Attack Pattern Analysis: Trends in Malware Variant Development

U.A.Abu Bakar[1], S.M.Furnell[1], M.Papadaki[2] and G.Pinkney[2]

[1] Network Research Group, University of Plymouth, Plymouth, United Kingdom
[2] Symantec, Berkshire, United Kingdom
e-mail: info@network-research-group.org

Abstract

This paper presents an investigation into recent trends and patterns in malware variant development targeting the Microsoft Windows environment. This research focuses on three significant malware threats: Beagle, Netsky and Mytob; which were all successful mass mailing worms and unique in terms of their propagation techniques and functionality. The results from this investigation showed that mass mailing worms still prove to be the preferable propagation method, but other techniques are also required to ensure it becomes successfully widespread. Mass mailing worms also continues to prove successful in terms of their propagation speed and widespread distribution.

Keywords

Mass-Mailing Worm, Malware Analysis, Beagle, Netsky, Mytob

1. Introduction

Modern malware poses a major security threat to computer systems due to the speed it can spread over the Internet, exploiting the flaws and vulnerabilities in many systems. The threat is growing because malware is continually propagating in smaller time periods and malware routines are getting more complex. Having early detection and taking early prevention steps is better than having to clean up systems after they have been infected. One approach that may aid the process of detection and prevention is to better understand the malware attack patterns and trends. This may help malware researchers to expand the existing information on malware, and thus improved preventive actions could be taken.

The objective of the research presented in this paper is to determine if there are any trends and patterns in malware attacks from the perspective of variant development, focusing on the Microsoft Windows platform. Three malware examples are chosen for this purpose: Beagle, Netsky and Mytob; which have all been responsible for significant and widely spread malware incidents. To aid discovery of any trends or patterns that may exist in these malware attacks, this research proposes to answer the following investigation questions: 1) Is there a correlation between nature of malware development and time? 2) Is the creation of new malware based upon re-use of existing code and techniques, or is it a creation of entirely new techniques?

This paper is organised as follows: Section 2 presents the definitions of malware and overview of the current malware scene. Section 3 outlines data techniques and results from the analysis on malware variant development. In Section 4, the analysis results are used to derive the comparison and summary on discovery of trends in malware variants development. Finally, Section 5 presents the overall conclusions of this research.

2. Malware Overview

This section outlines some definitions in malware taxonomy related to this research and an overview of recent malware attack trends.

2.1 Definitions

Malware is short form for "malicious software", and refers to software that contains code which is typically designed to perform malicious activity or damage to single computer, networked computers or servers (Microsoft, 2004). Examples of malware are worm, virus, and Trojan horses. Malware is grouped according to its unique characteristics. However these days, it is rather difficult to provide perfect definition and categorisation of malware because some malware may show behaviour that fits into one or more category. The following description of malware is not exhaustive but they provide the basic definition (Szor, 2005; Microsoft, 2004):

Virus	A program that is written to enable replication of itself and it need host to infect as it does not function on its own. Each time the host program is executed; the virus is executed as well and reproduces itself by attaching to other programs.
Worm	A stand alone program that can propagate, replicate and distribute itself on networks with or without user intervention.
Trojan horses	A Trojan horse disguises itself as a program with some useful functions but also contains hidden code that, when executed, perform some malicious function.
Bots	An agent designed to seek and infect vulnerable machines, run silently in the background to open network ports and allow outbound connection, typically to IRC channel which is remotely controlled by attacker (SANS, 2003).

2.2 Current Malware Scene

The scope for malware analysis presented in this paper is on malware that was recent at the time of the investigation. Thus, this section presents the discussion on some notable trends in current malware attacks scene from the period of Q1 2003 until Q1 2005.

i. *Mass mailing worms continue to be dominant and prevailing threat.* E-mail continues to be the most effective vector for malware propagation as many successful worm outbreak uses e-mail as their propagation vector. Since the first major outbreak of a mass mailing worm, Melissa in 1999, this technique has been employed widely by other malware. 2003 saw mass mailing worms like Sobig,

Klez and Bugbear dominating the top of the malware charts and their propagation techniques were basically derived from previous mass mailing worms. The trend continued in 2004, where mass mailing worms like Netsky, Beagle and MyDoom were discovered and they all dominate AV (Anti-Virus) vendor's malware charts until the end of the study period. These worms are also derivative upon previous mass mailing worms (Symantec, 2004b; Trend Micro, 2005).

ii. *Increase threats to confidential information.* Recent trends showed that the threat to confidential data has increased significantly each year (Symantec, 2003; Symantec, 2004a; Symantec, 2004b; Symantec, 2005a). Malware are created with malicious intentions of stealing confidential information, such as financial information, passwords and login cache from the compromised machines. This trend also shows companionship between malware to serve this purpose. For example, Mytob utilises Mydoom as its main distribution technique, and uses SDbot to serve the purpose of stealing information.

iii. *Bots and its numerous variants.* The current breed of bot worms showed an enormous amount of variants and the number is still continuing to rapidly increase. During 2004 alone, Trend Micro documented 2,830 bot programs (Trend Micro, 2004).

iv. *Increment and commonness of blended threats*. Blended threats use multiple combinations of malicious codes and exploit multiple vulnerabilities (Szor, 2002). The numbers of malware that employ blended threats continues to increase and they also demonstrate more complexity in routines and are also targeting more new vulnerabilities. Recent malware outbreaks employing blended attacks have caused major damage and are always rated with high severity, with examples including Netsky and Mydoom (Symantec, 2004a; Symantec, 2004b; Symantec, 2005a).

