CHARACTERISTICS BASED DETECTION OF INTERNET WORMS USING COMBINED MACHINE LEARNING METHODS AND WORM CONTAINMENT

CHAPTER 1

INTRODUCTION

1.1. Worm Attack

1.2. Categories of Worm

1.3. Research Motivation

1.4. Internet Worm Attack

1.5. Types of Internet Worms

1.6. Characteristics of Internet Worms

1.7. Internet Worm Defense Mechanisms

1.7.1. Detection Schemes

1.7.2. Containment Schemes

1.8. Problem Statement

1.9. Objectives of the Thesis

1.10. Significant Contributions of the Thesis

1.11. Organization of the Thesis

1.12. Chapter Summary

With the immense growth of Internet, the data transaction process through network is growing highly in communication fields. Along with these rapid developments of Internet applications, the spread of attacks are also increasing. Among the various attacks, worm attack is one of the serious security threats that affect the data [50] [2] [60] [74]. The Internet infrastructures are significantly damaged by Internet worms through their vulnerable exploitation in operating systems, programs and applications. Moreover, these threats are seriously increasing by the introduction of the Internet Of Things (IOT) [20].

1.1. Worm Attack

The worms in the network are computer programs which self-propagate by sending copies themselves from one node to another over a network. Such transmissions occur without any user intervention, thereby allowing them to be spread quickly and easily [40][65][30][67]. Worm propagating in the network infects the vulnerable hosts exploiting the operating system and program [48]. From that affected host, the worm starts its propagation inside the network through injecting other vulnerable hosts for further spreading. Spread of worms in the network is shown in the figure.1.1 below.

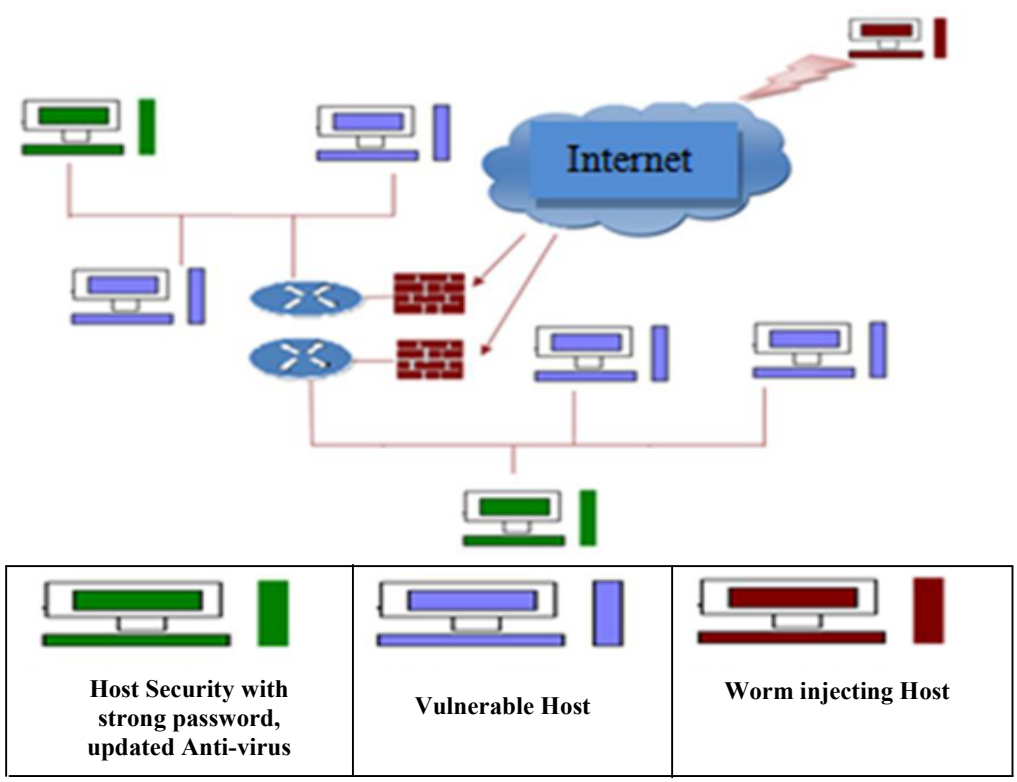

Figure.1.1. Worm Propagation in the Network

Now-a-days, worms are causing dangerous security threats in the network through blocking packets, creating delays and loss of data during transfer [33]. There are various types of worms existing in the network and it is explained below.

