STATISTICAL-BASED MECHANISM FOR DETECTING HYPER TEXT TRANSFER PROTOCOL DDOS ATTACKS

AYMAN IBRAHIM ALI GHABEN

UNIVERSITI SAINS MALAYSIA

2023

STATISTICAL-BASED MECHANISM FOR DETECTING HYPER TEXT TRANSFER PROTOCOL DDOS ATTACKS

by

AYMAN IBRAHIM ALI GHABEN

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

January 2023

ACKNOWLEDGEMENT

In the name of Allah (SWT), the most gracious and merciful, I am thankful to my Creator, who blessed me with the ability to complete my Doctorate.

I would like to express my heartfelt gratitude to Dr. MOHAMMED ANBAR for his significant contributions through the supervision of this thesis. He made invaluable inputs into the research through his guidance from the start of the research proposal to the ultimate completion of the thesis. He challenged me many times while writing the thesis and helped me refine and shape my ideas to fit the research parameters I was doing. In addition, he gave me valuable but detailed feedback. On top of that, he passed valuable literature he found related to my research topic and constantly reminded me of the importance of reflecting critically on my work. Through his guidance, I learned valuable research experience and skills to stay with me throughout my life.

I pay my profound gratitude and recognition to my co-supervisor, Dr. Shankar, who led my research activities in fixing the real research gaps. His precious time, advice, and concerns helped me immensely in improving and completing this thesis.

I dedicate this special document of my education to my parents M IBRAHIM GHABEN (father) and Madam INTISAR GHONAIM (mother), my lovely wife, REHAM GHONAIM, and my close family KHALED, ZAINA, TARIQ, and BILAL.

Lastly, to all my brothers who inspired, encouraged, assisted, and guided me in one way or the other but whose names could not be mentioned here, thank you, and may God richly bless us all.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

AWS	Amazon Web Services
C&C	Command and Control
DDoS	Distributed Denial of Services
DFT	Discrete Fourier Transform
DoS	Denial of Service
DWT	Discrete Wavelet Transforms
FN	False-Negative
FP	False-Positive
HTTP	Hypertext Transfer Protocol
IRC	Internet Relay Chat
SDG	Sustainable Development Goals
TN	True-Negative
TP	True-Positive
VNI	Cisco Visual Networking Index

LIST OF APPENDICES

MEKANISMA BERASASKAN STATISTIK UNTUK MENGESAN SERANGAN DDOS PROTOKOL PINDAHAN HIPERTEKS

ABSTRAK

Sejak sedekad yang lalu, pergantungan pada Internet semakin meningkat dengan ketara. Pada masa yang sama, terdapat peningkatan ancaman keselamatan yang bertujuan menghalang pengguna internet daripada mendapat akses kepada perkhidmatan. Ancaman keselamatan yang paling intrusif di Internet ialah serangan Nafi Perkhidmatan Teragih (DDoS). Pada masa ini, serangan DDoS pembanjiran Protokol Pemindahan Teks Hiper (HTTP) pada pelayan web adalah salah satu serangan siber terancang yang paling banyak mendatangkan kerosakan terhadap perkhidmatan dalam talian atau komputer pada rangkaian. Walaupun terdapat banyak pendekatan untuk mengesan serangan DDoS pembanjiran HTTP, masalah ini masih berleluasa kerana ia boleh memperdaya pendekatan pengesanan dengan meniru tingkah laku pengguna biasa. Oleh itu, tesis ini mencadangkan satu pendekatan untuk mengesan serangan DDoS pembanjiran HTTP ke atas pelayan web. Pendekatan yang dicadangkan terdiri daripada lima fasa untuk mencapai matlamat penyelidikan, seperti berikut: (1) Pra-pemprosesan data, (ii) Atribut paket agregat bertujuan untuk mengagregatkan paket setiap (t) masa berdasarkan tiga atribut iaitu (a) saiz paket, (b) keteraturan (antara masa ketibaan), dan (c) bilangan paket (iii) Pengesanan berasaskan anomali menggunakan empat penunjuk iaitu: (a) penjumlahan baris-lajur, (b) entropi Bayes, (c) pencongan agihan paket, dan (d) nombor Reynolds (iv) Mekanisme berasaskan pengundian, dan (v) Mekanisme berasaskan statistik. Mekanisme yang dicadangkan telah dinilai menggunakan dua set data penanda aras (CIC DDoS dan ISCX) dan hasilnya menunjukkan kadar ketepatan pengesanan adalah 96.03% dan 94.28% untuk set data CIC DDoS dan ISCX, masing-masing.

Tambahan lagi, kadar positif palsu ialah 14.28% dan 10.00% untuk set data tersebut.

STATISTICAL-BASED MECHANISM FOR DETECTING HYPER TEXT TRANSFER PROTOCOL DDOS ATTACKS

ABSTRACT

Over the last decade, the reliance on the Internet is noticeably increasing. On the other hand, there is also increases in the security threats aim to prevent the internet users from getting access to the services. The most intrusive security threat on the Internet is the Distributed Denial of Services (DDoS) attacks. At present, the Hyper Text Transfer Protocol (HTTP) DDoS flooding DDoS attack on webservers is one of the most destructive organized cyberattacks against online services or computers on the network. Despite the existence of many approaches to detect HTTP flooding DDoS attacks, the problem is still prevalent because it can bypass the detection approach by mimicking the behaviour of normal users. Thus, this thesis proposes an approach to detect HTTP flooding DDoS attacks on web servers. The proposed approach consists of five phases to achieve the goal of the research, as follows: (1) Data pre-processing, (ii) Aggregated packets attributes aim to aggregate the packets every (t) time based on three attributes which are (a) packet size, (b) regularity (inter arrival time), and (c) number of packets (iii) Anomaly-based detection using four indicators which are : (a) summation rows-columns, (b) Bayes-entropy, (c) skew of the packets distribution, and (d) Reynolds number) (iv) voting-based mechanism, and (v) statistical based mechanism. The proposed mechanism has been evaluated using two benchmark datasets (CIC DDoS and ISCX) and the results reveal that the detection accuracy rates are 96.03% and 94.28% when evaluated over

CIC DDoS and ISCX datasets, respectively. Furthermore, the false positive rates are 14.28%, 10.00% when evaluated over those datasets.

