Classification and identification of Malicious codes

Ankur Singh Bist

Govind ballabh pant University of agriculture and technology , pantnagar
ankur1990bist@gmail.com
India

Abstract

Malicious codes are serious threat to our society. The presented work discusses various aspects of technological level, detection, mitigation , identification and Classification of malicious codes. Detection method depicts how various antivirus technologies get used to fight with complex encrypting , decrypting and nature changing virus activities. Antivirus approach consists of waiting for a number of computers to be infected, detecting the virus, designing a solution, delivering and deploying a solution.

***Keywords* —** malware, worm , malicious , replication.

1. Introduction

Wide use of internet has made it a target for malicious code activities . Malicious codes are executable code able to reproduce itself [1]. Viruses are an area of pure programming, and, unlike other computer programs, carry intellectual functions on protection from being found and destroyed. They have to fight for survival in complex conditions of conflicting computer systems. That is why they evolve as if they were alive. Yes viruses seem to be the alive organisms in the computer environment, and yet their, another main goal is survival. That is why they may have complex encrypting/decrypting engines, which is indeed a sort of a standard for computer viruses nowadays, in order to carryout processes of duplicating, adaptation and disguise.[2]
It is necessary to differentiate between reproducing programs and its similar forms. Reproducing programs will not necessarily harm your system. There is big contest between virus creators and antivirus designers and it is becoming more complicated everyday, and will continue afterward . Actions are not an integral part of the virus by default. The enhanced knowledge about the certain trends , specifications can be developed and advancement in the various malicious code decaying codes (antivirus) can be evolved and incremented in most efficient manner[3].

2. Brief look on malicious codes history

1980’s, in the early 1980s, Fred Cohen did extensive theoretical research in computer viruses. Dr. Cohen's definition of a computer virus as "a program that can 'infect' other programs by modifying them to include a version of itself.

1981 Apple Viruses 1, 2, and 3 are some of the first viruses public domain. Found on the Apple II operating system, the viruses spread through via pirated computer games.

1987 In November, the *Lehigh* virus was discovered at Lehigh University in the U.S. It was the first "memory resident file infector". A file-infecting virus attacks executable files. It gets control when the file is opened. The Lehigh virus attacked a file called COMMAND.COM. When the file was run (usually by booting from an infected disk), the virus stayed in the resident memory.

1988 In March, the first anti-virus software was written. It was designed to detect and remove the Brain virus and immunized disks against Brain infection.

1990 Viruses combining various characteristics spring up. They included *Polymorphism* (involves encrypted viruses where the decryption routine code is variable), *Armouring* (used to prevent anti-virus researchers from dissembling a virus) and *Multipartite* (can infect both programs and boot sectors).

1992 Media mayhem greeted the virus *Michaelangelo* in March. Predictions of massive disruptions were made and anti-virus software sales soared. As it turned out, the cases of the virus were far and few between.

1994 A virus called *Kaos4* was posted on a news group file. It was encoded as text and downloaded by a number of users.
1996 *Concept*, a macro-virus, becomes the most common virus in the world.

1998 - The "Red Team" virus infects Windows executables dispatches the infected files through e-mail. The emergence of the sensational "Back Orifice" ("Backdoor.BO") - utility of that allowed hackers management of remote computers and networks.

1999 The *Melissa* virus, a macro, appears. It uses Microsoft Word to infect computers and is passed on to others through Microsoft Outlook and Outlook Express e-mail programs.

2000 The "I Love You Virus" wreaks havoc around the world. It is transmitted by e-mail and when opened, is automatically sent to everyone in the user's address book.

2001: The Code Red worm infects tens of thousands of systems running Microsoft Windows NT and Windows 2000 server software, causing an estimated $2 billion in damages. The worm is programmed to use the power of all infected machines against the White House Web site at a predetermined date. The Anna Kournikova virus hits e-mail servers hard by sending e-mail to contacts in the Microsoft Outlook address book.

2002: February 11: Simile (computer virus) is a metamorphic computer virus written in assembly. My life (computer worm) is a computer worm that spread itself by sending malicious emails to all the contacts in Microsoft Outlook.

2003: The "Slammer" worm infects hundreds of thousands of computers in less than three hours. The fastest-spreading worm ever wreaks havoc on businesses worldwide, knocking cash machines offline and delaying airline flights. it create trouble microsoft sql server.

