ANOMALY-BASED DETECTION APPROACH TO DETECT THE FLASH CROWD ATTACK DURING THE FLASH EVENT

SAMER ABDULSADA MUTLAG AL-SALEEM

UNIVERSITI SAINS MALAYSIA

2017

ANOMALY-BASED DETECTION APPROACH TO DETECT THE FLASH CROWD ATTACK DURING THE FLASH EVENT

by

SAMER ABDULSADA MUTLAG AL-SALEEM

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

December 2017

DEDICATION

To my appreciated father "Dr. Abdulsada Mutlag Al-Saleem."

To my dearest mother "Dr. Noria Flayyh Al-Joboori."

To my beloved wife "Marwah Abdulmonem Jameel."

To my lovely son "Mohammed Rayyan."

To my dearest sister "Samarah."

To the memory of my beloved Iraq

ACKNOWLEDGEMENT

بسم الله الرحمن الرحيم

{نَرْفَعُ دَرَجَاتٍ مَّن نَّشَاء وَفَوْقَ كُلِّ ذِي عِلْمٍ عَلِيمٌ} {76- سورة يوسف}

All praise and thanks are due to ALLAH SUBHANAH WA TAALA, the Lord of the world, for giving me the health, strength, knowledge and patience to complete this work.

Since the Prophet MOHAMMED "Peace be Upon Him" said: 'Whoever does not thank people (for their favours) has not thanked Allah (properly)', therefore, First, I would like to express my deepest gratitude to my supervisor, **Dr. Selvakumar Manickam** for giving me full support and faithfulness in all guidance and commitments upon on effort from the early stages of this study through to the completion of this thesis. His wide knowledge and understanding have been invaluable to me especially in giving constructive comments and advice throughout this study. Furthermore, my appreciation and sincere gratitude go to the co-supervisor **Dr. Mohammed Anbar** for his diversified help, support and the encouragement along with his contribute to complete the thesis.

I would like to express my gratitude and thanks to all the **academic staffs** in National Advanced IPv6 Centre (NAv6), Universiti Sains Malaysia for their dedication and persistent support. Besides that, the **administration and support staffs** inNAv6 deserve a great mention for keeping everything running smoothly.

My sincere, heartfelt gratitude to my parents, **Dr. Abdulsada Mutlag** and **Dr. Nooriya Flayyih**, Sister **Samara Abdulsada**, my wonderful wife **Marwah Al-Mufti** and my lovely son, **Mohammed Rayyan** for their endless love, help, and encouragement during this study period especially when I have endured a period of frustration. Word cannot truly express how much I owe you all.

Last but not least, collective and individual acknowledgments are also owed to my friends and colleagues in NAv6 who have helped and supported me over these years, especially **Dr. Esraa Saleh Hasoon** for assisting me in collecting my data to verify the work and the results and being there whenever I need.

Samer Abdulsada Mutlag Al-Saleem
Penang, Malaysia, February 2017

TABLE OF CONTENTS

CHAPTER 2 LITERATURE REVIEW

CHAPTER 3 METHODOLOGY OF THE PROPOSED APPROACH

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACO	Ant Colony Optimization
ADM	Angular Directional Model
AODV	Ad hoc On-demand Distance Vector
ARIB	Association of Radio Industries and Businesses
ASK	Amplitude Shift Keying
ASTAR	Anchor-based Street and Traffic Aware Routing
ASTM	American Society for Testing and Materials
AWGN	Additive While Gaussian Noise
BER	Bit Error Rate
BIT	Beacon Interval Time
B-MFR	Border-node based MFR
BPS	Bit per Symbol
BPSK	Binary Phase-Shift Keying
BSS	Basic Service Set
C2C-CC	Car-to-Car-Communication Consortium
CALM	Communications Access for Land Mobiles
CAR	Connectivity Aware Routing
CDF	Cumulative Density Function
CEN	European Committee for Standardization
CLWPR	Cross-Layer Weighted, Position-based Routing protocol
CMGR	Connectivity-Aware Minimum-Delay Geographic Routing
CR	Compass Routing
CTR	Communications Transmission Range

PENDEKATAN PENGESANAN ANOMALI UNTUK MENGESAN SERANGAN FLASH CROWD SEMASA PERISTIWA KILAT

ABSTRAK

Sepanjang dekad yang lalu, ancaman keselamatan yang paling mengganggu di Internet adalah *Distributed Denial of Services* (DDoS). Pada masa ini, di antara semua jenis serangan DDoS pada lapisan aplikasi, serangan *flash crowd* yang mensasarkan pelayan web semasa peristiwa kilat adalah dianggap paling mencabar; oleh itu, melindungi pelayan web daripada jenis serangan tersebut telah menjadi masalah kritikal yang sangat perlu ditangani. Matlamat tesis ini adalah untuk mencadangkan satu pendekatan yang dapat mengesan serangan *flash crowd* yang menyasarkan pelayan web semasa peristiwa kilat. Pendekatan yang dicadangkan terdiri daripada empat fasa untuk mencapai matlamat penyelidikan. Pendekatan yang dicadangkan telah dinilai menggunakan tiga set data yang tersebar di tiga senario. Senario pertama adalah bertujuan untuk mengesahkan ketepatan pengesanan pendekatan yang dicadangkan apabila pelayan web berada dalam keadaan biasa (ketika tidak ada ancaman yang mensasarkan pelayan web). Senario kedua (set data peristiwa kilat) bertujuan untuk meningkatkan ketepatan pendekatan yang dicadangkan dalam mengesan tingkah laku anomali apabila pelayan web menghadapi peristiwa kilat. Senario ketiga (set data serangan *flash crowd*) bertujuan untuk mengesahkan ketepatan pendekatan yang dicadangkan dari segi mengesan serangan *flash crowd* yang mensasarkan pelayan web semasa peristiwa kilat. Kajian ini mempertimbangkan keperluan untuk mengesan tingkah laku anomali semasa serangan di mana ia boleh digunakan untuk mengesan permintaan yang berniat jahat dan meningkatkan keselamatan rangkaian. Sumbangan utama penyelidikan ini adalah untuk mempromosikan pendekatan yang dapat mengesan permintaan serangan *flash*

crowd semasa peristiwa kilat. Hasil dan penilaiannya menunjukkan dengan jelas bahawa pendekatan yang dicadangkan dapat mengesan serangan *flash crowd* semasa peristiwa kilat dengan ketepatan 98% dari segi mengesan tingkahlaku anomali dan 100% dari segi mengesan serangan kilat orang ramai.*flash crowd.*