1.2. Categories of Worm

The worms are categorized based on their payload and the program format designed by authors. The worms start their target space from the categories listed below [2][58].

- Email Worms
- Instant Messaging Worms
- P2P Worms
- PDF Worms
- Internet Worms

i) Email Worms

These worms spread through the infected email messages [2][58]. The email is said to be infected, when the mail has its link or attachment with the infected website. The email worm starts its propagation immediately when the user clicks the attachment or the link. Email worms attacks the hosts systems and infect the social engineering through its spam characteristics [25] and infects as many host as possible.

ii) Instant Messaging Worms

The Instant Messaging (IM) Worms [2][58] spread through the Instant Messaging applications. This attack links the local contact list with the infected websites through the use of IM users IDs.

iii) P2P Worms

P2P worms spread through the peer–to–peer networks [58], through the shared folder. When the infected folder is shared between the local machines in P2P network, the worm starts its infection.

iv) PDF Worms

This type of worms uses Acrobat PDF Format as its platform for initial infection [58]. Here, Acrobat Reader Program is not directly infected, but only through the VB script embedded inside a PDF file, it is operated as worms.

v) Internet Worms

This worm scans for the vulnerable hosts in the Internet [2][58]. After finding the vulnerable IP address space, the worm starts its spread in the Internet through security flaws in systems.

Though various categories of worms affect the network, the Internet worms are causing high security threats in the Internet. The damages and threats caused by Internet worms motivated to research on Internet Worms.

1.3. Research Motivation

The damages caused by Internet worms during the past 20 years created serious security threats and large financial losses in the world [1][31][32][34][52][53][57][63][75]. Some of the infections and damages caused by Internet worms are listed in table.1.1 below.

Table.1.1. Potential Damages caused by Internet Worms

Year	Worm Names	Infection
1998	Morris	Infected - 60,000 computers Financial Loss - $100 million
2001	Code Red	Infected - 3,50,000 systems Financial Loss - $2.5 billion
2003	Slammer	Infected: 75,000 Computers Damaged: MS SQL servers Achieved : 55 million scans in 3minutes
2004	Witty	Infected: 12,000 hosts in 45 minutes
2007	Storm	Infected : Tens of Millions of hosts
2008	Conficker	Infected: 90% of susceptible hosts within minutes and controlled 6.4 million hosts
2012	Stuxnet	Created cyber war

To avoid expensive damages caused by Internet worms, worm defense schemes are very important for the safety and security of the systems connected to the network. To overcome the problems of Internet worm attack and provide effective defense mechanism, clear knowledge on Internet worms is necessary. The coming sections explain in detail the Internet worm attacks, their vulnerable propagation, various worm types and characteristics of worms to provide better detection and containment of Internet worms.

1.4. Internet Worm Attack

Internet worms [12][17][19][27][69] are malicious codes that propagate automatically by themselves without any human intervention. Internet worm launches most destructive attacks like stealing confidential information, deleting files, reducing the speed of network functioning, creating a Distributed Denial Of Service (DDOS) in the network through Internet communication.

