Behavioral Malware Detection in Delay-Tolerant Networks

[1] Najim Sheikh, [2] Priya Patil, [3] Sneha Pillewan, [4] Suman Suryavanshi, [5] Vaishnavi Nilapwar

[1] M-Tech 1st RGPV, [2] BE 2nd RTMNU, [3] BE 3rd RTMNU, [4] BE 4th RTMNU, [5] BE 5th RTMNU

Abstract - Delay Tolerant Networking (DTN) is pioneered as an approach in network architecture to address the technical problems in non-homogeneous networks that may reduce continuous network connectivity. The behavioral characterization of malware based on Naive Bayesian model is an alternative approach to pattern matching for detecting proximity malware. Computer is an important part of an everyday life to many people across the world. Computer in the hand of consumer to lack the knowledge of protection tools and to have limited administrator skill are vulnerable to Virus attack .these system are extremely valuable to intruders as they have lot of secret personal information about the users. Attacker exploit vulnerabilities in the software layers to install malicious program on users Machine to steal secret data for financial gains. Security protocols have been in place for some time to counter the threat posed by the attack however, despite the presence of such measures; the number of attacks on consumer computer is growing rapidly. A recent trend in attacks has been the attempt to disable security protocol In a place at the host machine. This type of attacks leaves the host computer completely defenseless and vulnerable to many further exploits through the internet.

***Keywords** - DTN,* **Signature-Based Malware Detection Techniques.**

1. Introduction

Malware is a collective term for any malicious software which enters system without authorization of user of the system. The term is created from merging the words ‘malicious’ and ‘software’. Malware is a very big threat in today’s computing world. It continues to grow in volume and evolve in complexity. As more and more organizations try to address the problem, the number of websites distributing the malware is increasing at an alarming rate and is getting out of control .Most of the malware enters the system while downloading files over Internet. Once the malicious software finds its way into the system, it scans for vulnerabilities of operating system and perform unintended actions on the system finally slowing down the performance of the system. Malware has ability to infect other executable code, data/system files, boot partitions of drives, and create excessive traffic on registry entries in Windows, buffer overflow, format string etc.

Network leading to denial of service. When user executes the infected file; it becomes resident in memory and infect any other file executed afterwards. If operating system has vulnerability, malware can also take control of system and infect other systems on network. Such malicious programs (virus is more popular term) are also known as parasites and adversely affect the performance of machine generally resulting in slow-down .Some malware are very easy to detect and remove through anti-virus software. These anti-virus software maintains a repository of virus signatures i.e., binary pattern characteristic of malicious code. Files suspected to be infected are checked for presence of any virus signatures. This method of detection worked well until the malware writer started writing polymorphic and metamorphic malware. These variant of malware avoid detection through use of encryption techniques to thwart signature based detection. Security products such as virus scanners look for characteristics byte sequence (signature) to identify malicious code. The quality of the detector is determined by the techniques employed for detection. A good malware detection technique must be able to identify malicious code that is hidden or embedded in the original program and should have some capability for detection of yet unknown malware .Commercial virus scanners have very low resilience to new attacks because malware writers continuously make use of new obfuscation method.

2. Malware Types

Malware can be broadly classified into following categories.

2.1 Viruses

Computer virus refers to a small program with harmful intent and has ability to replicate self. Mode of operation is through appending virus code to an executable file. When file is run, virus code gets executed. The original virus may evolve into new variants by modifying itself as in case of metamorphic viruses. A virus may spread from an infected computer to other through network or corrupted media such as, floppy disks, USB drives.

Viruses have targeted binary executable file (such as .COM and .EXE files in MSDOS PE files in Windows etc.), boot records and/or partition table of floppy disks and hard disk, general purpose script files, documents that contains macros.

2.2 Worms

Worms are self replicating programs. It uses network to send copies of itself to other systems invisibly without user authorization. Worms may cause harm to network by consuming the bandwidth. Unlike virus the worms do not need the support of any file. It might delete files, encrypt files in as crypt viral extortion attack or send junk email. Example Sassier, My Doom, Blaster, Melissa etc

2.3 Spyware

Spyware is a collective term for software which monitors and gathers personal information about the user like the pages frequently visited, email address, credit card number, key pressed by user etc. It generally enters a system when free or software trial is downloaded.

2.4 Adware

Adware or advertising-supported software automatically plays, displays, or downloads advertisements to a computer after malicious software is installed or application is used. This piece of code is generally embedded into free software. The problem is, many developers abuse ad – supported software by monitoring Internet users' activities .The most common adware programs are free games, peer-to-peer clients like Kazak, Bear Share etc.

2.5 Trojans

Trojan horses emulate behavior of an authentic program such as log in shell and hijacks user password to gain control of system remotely. Other malicious activities may include monitoring of system, damages system resources such as files or disk data, denies specific services.

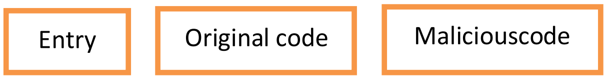
2.6 Botnet

A botnet is remotely controlled software collection of autonomous software robots. It is usually a zombie program (Worms, Trojans) under common control on public and private network infrastructure. Botnets are usually used to send spam /spyware remotely. Bot doesn't sit around on machine (infected machine) waiting for the instruction from a third party instead it looks for the communication with similar instances of bots awaiting instructions. Simplest bot configuration is where the bots are connected to single central hub. This configuration does not scale much because maintenance of various connections over single server is difficult. The next configuration is hierarchical structure where bot master connects to hundreds of bots which in turn is connected to many bots. Thus this configuration would scale.