CHAPTER 1

INTRODUCTION

This chapter introduces the research topic and the key aspects related to the proposed mechanism for detecting the Hypertext Transfer Protocol (HTTP) flooding distributed denial of service (DDoS) attacks in aggregated packets captured from the traffic in nine sections. First, Section 1.1 is the overview, followed by Section 1.2 and Section 1.3, which provide the research background and motivation. Next, Section 1.4 until Section 1.6 presents the research problems, objectives and goals, and contributions. Next, section 1.7 lists the research scope and limitations, then Section 1.8 briefly describes the research methodology, and finally, Section 1.9 summarizes this chapter.

1.1 Overview

The dependency on the Internet in all kinds of human activities has grown exponentially in the last few years. The growth is spurred by demands from all sectors (public and private) and the vast proliferation of emerging technologies such as smart mobile devices, financial technology (fin-tech), advanced communication technologies, and the Internet of Things and its applications. Unfortunately, the dramatic increase in Internet-connected devices and applications also brought many challenges, such as exposure to various cyber threats. The most common cyber threats are denial of service (DoS) and DDoS attacks. These types of attacks deliberately target the availability of network services or resources for legitimate users.

Consequently, many organizations had to spend substantial effort and resources to secure and protect their network services from such attacks. However, identifying and stopping these attacks is challenging due to the dynamic nature of these attacks. Therefore, to address the threat of DoS/DDoS attacks, understanding their nature is necessary before designing and engineering an effective countermeasure.

DoS and DDoS attacks are the two most dangerous and destructive threats affecting people, businesses, and governments (Alomari et al., 2012). DoS attacks are a persistent problem on the Internet, as evidenced by the continued attacks on commercial servers and ISPs. In addition, network outages have denied many users' accesses to essential online services and network resources, such as cloud services, web servers, mail servers, and domain resolution services. However, the most dangerous attack is the DDoS attack, the purpose of which is to deny legitimate users access to online services or resources (Park & Lee 2001).

The first well-documented DDoS attack occurred in August 1999. On 7th February the following year, believed to be the first large-scale DDoS attack – the Internet portal Yahoo! was unavailable within three hours of the attack. Subsequently, Amazon, Buy.com, CNN, and eBay were attacked by DDoS and shut down their web servers the next day (Bhuyan et al., 2014).

1.2 Background

In the cyber security science, there are many attacks against the webservers happened because of flooding DDoS attacks which can be caused by high-rate or low-rate, botnet is most commonly method used to trigger DDoS attacks. So, this section presents an introduction to Botnet, DoS, and DDoS and HTTP Get flooding DDoS attacks in Section 1.2.1, Section 1.2.2, and Section 1.2.3, respectively.

1.2.1 Botnet

A botnet is a collection of distributed infected devices connected to the Internet and controlled by a bot master through a command and control (C&C) server via a reliable system using communication protocols, such as DNS. Botnets engage in various malicious activities, including spam, and DoS/DDoS attacks, as shown in Figure 1.1. DDoS attacks are common attacks triggered by Botnet, and their severity increases as more internet-connected devices are compromised (Barford &Yegneswaran 2007; Kwon et al., 2017).

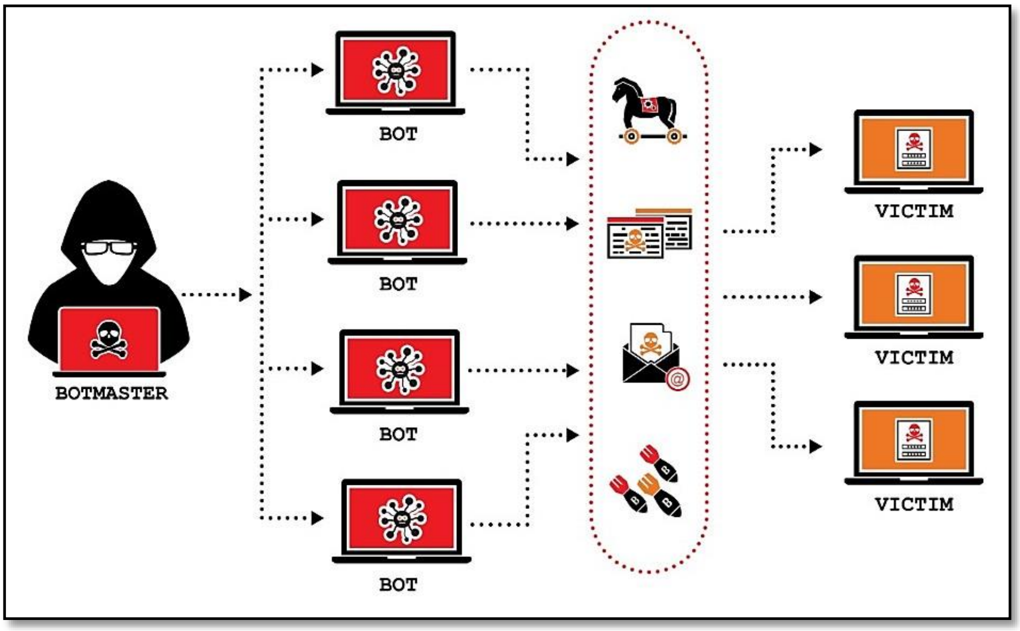

Figure 1.1 Botnet Malicious Activities

1.2.2 Denial of Service (DoS) and Distributed Denial of Services (DDoS)

DoS attacks are among the oldest types of botnet activities that pose a serious and continuous threat to users, organizations, and internet infrastructure that completely disrupt users' network connectivity or deny them access to network services (Elejla et al., 2018).

After launching a DoS attack, the attacker will eventually learn the defense mechanism that prevents and mitigate DoS attacks. Then, the attacker will launch a new DoS attack that uses distributed traffic attacks to avoid losing opportunities for future DoS attacks if detected and blocked by the network’s security mechanism. DDoS attacks are many forms of DoS attacks, which are large-scale coordinated attacks on the availability of Internet services and resources. DDoS uses the same technology as DoS, but on a larger scale, and come from multiple sources or locations simultaneously (Mirkovic & Reiher 2004). Figure 1.2 shows the general scenario of a DDoS attack.