1.5. Types of Internet Worms

There are different types of Internet Worms exist in the network, causing latency and bandwidth limitations. The Internet worms are mainly classified into Active and Passive worms. ***Active worms*** [61] propagate themselves into the network and impose threat to network security without user intervention. Various active worms are code Red II, Slammer, Witty, Nimda, C-Worm and Morris. ***Passive worms*** [50] are similar to virus and they require user intervention or any mechanism behavior to start their propagation. Different existing passive worms are Nelissa, VBS- Gnetella, W32.Gnuman, Fizzer, worm.Lolol.b and worm.Kitro. Some of the recent Internet worms are listed below:

- Polymorphic worms
- Benign worms
- AutoRun worms
- Divide-Conquer- Scanning worms
- Importance Scanning worms
- Self-Disciplinary worms

A. Polymorphic worms

These worms exploit the buffer overflow vulnerability and they have their structure summarized in network protocol frame [10]. Polymorphic worms at each execution of infection alter their byte sequences. These worms initialize their progress through network Protocol commands and exploit the target code.

B. Benign worms

Benign worms exploit software vulnerabilities [55]. Various types of benign worms are passive, hybrid, active and IDS based on spread strategies. SWORM and RWORM are two different benign worms.

C. Auto Run worms

These worms exploit and affect the removable devices [29]. They are dynamic threats. Controlling the usage of removable device decreases the spread of Auto Run worms and maximizes the recovery rate.

D. Divide-Conquer-Scanning worms

Divide-Conquer-Scanning Worms spread through the traditional process of random scanning [11]. Through various infected hosts using different scanning rate, probability and space, these worms infect their targets. These worms are strong epidemic attacks in the Internet.

E. Importance scanning worms

In internet, irregularly distributed vulnerable hosts are exploited by the importance Scanning worms [70]. Vulnerable host distribution is to be calculated and determined to identify and analyze these worms. Static optimal scanning, dynamic optimal scanning and self- learning worms are the different importance scanning worms.

F. Self-Disciplinary worms

Detection probability [62] is decreased by these worms by adapting propagation traffic patterns in infectious computers. To defend against suspicious countermeasures, achieving its exploitation and delay detection, these worms implement their own prorogation patterns. Dynamic and static self-disciplinary worms are of this category. The propagation factors of the different Internet worms are listed in table.1.2.

Table.1.2. Propagation Factors of Internet Worms

Name of Worms	Defensive Mechanisms		Speed		Security Level	
	Proactive	Reactive	Fast	Slow	Host	N/W
Polymorphic	✓	×	✓	×	×	✓
Benign	✓	×	✓	×	✓	×
Auto Run	×	✓	✓	×	✓	×
Divide-Conquer-Scanning	✓	×	✓	×	✓	×
Importance Scanning	×	✓	✓	×	✓	×
Self-Disciplinary	✓	×	×	✓	×	✓

There are different types of Internet Worms propagating through various methods. They are categorized under different specific characteristics and the worms fall under those classified categories.

1.6. Characteristics of Internet Worms

Internet worms' life is divided into four stages namely, discovering targets, transfer of worms, activation and infection [68][43]. Internet worm characteristics are categorized using first two phases, since the worms will be active in Internet during the first two phases only. Internet worms' characteristics are categorized into four main categories namely, Target Discovering Method, Propagation Strategy, Transmission Media and Payload Scheme as shown in figure.1.2.