ANOMALY-BASED DETECTION APPROACH TO DETECT THE FLASH CROWD ATTACK DURING THE FLASH EVENT

ABSTRACT

Over the last decade, the most intrusive security threat on the Internet is the Distributed Denial of Services (DDoS). Currently, among all types of application-layer DDoS attacks, the flash crowd attack that targets a web server during the flash event is considered the most challenging; therefore, protecting web servers from such type of attacks has become a critical problem that urgently needs to be addressed. The goal of the thesis is to propose an approach that can detect flash crowd attacks that target web servers during flash events. The proposed approach consists of four phases to achieve the goal of the research. The proposed approach is evaluated using three different datasets distributed in three scenarios. The first scenario is aimed to validate the detection accuracy of the proposed approach when a web server is in a normal situation (when there is no threat targeting the server). The second scenario (flash event dataset) is aimed to improve the accuracy of the proposed approach in detecting anomalous behaviour when the server is facing a flash event. The third scenario (flash crowd attack dataset) is aimed to verify and validate the accuracy of the proposed approach in terms of detecting the flash crowd attack that targets the web server during the flash event. The main contribution of this research is to promote an approach that responds to detect flash crowd attack requests during the flash event. The result and its evaluation clearly demonstrate that the proposed approach can detect the flash crowd attack during the flash event with an accuracy of 98% in terms of detecting the anomalous behaviour and 100% in terms of detecting the flash crowd attack.

CHAPTER ONE

INTRODUCTION

1.1 Overview

This thesis discusses the design of a framework for the detection of flash crowd attacks that occur against web servers during flash events. In this chapter, an introduction to Internet security is presented with discussions focusing on malware, botnet, distributed denial of service (DDoS) attacks, and flash events and flash crowd attacks in Sections 1.2.1, 1.2.2, 1.2.3, and 1.2.4, respectively. Sections 1.3, 1.4, and 1.5 present the research motivation, problem, and objectives, respectively. The scope of this thesis is stated in Section 1.6, and the research contributions of this study are presented in Section 1.7. Section 1.8 presents the research methodology. The organisation of this thesis is summarised in Section 1.9.

1.2 Internet Security Issues

In the past decade, the Internet has become the popular way to provide information and services to users dynamically (Angrishi, 2017). Figure 1.1 shows the number of hosts interconnected through the Internet over 10 years, indicating the increasing influence of the Internet on society.

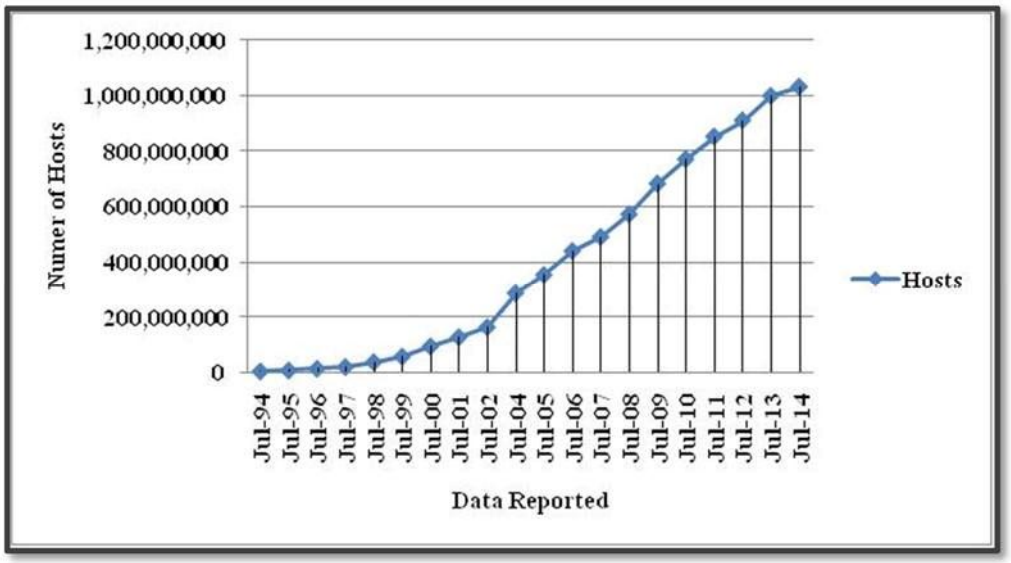

Figure 1.1: Number of Hosts Connected to the Internet

One of the major and widely used services after the email is the World Wide Web (WWW). The first-generation web was severely limited in its ability to provide any more information than a brochure one might receive in the mail. The rapid growth of the web can be attributed to the changes in traditional roles and in the way business is conducted using the web, allowing all transactions through the Internet (Berners-Lee & Cailliau, 1990). For example, the government uses the Internet to provide its citizens many information and governmental services. Furthermore, the web enables companies to share and exchange information among their divisions, suppliers, partners, and customers to increase operational efficiency. Research and educational institutions depend on the Internet as a medium for collaboration to enhance their research discoveries.

The Internet is an international collection of independent networks owned and operated by many organisations, and no central authority exists through which it can regulate the behaviour of its users. Therefore, network attacks have become more sophisticated by shifting from physical attacks (direct sabotage of digital resources)

to remote attacks (disruption or disabling of one or more targets, such as web servers).

According to statistics (TEAM, 2015), only 171 vulnerabilities were reported in 1995. The number of vulnerabilities increased to 7,236 by 2007. This number further increased to reach over 10,000 in 2013 and more than 15,000 in 2014**,** as shown in Figure 1.2.