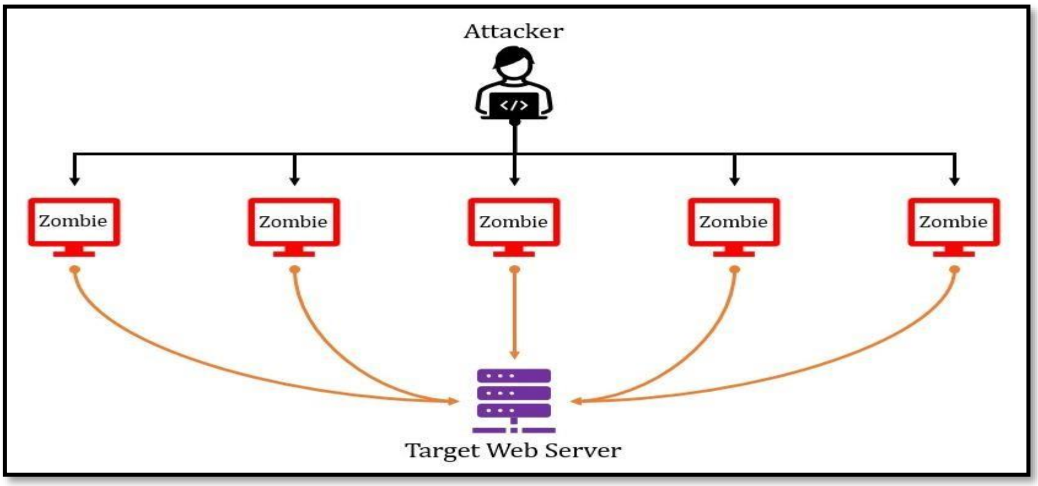

Figure 1.2 DDoS Attack

The DDoS attacks can be categorized into two classes as follows:

Network layer DDoS attacks attempt to disconnect legitimate users by exhausting the victim's network bandwidth or using specific functions in the victim's log. Sometimes, an attacker who usually uses IP address spoofing will send many spoofed packets to the victim's server (Lee & Lee, 2011).

Application layer DDoS attacks target the application layer to exhaust server resources, such as CPU, memory, and disk space. Generally, these attacks impact the service (communication interruption) because they target specific applications, such as HTTP, MQTT, XMPP, CoAP, and DNS (Asad et al., 2019).

1.2.3 HTTP DDoS attacks

These HTTP flooding DDoS attacks often rely on a botnet. These types of DDoS attacks target servers or applications by making them busy responding to tremendous requests, this will make the server unable to process the requests sent by legitimate users on time, resulting in the server dropping the legitimate requests due to time out (Esraa Alomari et al., 2012). In this way, the attacker hopes to overwhelm the server.

Generally, HTTP flooding DDoS attacks are challenging to detect and block for several reasons. First, HTTP flooding DDoS attacks use standard URL requests, making it challenging to distinguish between regular and malicious URL requests. Second, they do not rely on malformed packets, spoofing, or reflection techniques. Finally, they are challenging to detect since traffic volume in HTTP flooding DDoS attacks is generally below detection thresholds. Thus, standard detection mechanisms such as rule-based mechanisms that rely on fixed thresholds are ineffective in detecting these HTTP flooding DDoS attacks (Zhang et al., 2020).

1.3 Research Motivation

Most people and companies have become entrenched and more dependent on Internet-connected devices in the current web era. As a result, their daily life has gradually changed to virtual life on different networks, whether companies, governments, or academia. Unfortunately, there are continuous attempts to attack the networks by malicious users by targeting the weaknesses of these networks. Therefore, there is a pressing need to propose approaches with high detection accuracy to detect network attacks, especially the HTTP flooding DDoS attacks.

Day after day, the targets of HTTP flooding DDoS attacks keep increasing, including education, health, intellectual property, commercial, financial, and industrial domains. Figure 1.3 highlights the domains that experienced devastating HTTP flooding DDoS attacks in Mach 2018.

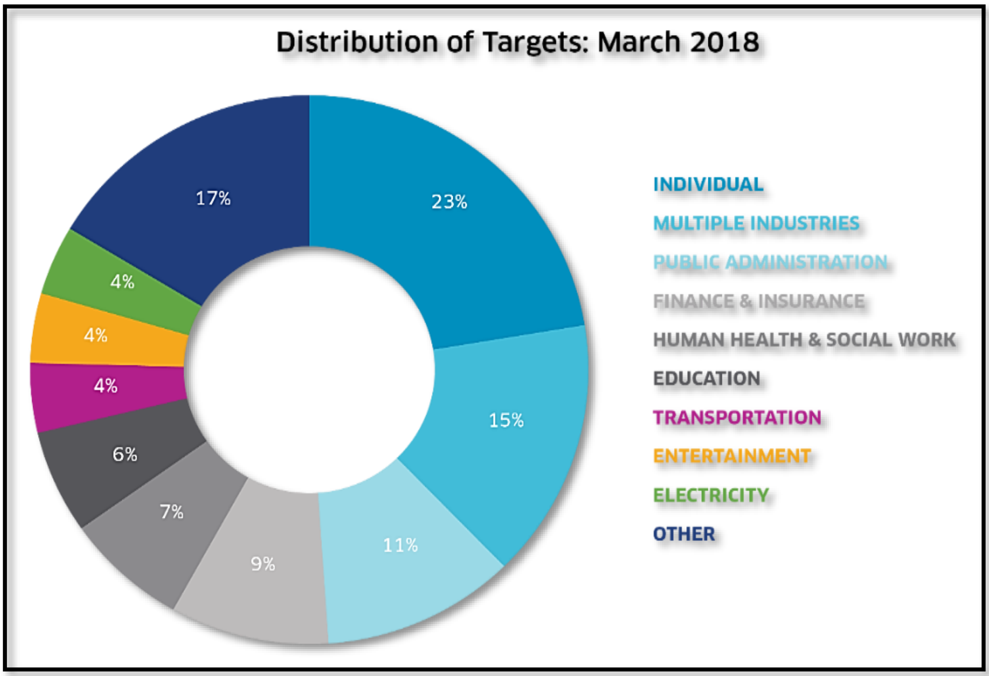

Figure 1.3 Targets of DDoS attacks

The issue of DDoS attacks linked to HTTP flooding will continue to exist if it is not resolved, which will decrease dealer confidence in their businesses, cause them to leave the market with significant losses, force them to close, cause the loss of their employees' jobs, and make it more difficult for the country to achieve the Sustainable Development Goals (SDG). To attain sustainable development goals in society for future generations, the nation's policies must motivate and encourage researchers to solve the problem of attacks by proposing efficient detecting techniques.