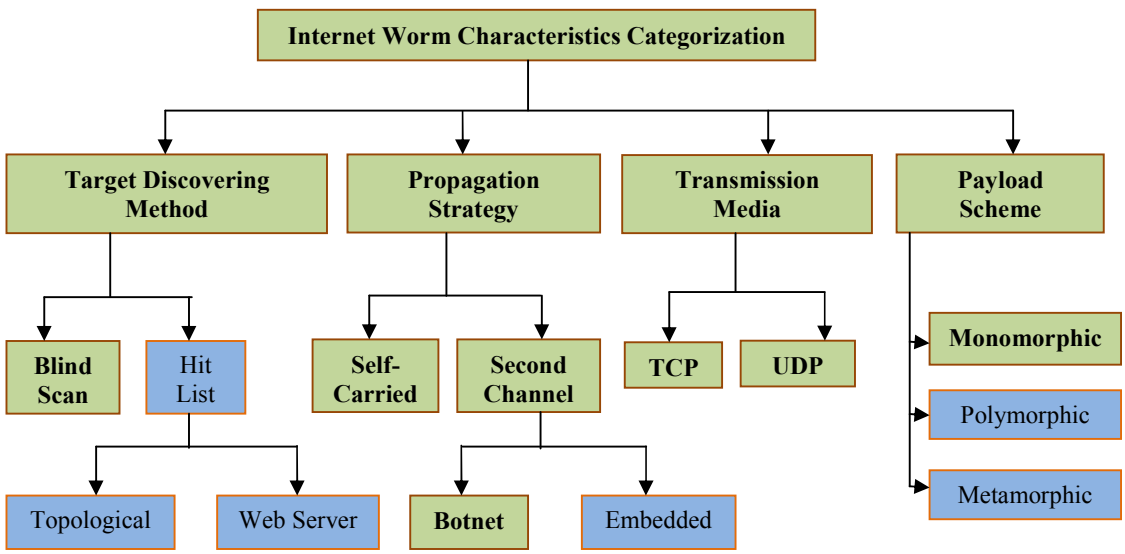

Figure.1.2. Classification of Internet worms characteristics

Target Discovering Method

When the worm enters into the network, its initial step is finding targets to spread and exploit [68][43]. Various target finding schemes are blind target scanning, hit list scanning, topological scanning and web search.

Blind scan Internet worms do not provide prior information about the targets. Blind scan worms create high failure connection rate. Hit list worms using the list of available pre-scanned vulnerable addresses perform its attack. In the Internet, hosts connected with the network stores the details of other hosts, and this type of task helps the attackers to identify those vulnerabilities. Topological worms gather the information and form the path to infect, through the topology of the network. The web search worms identify and discover its target information using the search engines.

Propagation Strategy

After the target is identified by the initial worm, the worm spreads copies of itself to other victims through various schemes [68]. Some of the propagation formats are self-carried, second channel and Botnet schemes. The propagation of self-carried worms are straight forward. Second channel worms propagate from the backdoor of the infected systems or through backdoor. The worms delivered through the Second Channel affects the network through their botnet propagation, where botnet propagation creates abnormal behaviors using different protocol implementation.

Transmission Media

Transmissions of worms are performed through Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) worms [68]. TCP worms are connection oriented and latency limited. These worms block the progress thread. UDP worms infect through self-carried. They are connectionless and bandwidth limited. UDP worms block resources in the network.

Payload Scheme

Payloads are referred as the worm code [68]. If worms are encoded, then worms are harder to detect in the network. Monomorphic, Polymorphic and Metamorphic worm schemes are different worm payload formats. The payloads that transfer the original worm codes without any change are Monomorphic worms. In Polymorphic worms, the forms of

payloads keep changing without affecting the function. The payload form and the function are changed regularly in Metamorphic worms.

The Internet worm based on the above characteristics discovers targets, drives, copies, and chooses transmission media and payload schemes to safe spread. Hence detection based on these characteristics provides better detection of Internet worms, as they are active in the Internet during the process. Better detection and containment of worms can be achieved by effective defense mechanism and is discussed in the next section.

1.7. Internet Worm Defense Mechanisms

Based on the Internet worm characteristics, defense schemes [68][43] have been applied to defend against Internet worms. The defense schemes are divided into two phases namely, worm detection and worm containment as shown in figure.1.3. Internet Worm Detection methods based on characteristics provide better detection because the worms are active on Internet during the above four stages. Internet worms' detection schemes are broadly classified into signature and anomaly based schemes [68][43]. Containment algorithm is used to eliminate the detected worms. Containment schemes are categorized into slowing down and blocking.