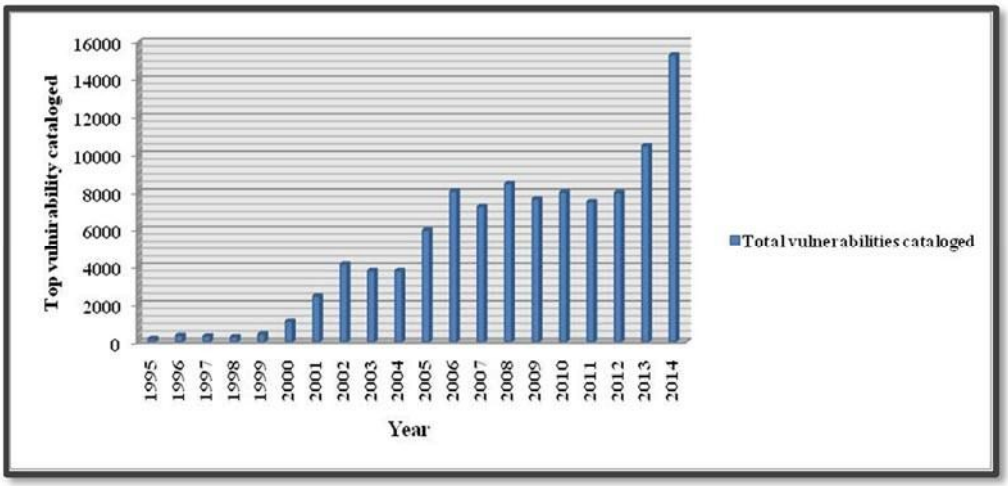

Figure 1.2: Number of Vulnerabilities Reported Over 20 Years

Securing web applications has become incredibly important as the information processed by web applications has become critical to corporations, customers, organisations, and countries. Web-based attacks are considered by security experts to be the greatest and oftentimes the least understood of all risks related to confidentiality, availability, and integrity (Ravikiran Kalava, 2012). The purpose of a web-based attack is significantly different from other attacks. In more traditional penetration testing exercises, a network or host is the target of attack. Web-based attacks focus on an application and function on Layer 7 of the Open System Interconnection (OSI) model. John Pescatore of the Gartner group claims that nearly 70% of all attacks occur at the application layer (Rights, 2001). All web application attacks are comprised of at least one normal request or a modified request aiming at

taking advantage of poor parameter checking or instruction spoofing. Application attacks have five fundamental categories (Rights, 2001):

1. **Spoofing:** Spoofing is the act of mimicking another user or process to perform a task or retrieve information that would normally not be allowed. An attacker could use a crafted HTTP request containing the session ID information of another user and retrieve the targeted user's account information (Persis, Nazareth, Dharmaraj, & Neil, 2014).
2. **Repudiation:** To tie specific actions of a single user, applications must have reasonable repudiation controls such as web access, authentication, and database transaction logs. Without corroborating logs, online web application users could easily claim that they did not transfer equities from one account to an external account of another. Without proof, all online brokerages would be required to reimburse the client for their lost funds. Aggregating and correlating logs from multiple sources (web application, middleware, and database) can prevent repudiation attacks (B. Wu, Chen, Wu, & Cardei, 2007).
3. **Information Disclosure:** Information disclosure is one of the biggest threats to large organisations that maintain private information about their customer base. When attackers can reveal private information about a user or users of a website, consumer confidence in that organisation can take drastic hits, thereby causing loss in sales, stock price, and overall marketability. To prevent this, applications require adequate controls to prevent the manipulation of user IDs and sessions (Chester & Srivastava, 2011).
4. **Evaluation of Privileges:** Authorisation controls, which are both reliable and staunch, are requisite for any system or application that guards sensitive

information. Escalation of privileges requires a malicious user to either already possess or gain through unlawful methods the authorisation privileges of a regular user. Once the malicious user is logged into the victim system, an attempt will be made to exploit an application through poor parameter checking or instruction spoofing. (Crist, 2007).

5. **Denial of Service:** DoS attacks are likely the most well-known of all application attacks. They are often generated by malicious users, competitors, or script kiddies. Motivations for this type of an attack range from personal to political reasons in hopes of stifling an organisation's ability (Rights, 2001).

Attacks launched over the web can be carried out from anywhere in the world. Unfortunately, no web-based application service is immune to these attacks. Therefore, the reliability and security of web applications are issues that affect not only online businesses, but also the national security. The new paradigm of these attacks, known as "malware," has become one of the most insidious threats in the world.

1.2.1 Malware

Over the last decade, malware has been the most intrusive security threat on the Internet and has risen to become a primary source for most of the scanning DDoS activities (Zhou & Jiang, 2012).

Malware is malicious software used by cybercriminals and hacktivists to disrupt computer operations, steal personal or professional data, bypass access controls, and cause harm to the host system. Malwares are of many different classes and possess varying means of infecting machines and propagating themselves.

Some of the most visible and serious problems facing the Internet today depend on malicious software and tools. Spamming, phishing, DoS attacks, botnets, and worms largely depend on some form of malicious code, which is commonly referred to as malware (Rudd, Rozsa, Gunther, & Boult, 2016). Malware is often used to infect the computers of unsuspecting victims by exploiting software vulnerabilities or tricking users into running malicious codes. Among the various forms of malicious software, botnet has been recently distinguished to be among the primary threats to computing assets. Like the previous generations of computer viruses and worms, bot is a self-propagating application that infects vulnerable hosts through direct exploitation or Trojan insertion. A detailed discussion of botnet is presented in the next section (C. Li, Jiang, & Zou, 2009).

1.2.2 Botnet

One form of the malware application is known as a bot. It is generated to automatically perform a specific operation, with the infected machine often being called a "zombie". (Arshad, Abbaspour, Kharrazi, & Sanatkar, 2011). Defrauding users into making drive-by downloads, exploiting web browser vulnerabilities, or tricking users into running Trojan are ways to help cybercriminals execute the malicious software needed to recruit a computer into a bot (Soltani, Seno, Nezhadkamali, & Budiarto, 2014).

