicly available data sets for network based anomaly detection system. The results show that our system could detect attack traffic with 96% accuracy, and also get high accuracy in classifying the attack packets as the DDoS attack types.

### REFERENCES