2004**:** The "My Doom" worm becomes the fastest-spreading e-mail worm as it causes headaches -- but very little damage -- almost a year to the day after Slammer ran rampant in late January 2003. My Doom uses "social engineering," or low-tech psychological tricks, to persuade people to open the e-mail attachment that contains the virus. It claims to be a notification that an e-mail message sent earlier has failed, and prompts the user to open the attachment to see what the message text originally said. Nuclear RAT (short for Nuclear Remote Administration Tool) is a backdoor Trojan Horse ,infects Window NT family systems (Windows 2000, Windows XP, Windows 2003).

2005:The Zlob Trojan, also known as Trojan. Zlob, is a trojan horse which masquerades as a required video codec in the form of ActiveX. It was first detected in late 2005 .

2008 : Rustock.C , a hitherto-rumoured spambot-type malware with advanced root kit capabilities, was announced to have been detected on Microsoft systems and analyzed, having been in the wild and undetected since October 2007 at the very least. Mocmex_ is a trojan, which was found in a digital photo frame in February 2008. It was the first serious computer virus on a digital photo frame.

2009 , 9 million computers running on Windows operating system were hit with ***"Conficker"*** worm. The malware spread via the Internet and the main tools that helped the worm spread were unpatched corporate networks and USB memory sticks. First discovered last October, it loads itself on to a computer by exploiting a weakness in Windows servers. Once it has infected a machine, the software also tries to connect to up to 250 different domains with random names every day.

2010: Stuxnet, a Windows Trojan, was detected. It is the first worm to attack SCADA systems. There are suggestions that it was designed to target Iranian nuclear facilities. It uses a valid certificate from Realtek.

The virus, called "here you have" or "VB Mania ", is a simple Trojan Horse that arrives in the inbox with the odd-but-suggestive subject line "here you have". The body reads "This is The Document I told you about, you can find it Here" or "This is The Free Download Movies, you can find it Here".

2011: The Morto worm emerged in the summer of 2011. It attempts to propagate itself to additional computers via the Remote Desktop Protocol (RDP). Morto spreads by forcing infected systems to scan for servers allowing RDP login. Once Morto finds an RDP-accessible system, it attempts to log in to a domain or local system account named 'Administrator' using a number of common passwords. A detailed overview of how the worm works—along with the password dictionary Morto uses—was done by Imperva.

. 3. Phases of virus:-

(1)Dormant phase

(2) Propagation phase

(3)Triggering phase

(4) Execution phase

4. Types of Viruses:

Boot sector virus , Polymorphic virus ,Time Bomb, Shell virus, Add-on virus ,Trojan horse ,Internet worms.

5.Virus Definitions[10]:-

(5.1)Boot Sector Virus:
Replaces or implants itself in the boot sector. This kind of virus can prevent you from being able to boot your hard disk.

(5.2)Macro Virus:
Written using a simplified macro programming language, these viruses affect Microsoft Office applications, such as Word and Excel. A document infected with a macro virus generally modifies a pre-existing, commonly used command (such as Save) to trigger its payload upon execution of that command.

(5.3)Email viruses and worms:
An e-mail virus moves around in e-mail messages, and usually replicates itself by automatically mailing itself to dozens of people in the victim's e-mail address book. A worm is a computer program that has the ability to copy itself from machine to machine. Worms normally move around.

(5.4)Multipartite Virus:
Infects both files and the boot sector-- a double whammy that can reinfect your system dozens of times before it's caught.

(5.5)Polymorphic and stealth virus:
Changes code whenever it passes to another machine. Stealth virus hides its presence by making an infected file not appear infected.

(5.6)Logic bomb and flooders:
Triggers action when condition occurs .Use to attack network computer system with large traffic cause dos(denial of services) .

(5.7)Root kit and Zombie:-
set of hacker tool used after attacker has broken into a computer system and gained Root level access zombie program activated on an infected machine that is activated to launch attack on other machines.

(5.8)Metamorphic virus:-
this virus mutates with every infection , metamorphic virus rewrites itself completely with every iteration ,so typical to detect, they change their behavior as well as appearance.