Multiple bots that communicate with each other are called a "botnet". It can help execute different types of attacks, such as DoS. The aim of a botnet is to control many computers, which is accomplished by installing a backdoor in each of them. The individual computers in the botnet then become zombies because they are under

remote control, but are usually referred to as bots. Bots can be given orders by the controller, which is often known as the botmaster, to perform various tasks, such as sending spam mails, adware, or spyware; collecting confidential information, such as passwords or encryption keys; performing DDoS attacks; or just searching for further potential targets to be recruited into the botnet (Fabian & Terzis, 2007; Fuchs & Brunner, 2013). Figure 1.3 shows how the botmaster controls the botnet to perform an attack.

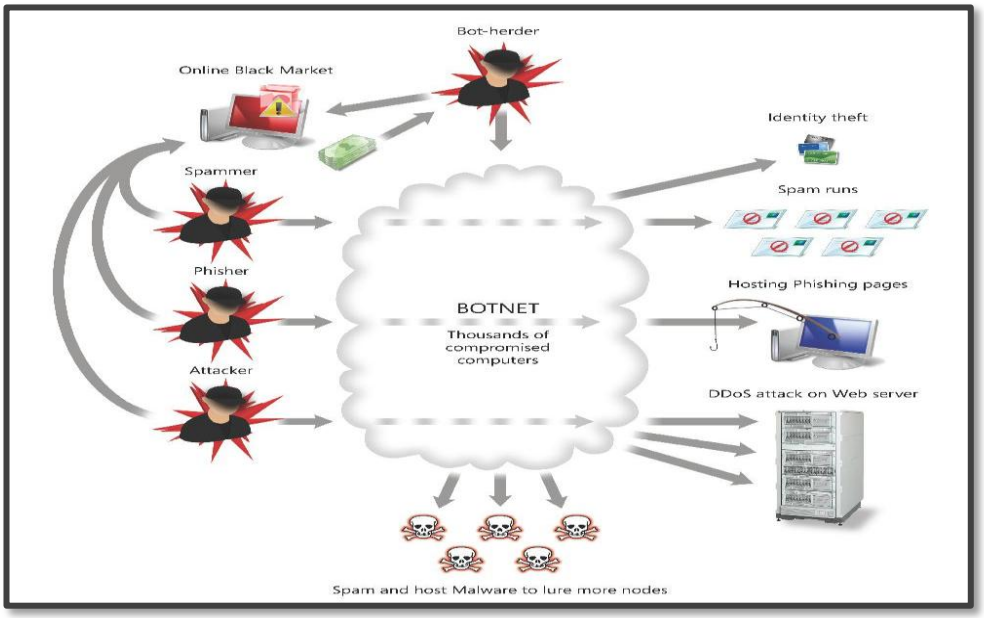

Figure 1.1: Botnet Scenario

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

DoS attacks are one of the oldest types of botnet activities, which pose a serious and permanent threat to users, organisations, and the infrastructure of the Internet. This type of attack is characterised by an intentional attempt by malicious users/attackers to completely disrupt or degrade the availability of services or resources to legitimate users (Feily, Shahrestani, & Ramadass, 2009).

According to the International Telecommunication Union (ITU) recommendation X.800, DoS is the "*prevention of authorized access to resources or delaying of time-critical operation".* (Telecom, 1996). The increasing velocity of such attacks has increased the risk of servers and network devices on the Internet. In February 2000, the first documented DoS-style attacks were launched when a 15-year old hacker started a series of attacks against e-commerce sites, such as Amazon.com and eBay.com (Calce & Silverman, 2008).

After DoS attacks are launched, the attacker eventually becomes aware of the defence mechanisms that are implemented to prevent and mitigate DoS attacks. To overcome the downfalls of aggregate DoS attacks, the attacker launches a new type of DoS attack, which uses distributed traffic to attack victims (DDoS). A DDoS attack is a multiple form of DoS attack. It is a large-scale coordinated attack on the availability of Internet services and resources. It uses the same technique as DoS but on a much larger scale and from more than one source and/or more than one location at the same time (Mirkovic, Dietrich, Dittrich, & Reiher, 2004). Figure 1.4 shows the general scenario of a DDoS attack.

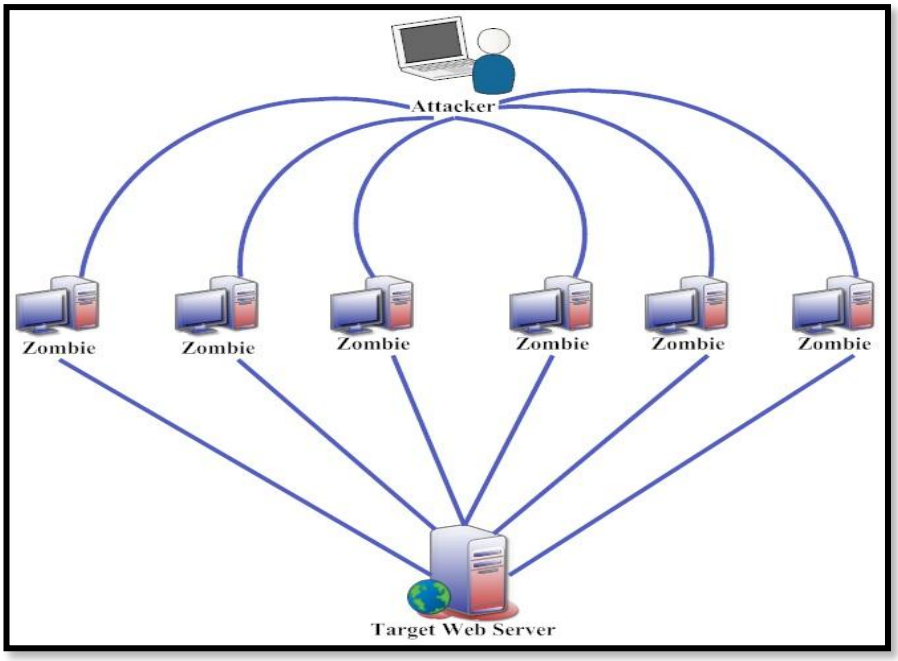

Figure 1.2: DDoS Attack

The DDoS attacks can be categorised into two classes as following:

- **Network-layer DDoS attacks:** The attacker in network DDoS attacks attempts to disrupt legitimate users' connectivity by exhausting the victim network's bandwidth or by exploiting a specific feature in the victim protocol. Sometimes, the attacker, who typically uses IP spoofing, sends a large number of bogus packets to the victimised server. To establish this type of attacks, the attacker mostly uses a TCP, UDP, ICMP, or DNS protocol packet (S. Lee, Kim, & Kim, 2011).
- **Application-layer DDoS attacks:** The attacker who targets the application layer attempts to exhaust the server resources, such as the CPU, memory, and disk space. In general, this type of attacks has the same impact to the services (disrupting the connectivity) because they target specific characteristics of applications, such as HTTP and DNS (Suryawanshi & Todmal, 2015).