For example, in 2013, a reflection DDoS attack, which is thought to be the greatest DDoS attack to date, targeted Amazon Web Services (AWS) and used a third-party server to create up to 2.3 Tbps of traffic. Based on the above, the negative impact of DDoS attacks leads companies to face significant revenue losses and additional operating costs due to recovery attempts for every hour the system is offline (Ahamed Aljuhani, 2021).

HTTP flooding DDoS attacks were among the most dangerous Internet threats that caused significant financial losses for commercial companies and government agencies. In recent years, HTTP flooding DDoS attacks have expanded from ordinary individuals to international organizations and businesses. For example, IBM reported the average financial losses due to DoS and DDoS attacks between 2016 and 2017, as shown in Figure 1.4.

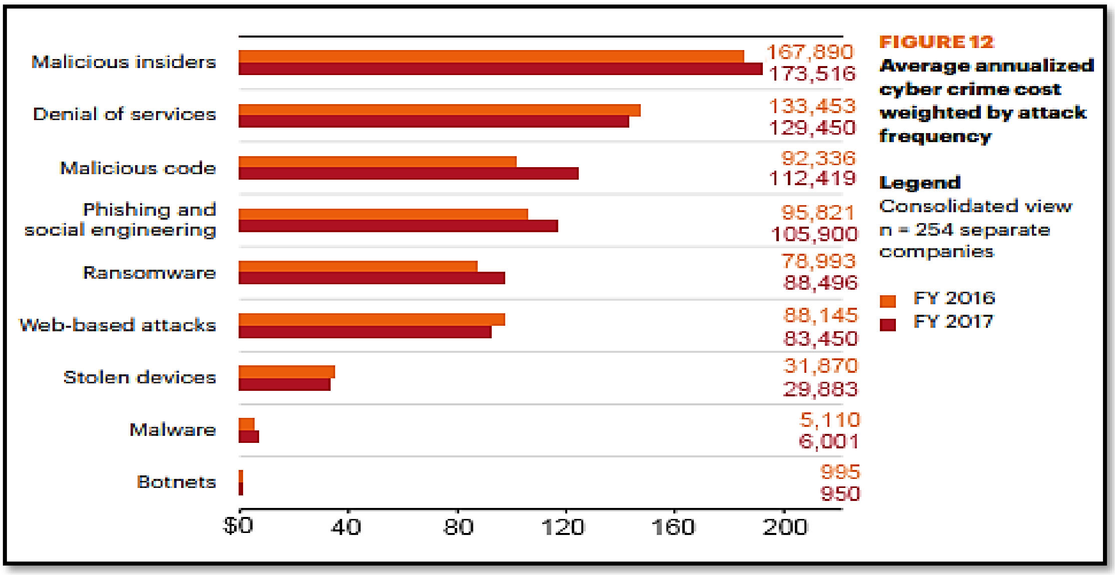

Figure 1.4 The diversity of DoS and DDoS attacks with the financial loses

Furthermore, Figure 1.5 shows that Cisco predicted the total number of DDoS attacks will double from 7.9 million in 2018 to slightly more than 15 million by 2023 (Paul Nicholson, 2020).

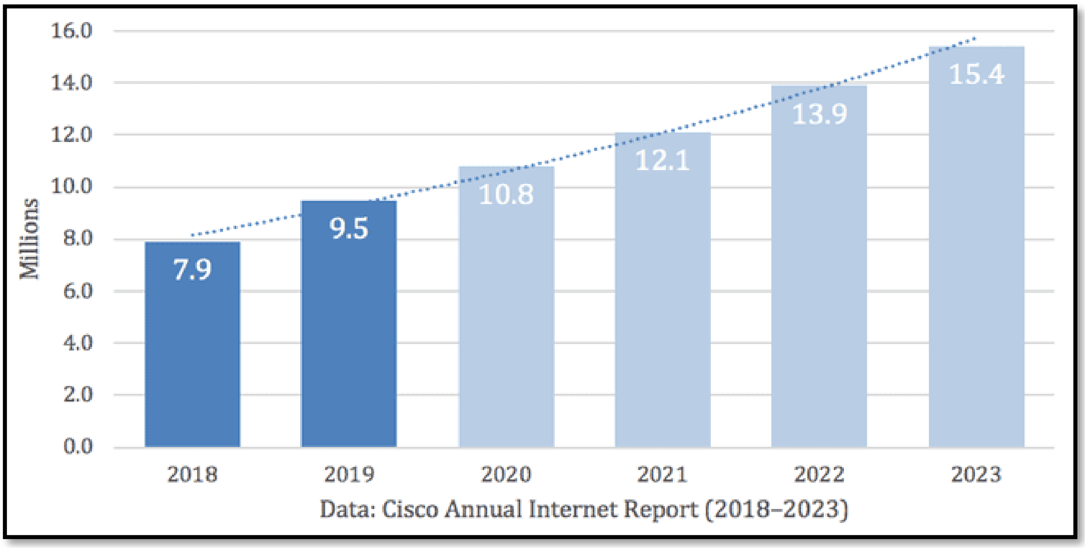

Figure 1.5 Cisco’s analysis of DDoS total attacks history and predictions

Cisco Visual Networking Index (VNI) projected the total number of DDoS attacks to double to 15.4 million by 2023, globally. In addition, Kaspersky reported that their products halted 975,491,360 browser-based attacks from around the world. The same report also indicates that 273,782,113 unique URLs were malicious (Aparna et al., 2021).

From the previous review of the reports of well-known organizations, it is clear to all that the dangers of HTTP flooding DDoS attacks causing painful financial losses, and since those attacks cannot be easily detected and distinguished from the normal network traffic, there is an urgent need to propose an effective mechanism to detect HTTP flooding DDoS attacks.

1.4 Research Problem

Financial losses occur due to HTTP GET flooding DDoS attacks flood web servers, systems or networks with traffic that overwhelms the victim's resources and makes it difficult or impossible for anyone else to access them. The presences of HTTP GET flooding DDoS attacks despite the high accuracy rates of the existing detection approaches indicate their ineffectiveness in countering these attacks.