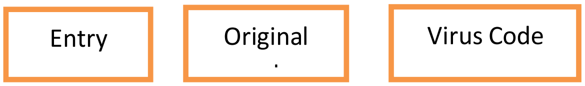
3. Related Work

3.1 Signature-Based Malware Detection Techniques

Commercial anti-virus scanners look for signatures which are typically a sequence of bytes within the malware code to declare that the program scanned is malicious in nature. Basically there are three kinds of malwares: basic, polymorphic, metamorphic malwares. In basic malware the program entry point is changed such that the control is transferred to malicious payload. Detection is relatively if the signature can be found for the viral code. Figure 1 show basic malware.



Polymorphic viruses mutates while keeping the original code intact. A polymorphic malware consists of encrypted malicious code along with the decryption module. To enable the polymorphic virus the virus has got polymorphic engine somewhere in the virus body. The polymorphic engine generates new mutants each time it is executed. Signature based detection for such a virus is difficult because each variant new signature is generated which makes signatures based detection difficult. Strong static analysis based on API sequencing is used for polymorphic virus detection [9].Figure 2 shows polymorphic malware's.



Metamorphic malware can reprogram itself using certain obfuscation techniques so that the children never look like the parent [4]. Such malwares evade the detections from the malware detector since each new variant generated will have different signature, hence it is impossible to store the signatures of multiple variants of same malware sample. In order to thwart detection a metamorphic engine has to be implemented with some sort of disassemble in order to parse the input code. After disassembly, the engine will transform the program code and will produce new code that will retain its functionality and would still look different from the original code Figure 3 shows metamorphic malware and multiple signatures for multiple variants.

3.2. Behavior based Detection

Behavior based detection differs from the surface scanning method in that it identifies the action performed malware rather than the binary pattern. The programs with dissimilar syntax's but having same behavior are collected, thus this single behavior signature can identify various samples of malware. This types of detection mechanisms helps in detecting the malwares which keeps on generating new mutants since they will always use the system resources and services in the similar manner. The behavior detector basically consists of following components which are as follows Data Collection: This component collects the dynamic / static information's are captured. • Interpretation: This component converts the raw information collected by data collection module into intermediate representations. • Matching Algorithm: It is used to compare the representation with the behavior signature.

4. Flowchart

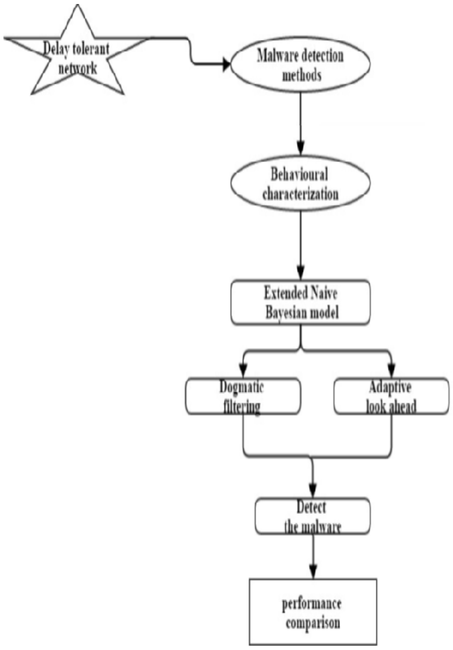


Fig. 1 Flowchart

5. Result

In this research, we have proposed a malware detection module based on advanced data mining and machine Learning. While such a method may not be suitable for home users, being very processor heavy, this can be implemented at enterprise gateway level to act as a central anti-virus engine to supplement anti viruses present on end user computers. This will not only easily detect known viruses, but act as a knowledge that will detect newer forms of harmful files. While a costly model requiring costly infrastructure, it can help in protecting. Invaluable enterprise data from security threat, and prevent immense financial damage.

6. Conclusion

In this survey a series of malware detection techniques have been presented. The problems related to traditional signature based detection method is also highlighted. Rate of new malware's causing destruction to systems worldwide is increasing at alarming rate. Detection of malware's changing their signatures frequently has posed many open research issues. Challenge lies in the development of good disassemble, similarity analysis algorithm so that the variants of malware's can be detected in shorter time there by reducing the computation over head. Behavioral characterization of malware is an effective alternative to pattern matching in detecting malware, especially when dealing with polymorphic or obfuscated malware. Naive Bayesian model has been successfully applied in non-DTN settings. The general behavioral characterization of DTN-based proximity malware with look ahead proposed, along with dogmatic filtering and adaptive look ahead, to address two unique challenging in extending Bayesian filtering to DTNs: "insufficient evidence versus evidence collection risk" and "filtering false evidence sequentially and in a distribute manner."

References

[1] Trend Micro Inc. SYMBOS_CABIR.A., http://goo.gl/aHcES, 2004.
[2] http://goo.gl/iqk7, 20 13.
[3] Trend Micro Inc. IOS_IKEE.A., http://goo.gl/z0j56, 2009.
[4] P. Akritidis, W. Chin, V. Lam, S. Sidiroglou, and K. Anagnostakis, "Proximity Breeds Danger: Emerging Threats in Metro-Area Wireless Networks," Proc. 16th USENIX Security Sympssssss., 2007.
[5] A. Lee, "FBI Warns: New Malware Threat Targets Travelers, Infects via Hotel Wi-Fi," http://goo.gl/D8vNU, 2012.
[6] NFC Forum.about NFC, http://goo.gl/zSJqb, 2013.