HTTP-flooding attacks are well-known botnet type of application-layer DDoS attacks that cause severe damage to servers and even greater disruption to the development of newer Internet services. The situation becomes more serious when attackers try to mimic the flooding of legitimate user requests, which is known as flash crowds.

1.2.4 Flash Event and Flash Crowds Attacks

Flash crowd refers to a huge number of people who assemble in one place for the same reason for a brief time (Sachdeva & Kumar, 2014). In computer networks, flash crowd is a rush in traffic to a specific website over a comparatively short period. It happens due to special events, such as breaking news and release of a popular product. A flash event sometimes occurs when a popular website links to a smaller site, thereby causing a massive increase in traffic, which is also known as a flashdot effect (Pai, Druschel, & Zwaenepoel, 2012).

Both flash events and the flashdot affect the network infrastructure and the server operation of websites because overcrowding at the network layer can prevent some user requests from reaching the server. The requests may reach the server after a considerable delay caused by packet loss and resend requests. Certain web server configurations and descriptions cannot handle the volume of the flash event requests. In the end, users who try to reach the website in a flash event will be disappointed due to the long wait or failure to reach the target.

Flash crowd attacks attempt to make Internet resources and services unavailable to its intended users. A very common method of flash crowd attack involves saturating the victim machine with external communication requests so that

it cannot respond to legitimate traffic. Moreover, flash crowd attacks attempt to do so by sending these external requests from many compromised machines (zombies, daemons, agents, slaves, etc.) distributed around the world. These legitimate-looking requests bring down the victim server by consuming scarce resources, for example, the CPU cycles, memory, and bandwidth of the victim machine or network (Bhatia, Mohay, Tickle, & Ahmed, 2011). Table 1.1 lists the differences between the flash event and the flash crowd attacks.

Table 1.1: the difference between flash event and flash crowd

CHARACTERISTIC	FLASH CROWD	FLASH CROWD ATTACK
Traffic volume	Both have a noticeable increase in the number of requests.	
Number of clients and their distribution	Caused mostly by an increase in the number of clients accessing the site. Client distribution can be expected to the follow population distribution among ISPs and network.	Caused either by an increase in the number of clients or a particular client sending requests at a high rate. Client distribution across ISPs and networks does not follow population distribution.
Cluster overlap	Significant overlap between clusters a site sees before and during flash events.	Cluster overlap is small.
Per-client request rates	Because a server usually becomes slower during the flash event, per-client request rates are lower during the flash event than usual. This indicates that legitimate clients are responsive to the performance of the server unlike flash crowd attackers who generate requests by pre-determined time distribution.	Some flash crowd attacks involve a few clients emitting very high request rates and a large number of clients generating a low request rate, but in both cases, the per-client request rate is stable during the attack and significantly deviates from normal.

1.3 Research Motivation

More than ever, the dependency on web technology is increasing. Meanwhile, destructive attempts to disrupt web users are also increasing. Flash events and flash crowd attacks are among the most dangerous Internet threats. Like flash events, flash crowd attacks can also have a significant financial implication. An example of a recent flash event that caused a substantial financial loss occurred on the 20 November 2012, when Click Frenzy, a national online shopping initiative in Australia, which is similar to the US-based Cyber Monday event, was launched after a heavy media and online publicity (Alsaleem, Manickam, Anbar, Alnajjar, & Saleh, 2017). The website experienced a traffic volume many times that of its anticipated traffic volume, leading to a dramatic increase in page-load times and the failure of the website within minutes after its launch. The organisers of the website had pre-arranged sales partnerships with leading Australian and international retailers and brands, many of whom had paid large amounts for advertisement and asked for a refund (Bhatia, 2013).

Flash crowd attacks and flash events can both overload the server or the server's Internet connection and result in partial or complete failure. Unlike flash crowd attacks, which are simply malicious requests that do not have to be handled by a website, flash events consist of legitimate requests (Thapngam, Yu, Zhou, & Beliakov, 2011). The web server has the responsibility to try and handle as many requests as possible during a flash event. By doing so, the site may increase its overall profile on the web, thereby resulting in additional revenue. If a flash crowd attack occurs during a flash event, the web server should ignore flash crowd attack requests and handle the legitimate requests. This requires the website to be able to distinguish the two sets of requests (Jung, Krishnamurthy, & Rabinovich, 2002).

1.4 Research Problem

The main challenge for the web server is to differentiate between a flash event request (from legitimate users) and a flash crowd attack request during a flash event because both have the same symptoms: (i) delayed response to legitimate users or (ii) complete crash of the web server (Jung et al., 2002). The existing techniques cannot differentiate between the flash event and flash crowd attack requests, thereby leading to low accuracy in terms of detecting flash crowd attacks. This limited capability can be attributed to the following reasons (J. Yu, Li, Chen, & Chen, 2007):

1. The existing techniques for detecting flash crowd attacks do not consider all the features in a web access log file that can contribute in detecting flash crowd attacks, such as the timeslot monitoring model which will be explained in Chapter 2.
2. The majority of the existing flash crowd attack detection methods that rely on analysing user requests in the web access log suffer from significant false positive and false negative requests because of the shortcomings in existing machine learning techniques.

1.5 Research Objectives and Goal

The main goal of this thesis is to propose an approach for flash crowd attack detection during flash events. The following objectives are set to achieve this goal:

1. To identify the most contributed features in the detection of flash crowd attacks by conducting an experiment using the web access log and verifying the result by employing two different algorithms.
2. To propose a rule-based mechanism to detect the misbehaviour of flash crowd attacks.

3. To verify and validate the effectiveness of the proposed approach in terms of detection accuracy and compare it with existing approaches.