6. Detection and classification method:-

1. Heuristic technology

2. Rule based system

3. Checksum

4. Scan string

7. Logic for compressed virus:-

```
Program CV
{ goto   main:
  01234567;
Subroutine infect executable=
  {
    loop:
          file:=get random-executable-file;
if(first-line-of-file=01234567)then goto loop;
compress   file;
       prepend   cv to file;
        }
Main: main program=
          {
      if ask permission then infect-executable;
  uncompress  rest-of-file;
   run compressed file;
   }
         }
```

* here virus does nothing other than propagate .

8.Viruses detection techniques:-

Virus monitors detection by behavioral abnormality. In this approach to virus detection, the machine is booted from uninfected files and a virus monitor is installed that monitors various activities of the machine while in day to day use.

(8.1)Detection by emulation:-

In this scheme the program under test is emulated by the virus detection program, which attempts to determine the run-time behavior of the program. This is different from monitoring in that the program is not observed while it is actually executing but is emulated with sample input(s)[10][11].

(8.2)Detection by static analysis/ policy adherence:-

This method examines a program to decide whether it meets a pre specified policy requirement, one which may include integrity requirements for the detection of viruses.
To determine whether an arbitrary program contains a virus is undecidable , but conservative decision on the presence of viruses may be possible[10][11].

(8.3)Detection by check summing:-

To protect programs against unwanted modification by viruses, a checksum based on the contents of the program is computed and stored in encrypted form within or outside the program. The encryption is done using a one-way function so forgery of a correct checksum after infection is computationally very hard.

(8.4)Detection by signature scanning:-
Using the “signature” of a virus to detect its presence in an executable is the simplest and most common approach to known virus detection. Once a virus is isolated, a sequence of bytes (unique sequence) from its code is taken as the identifying string for that virus[10][11].

(8.5)Heuristic analysis:-
It is very useful method to know unknown viruses but it may produce false positive outputs. This technique analyze code behavior and organization and tries to gather information.

9. Latest research Approaches in detection :-

1. Data mining techniques in worm detection

2. Natural virology mapping

3. Layering techniques

4. Virus detection using artificial intelligence

10.Malicious code environment

It is important to know about the particular execution environments to understand about Computer Viruses . A successful penetration of the system by a viral code occurs only if the various dependencies of malicious code match a potential environment.

1) Computer Architecture Dependency
2) CPU Dependency
3)Operating System Dependency and Operating System version Dependency
4) File System Dependency
5) File Form Dependency

11. Anti Virus Software Works strategy:-

1. Scanning for known files

2. Scanning for malicious infection vectors

3. Heuristic scanning and by using other defined detection mode.

4. Comparison of different detection measures:-

Various virus detection methods can be speed up using hashing and hybrid architecture approach but still the impact of false negative alarms are there .most of detection method are not powerful against evolutionary advanced or new viruses. String scanning techniques promise perfect disinfectant but in case of encrypted and new viruses it is not effective . Other techniques that tries to map with new viruses are no so effective since the problem of self and nonself is still present there. the dynamic decryptor approaches suffer from same scenario. the undecidability of malicious codes and their active environment cause various problems in front of us.

12. Conclusions and future work

In this survey , I have gone through the basic definitions of various malicious codes and a brief history of viruses irrespective of this various behavior come into picture while describing different type of malicious codes ,and it includes various approaches used by antivirus methodology and virus preventing measures . In future a strong biological antiviral model Can be modeled , strong formation of gene library, strong identification of non self and strong signature matching techniques, mutating nature prevention ,strong cbr (case based reasoning) techniques, of viruses can lead to design of more advanced platform for detection .

REFERENCES

[1] dr. Solomon's virus encyclopedia , 1995,isbn 18097661002,abstract at http://vx.netlux.org/lib/aas.html
[2] http://www.bartleby.com/61/97/0539700.html
[3] "what is computer virus?" actlab.utexas.edu1996-03-31 retrieved 20 10-08-27
[4] von Neumann 1996 "theory of self replicating automata"
[5] kaspersky lab virus kit
[6] http:// virus.wikia.com/wiki/
[7] http:// fsecure.com//
[8] kimono (march 3,2008) "mbr rootkit,a new breed of fsecure retrieved"
[9] http:// www.precisecurity.com/rogue/xp-antispyware-2011//
[10] cryptography and network security by William stalling
[11] classification and identification of malicious code thesis//Hamburg 2003