The existing mechanisms for detecting HTTP GET flooding DDoS attacks against web servers are commonly grouped into (i) machine learning-based approaches, (ii) data mining approaches, and (iii) statistical approaches like (a) statistical-based mechanisms, (b) soft computing-based approaches, (c) knowledge-based approaches, and (d) clustering techniques. The mathematical-based approaches are one of the most effective approaches and are considered a base for other approaches (Gogoi et al., 2011).

Despite the existence of many statistical-based approaches to detect DDoS attacks, the problem is still prevalent. Many statistical-based approaches, such as (Jalili et al., 2005; Johnson Singh & De 2017; Liao et al., 2015; Lu et al., 2007) suffer from the high false-positive rate. The high false-positive rate is because these approaches use only limited mathematical functions, such as statistics, probability, and simple algebra, but disregarding the more advanced mathematical functions, such as calculus, and applied formulas. Advanced mathematical functions are widely employed in other fields, such as civil and mechanical engineering and show impressive results. However, advanced mathematical functions have not been used widely for detecting HTTP DDoS attacks. Additionally, advanced mathematical functions can better model the DDoS behaviours which reduce the false positive rates and increase the detection accuracy of DDoS.

Moreover, the reported result of high detection accuracy of most of the existing statistical-based approaches is doubtful, which could be because of miscalculation of the approaches' detection accuracy rate by partially using the metrics when calculating the detection accuracy. The calculations that employ all four metrics (True-Positive, True-Negative, False-Positives, and False-Negative) will achieve a genuine detection accuracy (Jain et al., 2019; Moayedi, Nguyen, & Rashid 2021).

Given the disadvantages, a detection approach that could detect HTTP flooding DDoS attacks on web servers based on the aggregated packets captured from the network traffic with better detection accuracy is needed. Thus, the problem statements of this research are summarized as follows:

1. The existing HTTP flooding DDoS attacks mechanisms that using the statistical approach cannot effectively model the DDoS behaviors which reduces the detection effectiveness depending on usage of simple mathematical functions, such as logical, statistical, or probability functions only with ignoring advanced mathematical functions, such as calculus or applied functions, which will increase the high false-positive.

2. The detection accuracy results of the existing statistical approaches either do not consider the full metrics of detection accuracy or do not justify why some metrics are left out. Thus, the reported detection accuracy results are in doubt, especially for future

researchers who want to benchmark their work with the existing statistical -based mechanisms.

1.5 Research Objectives and Goal

The main goal of this research is to propose a statistical-based approach for HTTP GET flooding DDoS attacks on web servers which is based on fourindicators (summation rows-columns, Bayes-entropy, skew of packets' distribution and Reynold's number, which is novel of this thesis) named as SB-ESR in the anomaly phase. This goal is further broken down into the following objectives to facilitate its accomplishment:

1. To propose a set of mathematical indicators that significantly contribute to detecting HTTP flooding DDoS attacks.

2. To adopt a voting-based mechanism working on the aggregated packets to detect the abnormality of HTTP flooding DDoS attacks according to the set of the mathematical indicators mentioned in the first objective.

3. To propose a statistical-based mechanism to evaluate the genuineness of the reported detection accuracy of the proposed approach using a set of qualitative and quantitative metrics.

1.6 Research Contributions

The main contribution of this research is a mathematical-based mechanism for detecting HTTP flooding DDoS attacks on web servers with a better detection accuracy rate. The contributions of the present research are as follows:

1. A set of mathematical indicators contribute to the detection of the abnormality of HTTP flooding DDoS attacks on a webserver; those indicators are (summation rows-columns; by checking the content of the suspicious packets, Bayes-entropy; by computing the probability and the randomness of attack packets in the traffic, skew of the distribution of the packet; by observing distribution shape of the equal packet size, and Reynold's number; by studying the proportionality of density, velocity, length and viscosity in each aggregation packet.

2. Adoption of voting-based mechanism to detect the presence of HTTP flooding DDoS attacks on a webserver. This mechanism detects the abnormal HTTP flooding DDoS using the four mathematical indicators mentioned above.

3. A statistical-based mechanism to evaluate the genuineness of the detection accuracy of the proposed approach and compare it with existing HTTP DDoS attack approaches by computing Spearman rank correlation between the quantitative and qualitative metrics. Table 1.1 shows the mapping between the problem's challenges, research objectives, and research contribution.

Table 1.1 Mapping between problem's challenges, research objectives, and research contribution

Problem's challenges	**Research objectives**	**Research contribution**
The existing statistical approaches use simple mathematical functions, which degrade the detection accuracy. In addition, advanced mathematical functions, such as calculus or applied functions, are not considered.	RO # 1 and RO # 2	RC # 1 and RC # 2
The detection accuracy results of the existing statistical approaches either do not consider the full accuracy parameters or justify why some parameters are neglected. Thus, the detection accuracy is in doubt, especially for future researchersto benchmark their work with the existing mathematical-based mechanisms.	RO # 3	RC #3
The existing statistical-based approaches rely on the packet representation of the network traffic, which leads to packet drops, especially in high-speed networks.	RO # 2	RC # 2

*RO = Research Objective RC= Research contribution

1.7 Research scope and limitations

The proposed approach is limited to detecting the HTTP flooding DDoS attacks on web servers. In addition, the proposed approach relies on the incoming aggregated packets captured from the network rather than the access log file of the targeted web server. Table 1.2 summarizes the scope and limitations of the proposed approach.

Table 1.2 Research scope and limitations

1	Attack type	HTTP GET Flooding DDoS attacks
2	Target	Application layer
3	Dataset	Benchmark dataset
4	Evaluation metrics	Detection accuracy and false-positive rate

1.8 Research Steps

The research is divided into four steps to achieve the goal of detecting HTTP flooding DDoS attacks in the incoming aggregated packets captured from traffic and to fulfill the objectives of this research (Section 1.4). The steps are (i) understanding the DDoS attacks nature and concept through reviewing the literature, (ii) proposing a set of mathematical functions related to the networking security science (called the qualitative metrics), building the proposed detection mechanism of the HTTP flooding DDoS attack based on the proposed mathematical set containing five phases (pre-processing, aggregated packets attributes, anomaly-based detection, voting-based mechanism and the statistical based mechanism); building four indicators (summation rows-columns, Bayes-entropy, skew of packets' distribution and Reynold's number) in the anomaly phase, (iii) designing and implementing them on two different datasets (CIC DDoS and ISCX); and finally, (iv) testing, comparing and evaluating the results). Figure 1.6 illustrates the research steps.