1.6 Research Contributions

The main contribution of this research is the proposed rule-based approach which is designed to detect flash crowd attacks during flash events at the application layer with a better detection accuracy rate. The contributions of the present research are as follows:

1. An approach for the detection of flash crowd attacks against the Apache web server with a better detection accuracy rate.
2. A hybrid approach to extract the most effective features that most contribute to the detection of flash crowd attacks during flash events.
3. A rule-based mechanism to detect the anomalous behaviour of a flash crowd attack.

1.7 Research Scope

The proposed approach is motivated by the detection of flash crowd attacks against web servers at the application layer during a flash event. However, we are currently focused on the detection of flash crowd attack against Apache web servers. The rationale for focusing on Apache web servers is that nearly **70%** of all web servers globally that are subject to such attacks target the Apache server, which runs predominantly on Linux. In addition, the proposed approach retrieves the common web access log file from the web server as an input data for use in detection. Furthermore, the proposed approach does not address the flash crowd attacks at the network layer. Table 1.1 summarises the scope and limitation of the research.

Table 1.1: research Scope and limitation

Items	Scope of Research
Attacks type	Flash Crowd Attack
Target	Application Layer
Period of attack	During the Flash Event
Server type	Apache web-server
Log file type	Common web access log
Request type	GET Request
Detection	Anomaly- based detection

1.8 Research Methodology

To achieve the goal of detecting flash crowd attacks during flash events and to fulfil the objective of this research as stated in Section 1.4, the research process is divided into four phases: (i) understanding the flash event and flash crowd concepts through a review of the literature, (ii) proposing a new approach to detect flash crowd attacks during flash events, (iii) designing and implementing the proposed approach, and (iv) testing and evaluating the result. Figure 1.5 illustrates the research phases.

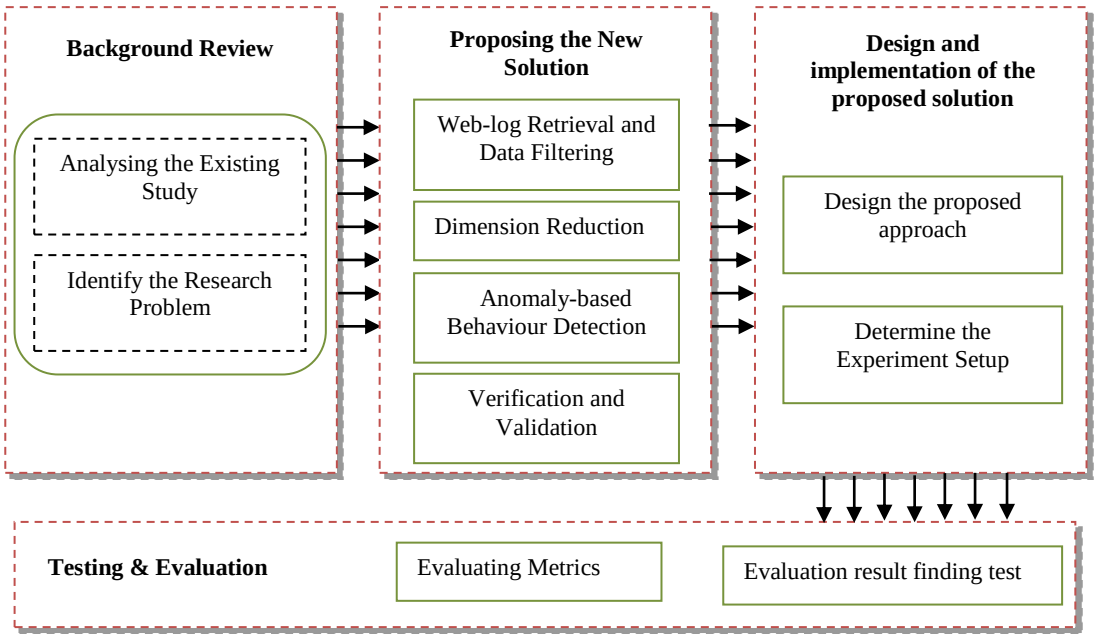

Figure 1.3: Step of Research Methodology

In the first phase, the research problem is clarified through the review of existing studies and understanding the background. This elucidates the dimension of the problem, the existing solution, and future scope to detect flash crowd attacks during flash events.

In the second phase, the solution to the research problem is proposed. The solution consists of four steps to detect flash crowd attacks by improving the detection accuracy. The proposed approach employs a rule-based mechanism to detect flash crowd attacks during flash events.

The third phase shows the design and the implementation of the proposed solution which uses the most contributed fields from the dataset to improve the efficiency in term of feature (field) selection, model training, and anomaly detection.

The fourth phase is mainly concerned with testing and evaluating the result to achieve the research objective. The proposed approach is tested and evaluated based on its effectiveness in increasing the detection accuracy using a real dataset. The approach is compared with existing behaviour-based detection approaches of application-layer DDoS attacks during flash events.

1.9 Thesis Organisation

This thesis comprises six chapters:

CHAPTER 1 presents the motivation, scope, objectives, and contributions of this work. This chapter also discusses the need for server-side protection strategies for the detection of flash crowd attacks.

CHAPTER 2 discusses the research background and related studies. This chapter critically reviews the existing solutions for the detection of flash crowd attacks. Furthermore, this chapter comprehensively discusses the gaps in the research that has been performed.

CHAPTER 3 explains the integrated phases of the proposed framework as well as the methods adopted for the detection of flash crowd attacks against web servers.

CHAPTER 4 presents the design and implementation of the proposed framework. This chapter contains the design principles of the test bed and presents details regarding the dataset generation. This chapter also explains the implementation of the phases in detail.

CHAPTER 5 reports the experiments and their results. It also presents a comprehensive analysis of the results achieved using the proposed framework. In addition, this chapter evaluates the performance of the proposed framework in comparison with existing schemes.