2.3 Recent Successful Malware

The trends in recent malware scene have drawn the attention to three particular cases: Beagle, Netsky and Mytob. The following will discuss the background of these examples.

i. *Beagle* – Beagle or Bagle was first reported by Symantec on 18 January 2004. As a mass mailing worm, it propagates using its own SMTP engine, scans for e-mail addresses in local drives with various file extensions and uses spoofed sender addresses. Furthermore, Beagle employs social engineering techniques by using a pre-configured list of e-mail subject and message body, and spread itself in the e-mail attachment with various extensions such as .exe, .scr or .zip. It also propagates via peer-to-peer (P2P) and shared folders. Later variants also utilise its own DNS server if MX record is unavailable on local drives. E-mails sent by Beagle may clog communications and degrade network performance. Beagle also installs a backdoor and opens network ports to allow remote command execution, and subsequently attempts to connect to lists of websites via HTTP GET request and relay back information about the infected machine to its attacker server. Beagle also attempts to terminate AV and security related

services. Beagle also created mutexes which may be used by Netsky in attempt to prevent Netsky from executing in Beagle's infected machines. Later variants saw Beagle start to send copies of Trojan Tooso via its mass mailing capability. This Trojan attempts to disable security-related services, as well as degrade computer performance when it performs remote file download from pre-configured list of websites (Symantec, 2004b; Symantec, 2005a; Symantec, 2005b;Trend Micro, 2004; Trend Micro, 2005).

ii. *Netsky* – First reported on 16 February 2004 by Symantec, Netsky is a mass mailing worm that propagates via e-mail, LAN and P2P networks. Like Beagle, it also adds a value to a Registry key so that the worm runs automatically during Windows startup. Netsky spreads inside e-mail attachments and employs social engineering techniques to trick users into opening the e-mail and attachment. E-mails sent by Netsky contain a spoofed sender address with variable names for e-mail subject and message body. Attachments that contain the body worm are usually in .zip, .pif, .scr or .exe file. The file that contained the worm is usually compressed with various types of packers such as UPX, Petite, FSG and so on. Netsky also propagates via a known vulnerability exploit. After July 2005, no new version of Netsky has been discovered, but the infection from old variants from 2004 was still widespread (Symantec, 2004b; Symantec, 2005a; Symantec, 2005b).

iii. *Mytob* – Discovered by Symantec on 28 February 2005, Mytob demonstrate a new wave in worm creation where it combines the code and functionality of an early version of the mass mailing worm Mydoom and the IRC bot SDbot. As part of Mydoom, Mytob arrived as an e-mail that has a pre-configured list of subjects, message body and attachment with file extensions of .pif, .zip, .exe, .scr, .cmd or .bat. Mytob e-mailed itself by gathering addresses from local drives with various file extensions including .htm, .wab, .asp, and others. It also propagates via P2P networks and via two known Microsoft vulnerability exploits. As part of SDbot, it has backdoor capability to open ports and connect to pre-configured list of IRC channel in order to listen for remote command from the attacker which then performs action like update itself or stealing information from the compromised machines. Mytob became widespread within a short period of time, and with its ability to compromise machines and steal confidential information, Mytob was rated as high severity (Symantec, 2005a; Symantec, 2005b; Trend Micro, 2004; Trend Micro, 2005).

3. An Analysis of Malware Variants

Having identified the three worms as notable examples of recent developments, this section discusses the techniques used for data collection and analysis within the study, and will then examines the three worms in more detail and considers the characteristics that helped to make them successful.

3.1 Data Collection and Analysis Techniques

The analysis performed in this paper is done upon set of data that was collected from Chronological Virus List on www.secunia.com (Secunia, 2005). This source was selected because it provided comprehensive reports on new malware discovery everyday and organised it chronologically. The malware reports were basically derived from seven well-recognised AV vendors. All data in Secunia was grouped and indexed, and it contains information on the date a threat was reported and updated, its aliases, file sizes, severity ratings, description and links to its reporting vendor. The period for data collection was from 1 March 2004 until 1 August 2005. To verify the data collected, a number of random entries from the collected data are picked and information on the malware is checked with the original AV vendor website.