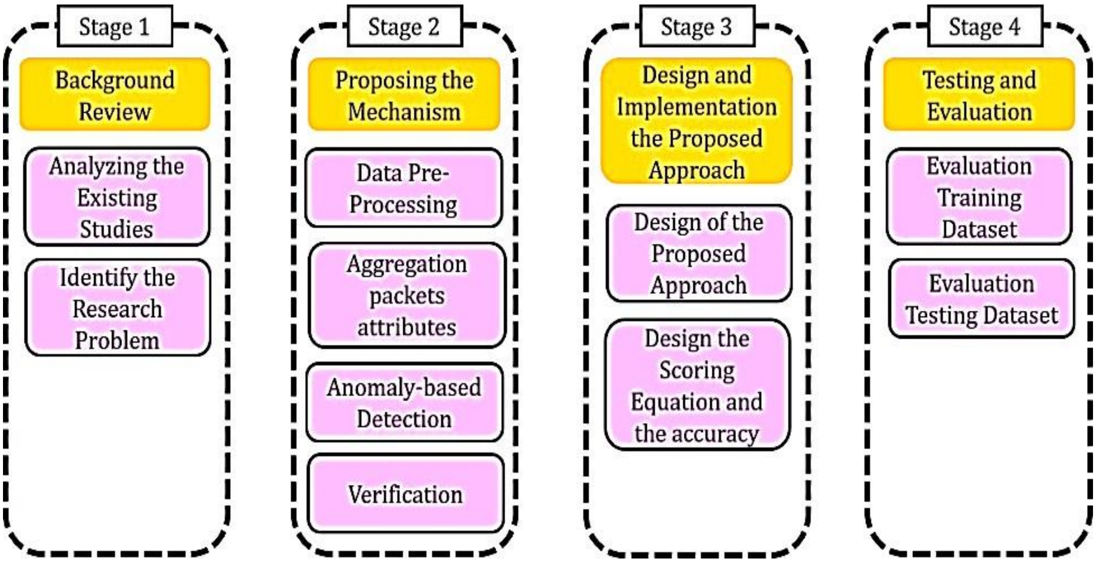
Stage 1
Background Review
Analyzing the Existing Studies
Identify the Research Problem
Stage 2
Proposing the Mechanism
Data Pre-Processing
Aggregation packets attributes
Anomaly-based Detection
Verification
Stage 3
Design and Implementation the Proposed Approach
Design of the Proposed Approach
Design the Scoring Equation and the accuracy
Stage 4
Testing and Evaluation
Evaluation Training Dataset
Evaluation Testing Dataset

Figure 1.4 Research Steps

In the first step, through reviewing the previous studies, understanding their works and background, the research problem is clarified, which explains the dimension of the problem, the existing solutions, and future scope to detect HTTP flooding DDoS attacks in network traffic.

The second step involved proposing the solution to the research problem. The proposed mechanism comprises five phases (pre-processing, aggregated packets attributes, anomaly-based detection, voting-based mechanism, and statistical-based mechanism) to detect HTTP flooding DDoS attacks through the enhancement of the detection accuracy. The anomaly-based detection phase comprises four indicators (summation rows-columns, Bayes-entropy, skew of the packet's distribution, and Reynold's number) applied to the aggregated packets captured from the traffic to determine whether the packets are normal, HTTP DDoS attacks, or suspicious.

The third step presented the proposed mechanism's design and implementation, using the four proposed algorithms, which represented the previous indicators in the proposed mechanism on the datasets.

The fourth step involves testing and evaluating the result to achieve the research objectives. Finally, the proposed mechanism will be tested and evaluated for its effectiveness in increasing the detection accuracy using a real dataset.

1.9 Thesis Organization

This thesis comprises the following chapters:

CHAPTER 2 will discuss the research background and related studies. This chapter critically reviewed the existing solutions for the detection of HTTP flooding DDoS attacks. Furthermore, this chapter comprehensively discussed the gaps in the research that exists. Finally, this chapter achieved the thesis's first objective by proposing the set of the mathematical functions related to the networking security science (the qualitative metrics) and joined it with the quantitative metrics by applying Spearman rank correlation.

CHAPTER 3 will explain the integrated phases of the proposed mechanism and the methods for detecting HTTP flooding DDoS attacks against web servers as a first step to achieve the second objective of this thesis.

CHAPTER 4 presents the design and implementation of the proposed mechanism. This chapter contains the design principle of the testbed and gives details of the dataset generation.

CHAPTER 5 will report the experiments and their results. It also presents a comprehensive analysis of the results achieved using the proposed mechanism. In addition, this chapter also evaluates the proposed mechanism's performance compared to existing mechanisms.

CHAPTER 6 presents the conclusion drawn from our work and suggests possible directions for future research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter presents a background of botnets, DoS and DDoS, HTTP Get flooding DDoS attacks, reviews the literature on the existing statistical -based detection approaches for detecting HTTP flooding attacks, and highlights the limitations of these approaches that serve as the motivation of this work.

This chapter is organized as follows. First, Section 2.2 provides the background. Then, section 2.3 covers related works, followed by Section 2.4 and Section 2.5 that present a critical review and statistical analysis of the related works, respectively. Finally, Section 2.6 summarizes the chapter. Figure 2.1 shows the structure of Chapter 2. Figure 2.1 shows the structure of Chapter 2.