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

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

The importance of addressing the problem of flash crowd attacks was described in Chapter 1. In this chapter, the background knowledge required for better understanding the problem is provided. The literature review in this thesis is organised into eight sections. The underlying concepts of the flash event and flash crowd attack are described in Section 2.2 and Section 2.3, respectively. Section 2.4 presents the classification of flash crowd attack detection approaches, and Section 2.5 explains the supervised and unsupervised machine learning methods adopted for detecting the flash crowd attacks. Section 2.6 presents the related studies conducted in this area. Finally, Section 2.7 summarises the chapter, justifying the need for an efficient approach to detect the flash crowd attacks against web servers during flash events.

2.2 Underlying Concept of Flash Event

Over the past decade, advances in Information and Communication Technology (ICT) have significantly transformed the way in which information is accessed and communicated, particularly via the Web (K. R. Lee, 2002). The range of services supported by ICT have constantly been expanding and in recent years have even included the control and monitoring of key systems such as power, water, gas, etc., also known as Critical Infrastructure (CI). This evolution of ICT has also entailed a significant dependence of society on the systems used for storing, processing, and communicating information (Unesco, 1996). As a result, any malfunction in these

information and communication systems directly affect, in one way or another, nearly all major aspects of contemporary society.

The delivery of an online service can be adversely affected because of legitimate user activity without any malicious intent. Such situations arise when a large number of users concurrently access a web server, either following some newsworthy event (e.g., the Olympics, the 9/11 attacks), or as a result of redirection from widely followed websites such as Slashdot or other social media like Facebook or Twitter. These situations are called flash events (Ari, Hong, Miller, Brandt, & Long, 2003). These events represent anomalies in the normal Internet traffic with anomalous characteristics such as a substantial increase in the incoming network traffic, overloading of the servers providing the services, and a degradation in the delivery of a service (Bhatia, 2013). Flash events can also have a significant financial implication. A recent example of a flash event causing a substantial financial loss occurred on the 20th of November 2012, when Click Frenzy, a national online shopping initiative in Australia similar to the US-based Cyber Monday events, was launched after big media and online publicity. The website (clickfrenzy.com.au) experienced many times its anticipated traffic volume, leading to a dramatic increase in page-load times, and the failure of the website within minutes of its launch. Organisers of the website had pre-arranged sales partnerships with leading Australian and international retailers and brands, many of whom had paid large amounts for advertisements, which they asked to be refunded (Khanna & Sampat, 2015).

Flash crowd attacks as a result of illegitimate user activities occur almost daily. Even favourite websites, such as Twitter, Facebook, Google, and other popular search engines, cannot escape these attacks, which affect countless users. An eye-

opening case was the DDoS incident that targeted important websites such as the White House, FBI, DOJ, Recording Industry Association of America, Universal Music websites, and Hong Kong Stock Exchange (Company, 2014). During this attack, a total of 80 computers were compromised by the botnet and up to 250,000 were infected with malware. The attack traffic consumed 45 gigabytes per second (Gbps), according to the 7th Annual Report from Arbor Company in 2010. The outage lasted for seven days and was the longest recorded in 2010. In 2011, the longest attack ever recorded was launched targeting a travel company; it lasted for 80 days, 19 h, 13 min, and 5 s. The average duration of a DDoS attack is 9 h and 29 min. In 2012, another large attack reported consumed 60 Gbps, whereas in 2013, the DDoS trend on the application layer raised again and registered 300 Gbps, and the DDoS attack continues in growing in 2014 reached to 400 Gbps. All the above numbers prove that the application-layer attack is the most dangerous and need to be prevented (Network, 2015). Figure 2.1 shows the dramatically increasing bandwidth consumption in DDoS attacks initiated by botnets, as provided by Arbor Networks.

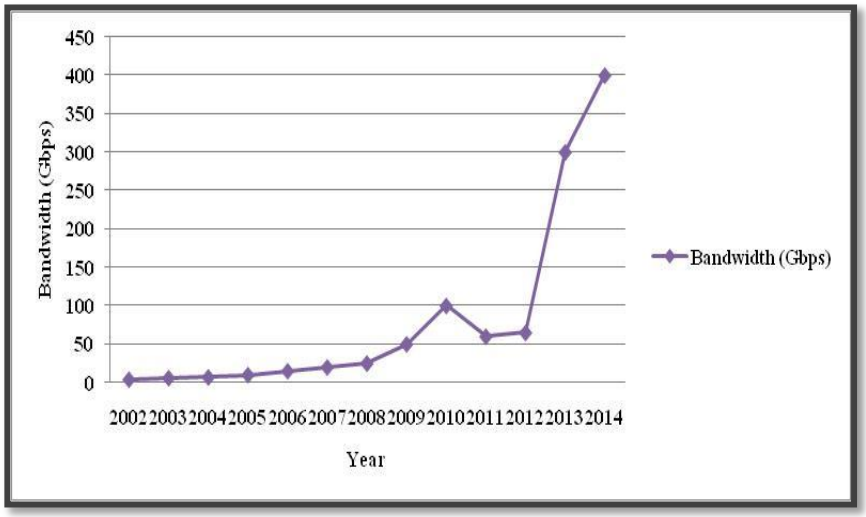

Figure 2.1: Survey of DDoS Attacks (http://www.arbornetworks.com/)

2.2.1 Flash Event Classification

It is argued in this section that flash events may be divided into three broad categories: predictable, unpredictable, and secondary (Pai et al., 2012). In this section, brief descriptions of the three categories are presented. Figure 2.2 shows the three flash event categories.

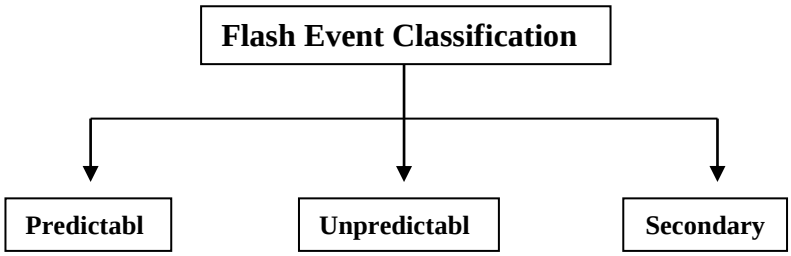

Figure 2.2: The Categories of Flash Crowd

Predictable Flash Event: For this category, the time of occurrence can be expected, allowing network administrators to prepare for the flash event by using various provisioning techniques, such as the use of load-sharing mechanisms or Content Distribution Networks (CDNs). Some popular examples are product releases (e.g. by hi-tech companies like Apple), widely followed sporting events such as the Olympics, or online play-along websites for popular television programs, where the expected time of the incoming traffic burst is well known in advance. The time when the incoming traffic will hit its peak can also be accurately estimated, permitting better handling and provisioning in the case of such events. Moreover, most of the predictable flash events are directed against servers owned by big companies that can handle the constant load or use content-sharing techniques to mitigate the effects of flash events (Bhatia, Mohay, Schmidt, & Tickle, 2012).