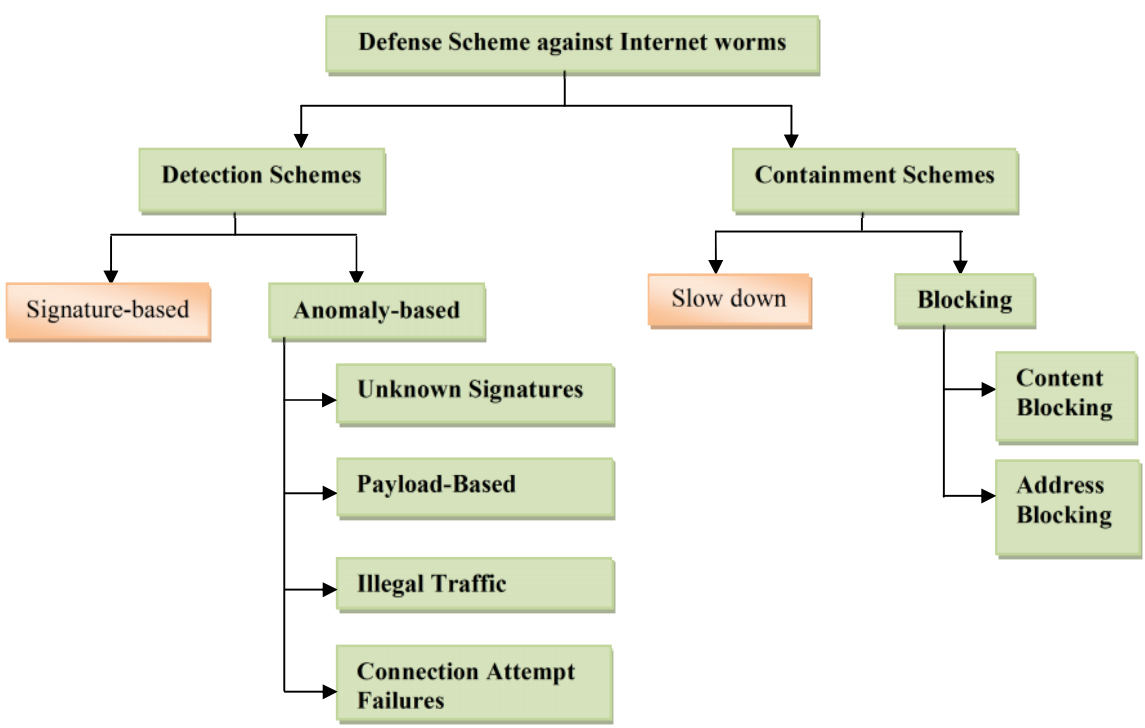

Figure.1.3. Internet Worm Defense Scheme Classification

1.7.1. Detection Schemes

In worm detection, the activities of the Internet worms are analyzed and defended [68]. Worm detection in the defense scheme consists of two types namely, Signature-based and Anomaly-based detection schemes.

Signature-based Detection Schemes

The signature based detection is a technique through which known attacks are detected [68]. The known attacks such as irregular patterns, repeated strings and payload packets are modeled and stored as known signatures. Each and every packet taken into the process is examined with the known signatures.

Anomaly-based Detection Schemes

Anomaly-based detection is used to detect the unknown abnormal behaviors [68]. A threshold is fixed for normal behaviors in the process. If the process exceeds the threshold, an alarm will be triggered, which denotes the occurrence of anomalous behavior. Anomaly–based detection scheme is further classified based on Unknown Signatures, Payload-based, Illegal Traffic and Connection Attempt Failures as shown in figure.1.3.

Unknown Signatures

The signatures detected newly from the process are stored as unknown signatures [68]. The unknown signature database doesn't store all the signatures. After a specific period of time, the signatures in the database will be removed. Hence, the database would not grow, thus reducing the processing time in comparing signatures and the storage space of the database.

Payload-Based

When strings are repeated frequently within a specific period of time, then it is stored as payload [68]. The detection of worm of these types is called as a payload based detection.

Illegal Traffic

The maximum part of the address is not used in the Internet [68]. The packet sent to such unused addresses is done by worm attack. The network traffic created by unused addresses leads to illegal traffic, which is easy for the worms to propagate.