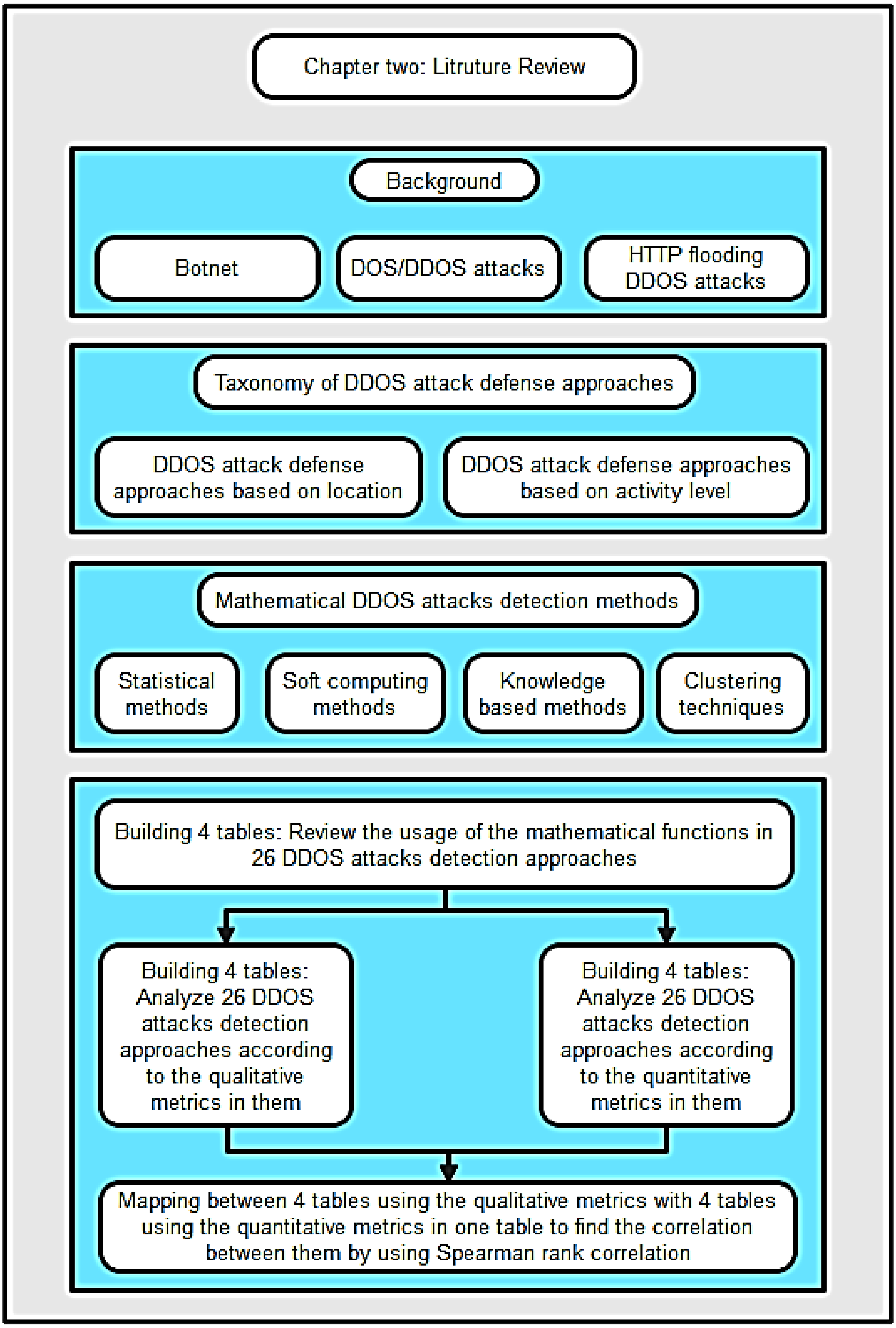
Chapter two: Litruture Review
Background
Botnet
DOS/DDOS attacks
HTTP flooding DDOS attacks
Taxonomy of DDOS attack defense approaches
DDOS attack defense approaches based on location
DDOS attack defense approaches based on activity level
Mathematical DDOS attacks detection methods
Statistical methods
Soft computing methods
Knowledge based methods
Clustering techniques
Building 4 tables: Review the usage of the mathematical functions in 26 DDOS attacks detection approaches
Building 4 tables: Analyze 26 DDOS attacks detection approaches according to the qualitative metrics in them
Building 4 tables: Analyze 26 DDOS attacks detection approaches according to the quantitative metrics in them
Mapping between 4 tables using the qualitative metrics with 4 tables using the quantitative metrics in one table to find the correlation between them by using Spearman rank correlation

Figure 2.1 Structure of chapter 2

2.2 Background

This section provides a comprehensive explanation of Botnet, DDoS, and HTTP Get flooding DDoS attacks in Section 2.2.1, Section 2.2.2, and Section 2.23, respectively.

2.2.1 Botnet

A botnet is a group of Internet-connected devices, each of which runs one or more bots. Botnets can be used to perform Distributed Denial-of-Service attacks, steal data, send spam, and allow the attacker to access the device and its connection. The owner can control the botnet using command and control software. The difference between botnets and other types of malware lies in the command and control (C&C) infrastructure that uses IRC, HTTP, or P2P as a C&C channel, enabling bots to execute malicious commands from the botmaster (Barford & Yegneswaran 2007; Thapngam et al., 2011).

2.2.2 Denial of Service/Distributed Denial of Service Attack

DoS/DDoS attacks continue to cause problems for the internet service providers as well as internet users. As a result, researchers have been trying to unravel the inner working of DoS/DDoS attacks since their inception.

DoS/DDoS attacks are malicious attempts to prevent legitimate users from accessing online services or resources by shutting down the server that provides the service or interrupting the network connection between the user and the server. DoS/DDoS attacks accomplish this by flooding the target with traffic or sending specially designed packets that cause failure. In both instances, the DoS/DDoS attacks deprive legitimate users (i.e., employees, subscribers, or account holders) of the online services or network resources (Yan et al., 2016).

DoS/DDoS attackers usually target the web servers of large organizations, such as banks, retail companies, media companies, government agencies, and commercial organizations. Although DoS attacks generally do not result in the theft or significant loss of information or other assets, they can cost victims a lot of time, money, reputation, and opportunities (Gupta & Badve 2017).

The difference between DoS and DDoS attacks is in the number of attack sources involved. A single attacker triggers a DoS attack. In contrast, multiple attackers trigger a DDoS attack; and by taking advantage of distributed attacks via botnets, a DDoS attack is several times more destructive than DoS.