Unpredictable Flash Events: The time of occurrence of this type of flash event is entirely unexpected, and these events also cause a sudden and dramatic surge

in network traffic to a site that is supposed to describe the event or provide further leads. The term unpredictable flash event is used to describe the ensuing burst of network traffic. Provisioning for these events in advance is akin to preparing for natural catastrophes like a tsunami or an earthquake (Bhatia et al., 2011).

Designing systems to handle unpredictable flash events is possible but may be economically infeasible due to their unpredictability and rarity. The 9/11 terror attack led to such an unpredictable flash event when major news websites like CNN and MSNBC were overwhelmed by the amount of incoming traffic, pushing their availability close to 0% within minutes after the event occurred. The start and peak-load time of such events are unpredictable and sometimes difficult to identify even post hoc (Hu & Sandoval, 2001).

Their frequencies of occurrence are relatively lower than that of predictable flash events. Figure 2.3 shows an example of an unpredictable flash event when the popular website Wikipedia experienced a sudden increase in its hourly hits following the death of Steve Jobs. Similar traffic was also observed following the death of Michael Jackson.

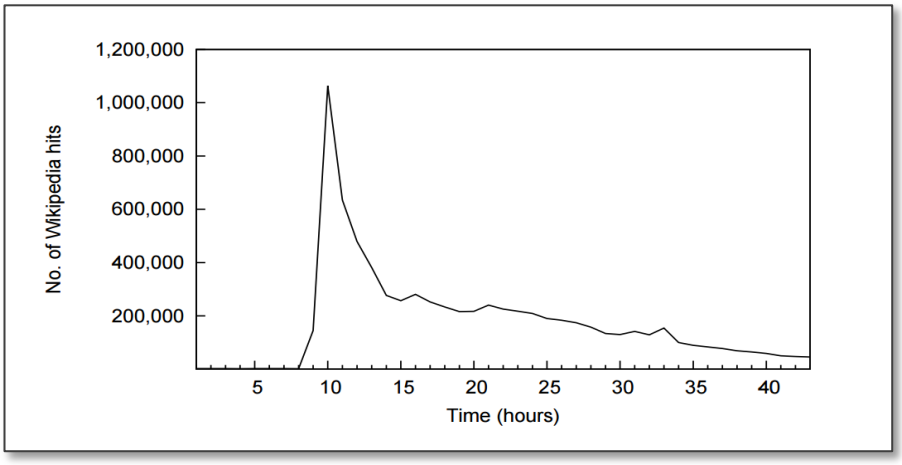

Figure 2.3: Hourly Hits Following the Death of Steve Jobs

Secondary Flash Events: These events usually occur when a brief article, along with a web link is posted on widely followed websites like Slashdot or Facebook. This link is often related to an interesting news item, although not as newsworthy as an event on a world-wide scale that causes an unpredictable flash event. This can capture the attention of a vast number of followers and redirect a high percentage of them to another website in search of additional information. When the usually user-posted articles contain links to poorly resourced websites, this can easily result in the redirection of an unprecedented amount of traffic to those small websites that exceeds their available resources and eventually cripples them. Once again, such events are unpredictable, and the peak-load time is likewise relatively difficult to predict. Provisioning for such events can be challenging but is more feasible than for unpredictable flash events, due to the smaller nature of the event (Anderson, 2008).

The last decade has seen a large number of flash events resulting in website outages. The recent flash events are categorised according to the reason for traffic surge as follows (Dhingra & Sachdeva, 2014).

Flash Events Due to Natural Disasters: In 2012, Hurricane Sandy hit the eastern coast of the United States. The Internet usage on 31 October increased by 114%. Netflix witnessed a traffic volume increase of 150%, while Skype witnessed a service increase of 122%, with notable spike Internet traffic for the day.

Flash Event due to Sports: The 2010 FIFA World Cup, in South Africa, had Internet traffic exceed all the previous records. The leading social website, Twitter, became the primary victim. Normally, it saw 750 tweets per second on an average day, but the traffic rose to approximately 2940 tweets per second, whenever a goal

was scored. These traffic spikes overburdened Twitter's internal network capacity. It saw outages and maintenance downtime throughout the World Cup.

Flash Event due to the launch of new software product: According to TechCentral, Ireland's technology news resource, a unique breakdown occurred at Microsoft, in June 2014 when Exchange Online and Lync Online, part of Microsoft Office 360, were unavailable for hours together. The previously unknown flaw had been detected in the directory partition due to which a large number of customers could not access the email services. Even though connectivity was resumed, the resulting traffic surge overwhelmed a large number of network elements, thus leading to the unavailability of the Lync functionality for a little longer time.

On 18 September, 2013, Apple launched iOS7. Upon the release, these updates caused almost 20% of total network traffic. Thousands of students at various universities in the US began to download it. This led to traffic surges as high as five times the normal traffic levels. Student newspapers also reported outages or slowdown of campus networks.

Flash event due to celebrities: The websites of celebrities also sometimes get affected by flash events. In August 2013, an unusual trigger caused all the previous records of the Twitter's tweets-per-second to be destroyed. It was the broadcast of anime master Hayao Miyazaki's most famous movie "Castle in the Sky". Hundreds and hundreds of Japanese fans of the movie tweeted a magic word used in the classic anime (short for animation), all at once. The word typed was "balse" spoken during the movie's climax scene. The flood of tweets peaked at 143,199 tweets-per-second. Other websites such as Amazon, PlayStation, KFC, and Nissan experienced the failure as soon as the button was pressed.