Detecting an ongoing DDoS attack is challenging for several reasons (Jin & Yeung, 2004). First, the detection process is interactive and provides security professionals and system administrators with a short time to detect and verify ongoing attacks, leading to misclassification. Second, the standard network security preventive measures are limited (e.g., packet filtering, software parameters tweak, rate-limiting, etc.) despite their usefulness in preventing damages to vital network resources. Finally, the similarity of their network traffic makes it difficult to distinguish DDoS attacks from outbreaks. Therefore, there is a need for an effective and accurate mechanism to detect DDoS attacks in the network.

2.2.3 HTTP Get flooding DDoS attacks

HTTP is an application layer protocol exploited by malicious users for GET flooding attacks. In addition, these attacks usually support POST messages and find request ways used for submitting sensitive information to the Internet (Singh, Singh, & Kumar, 2017).

A low-rate HTTP flooding DDoS attack causes an online server to be unavailable. However, it is challenging to detect in the network because its traffic patterns are like legitimate users' traffic (Hong et al., 2018). Moreover, HTTP GET flooding DDoS attacks used standard HTTP protocol within the traffic (Idhammad, Afdel, & Belouch 2018; Lee & Lee 2011).

2.2.4 Methods of DDoS attack defence approaches

This section discusses the methods of DDoS attack defense approaches, categorized into two groups based on their activity level or location (Mirkovic & Reiher, 2004).

2.2.4(a) DDoS attack defense approaches based on activity level

This section provides the categorization of activity level-based DDoS attack defense approaches, as shown in Table 2.1.

Table 2.1 Categorization of activity level-based DDoS attack defence approaches (Mirkovice & Reiher, 2004)

DDoS defence by activity level	Preventive	Attack Prevention	• System security: allow victims to survive attack attempts without refusing services to legitimate users. • Protocol Security: tackle bad protocol design problems.
		DDoS Prevention	• Resource accounting: services police each user’s access to resources based on the user’s rights and actions. • Resource multiplication: offer many tools to combat DDoS attack threats.
	Reactive	Reactive by detection strategy	• Signature-based IDS: deploy pattern detection to store recognized attack signatures in a database. • Third-party: implementations do not manage the detection process itself. • Anomaly-based IDS: attacks detected based on abnormal traffic behaviour vs. attack signature. • Hybrid: combine pattern-based detection with anomaly-based detection.
		Reactive by response strategy	• Agent identification: provide the identity of the attackers’ machines to the victims. • Rate-limiting: impose a limit on traffic rate once detected as malicious. • Filtering: makes use of characterization generated by a detection mechanism to filter out the attack stream fully.
		Reactive by cooperation degree	• Autonomous: the mechanism conducts independent defence at the deployed location. • Cooperative: the mechanism is capable of autonomous detection and response but can achieve significantly better performance through cooperation with other entities. • Interdependent: the mechanism relies on other entities for attack prevention or detection because it cannot operate independently.

The activity level-based DDoS attack defense mechanisms have two different sub-categories: preventive and reactive. The preventive category has two groups: attack prevention and DDoS prevention, and the reactive category has three groups: detection strategy, response strategy, and cooperation degree.

2.2.4(b) DDoS attack defense approaches based on location

The second category of DoS/DDoS attack defense mechanisms is based on the mechanism's location in the network, as shown in Figure 2.2.

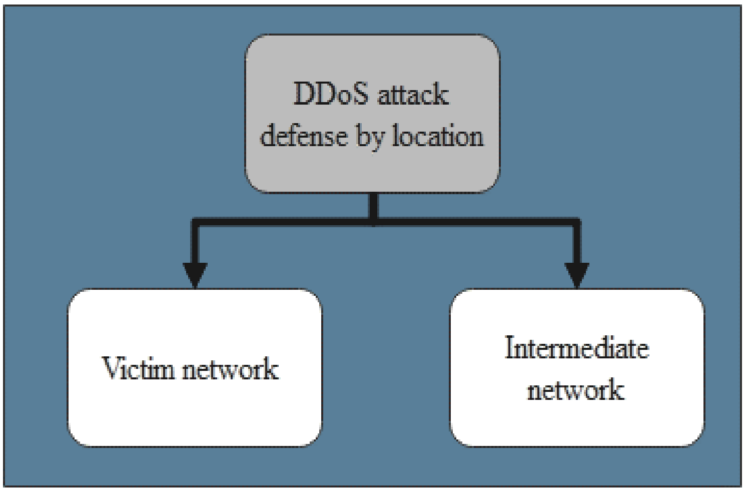

Figure 2.2 Categorization of location-based DDoS attack defence approaches

The two sub-categories for location-based DDoS attack defense mechanisms are victim network and intermediate network; victim network is for mechanisms that run on the victim network, whereas the second sub-category is for mechanisms deployed in the intermediate network (Mirkovic & Reiher, 2004).

2.2.5 Highlights on DDoS attacks

This section aims to highlight DDoS attacks and their nature by reviewing the efforts of previous researchers working in the DDoS attack field.

The early definition of DDoS defined it simply as a tool used by attackers (Sahoo et al., 2018). However, another definition is an attack launched by a botnet through a network of controlled computers (Yan et al., 2016).

Some researchers categorized the types of DDoS attacks according to their intended target. There are four categories of targets—first, service infrastructures (DNS server, email server, and webserver). Second, network infrastructures (routers, switches, etc.). Third, end-user to reach service provider infrastructure (spams attack customer to migrate to SPI) as (end-users/subscribers, e-commerce, financial service, governments, gaming, gambling, utility manufacturing, others). Lastly, the fourth target is protocols using UDP/53 or TCP/80 ports in layer 4 (Elejla, Anbar, & Belaton 2017; Sahoo et al., 2018).

Other researchers defined the first goal of DDoS attacks as depriving authorized clients of benefits from online services. The second goal of DDoS attacks is to disrupt the flow of data in the network by generating enormous traffic between the systems, eventually denying users of services or degrading the quality of services (Yan et al., 2016).

Some researchers also categorized DDoS attacks based on the perpetrator's plausible motives, including renegade, enmity, hate, competitions, ransom, and political affairs (Abhishta et al., 2020).

And they bordered the DDoS attacks as being the first border (establishing a network of computers) to generate a massive volume of traffic needed to deny services to legitimate users or victims. The second border is (discovering vulnerable hosts on the network), and the third border is (master computer orders the zombies (bots) to run the attack tools to send a massive volume of data to the victim) (Gupta & Badve, 2017).