The data collection was performed by utilising a Bash script which automates the collection process. The script functioned by downloading the relevant page from Secunia and then stripping out the HTML and finally converting the output into a CSV file for easy manipulation. This data was then organised and split into multiple separate CSV files according to type of malware and its reporting vendor to improve the process of identifying trends. From the sorted data, type of information that could be extracted was the discovery date of new malware and its variants. From here, the selected data on Beagle, Netsky and Mytob was plotted into cumulative distribution plots and the distributional information is analysed. The analysis was performed using qualitative analysis where information and description of the worms' variants were gathered and compared to find trends.

3.2 Summary of Analysis Results

The results of analysis for each worm variants development are summarised in the following sections.

3.2.1 Beagle

The nature of Beagle development is shown in Figure 1, where the first few releases were in parallel with Netsky. The competitive growth was resulted from the effect of Netsky attempt to delete Registry keys used by Beagle and Beagle keep releasing more variants to extend distribution.

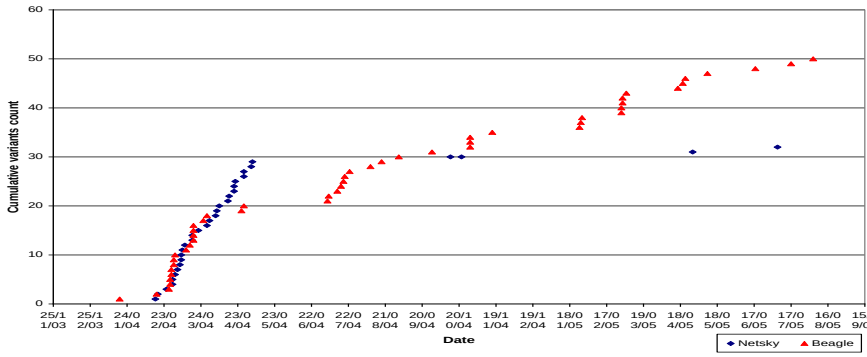

Figure 1: Development of Beagle variants versus Netsky variants

From the initial release, the trend sees clusters of variants released with rapid succession and the overall trend appears roughly linear throughout. Analysis on each variant clustering has shown correlation with its functionality, and each clustering shows improvement over time. The first cluster sees various techniques being tested, where it starts to include its own DNS server, terminate security-related services, create random ID for the infected machines, delete Netsky registry key and drop Trojan Mitglieder. The second clustering sees a similar technique, but with improvement in terms of added pre-configured list of e-mail subject and message body, and the latest clustering see that Beagle showing "new faces" in its pre-configured list of e-mail subject and bodies and it also sends out copy of Trojan Tooso via its mass mailing technique.

3.2.2 Netsky

Figure 1 shows the speedy release of Netsky in four months since its first outbreak. Note that the development was in parallel with Beagle, and the Figure is indicative of the competition between them, which explains why the curve of the both Beagle and Netsky early variants release was steep. Netsky's author was arrested in May 2004 and four variants discovered after that only reflect modification and re-use of the worm's code by somebody else. As a mass mailing worm that did not produce any more new iterations, Netsky proved to be very successful based upon its prevalence in malware charts. Figure 2 shows the number of infected machines for each Netsky variant since their first release and illustrate Netsky.P as the most successful variant of the family. The reason for this is because it employs powerful social engineering, built in SMTP engine, redundancy technique to retrieve SMTP server, exploit IE Incorrect MIME Header vulnerability and spread to P2P, LAN, FTP and HTTP server's folders, all in one variant.

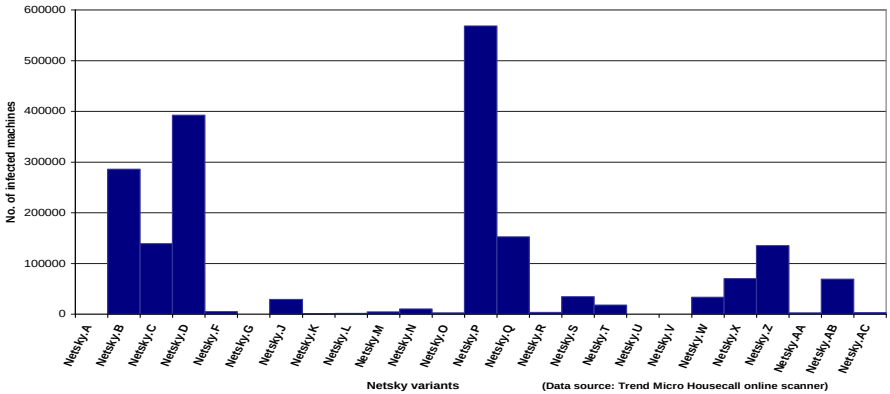

Figure 2: Number of infected machines by Netsky family

3.2.3 Mytob

Figure 3 shows the rapid development of Mytob in less than 6 months from its first outbreak.

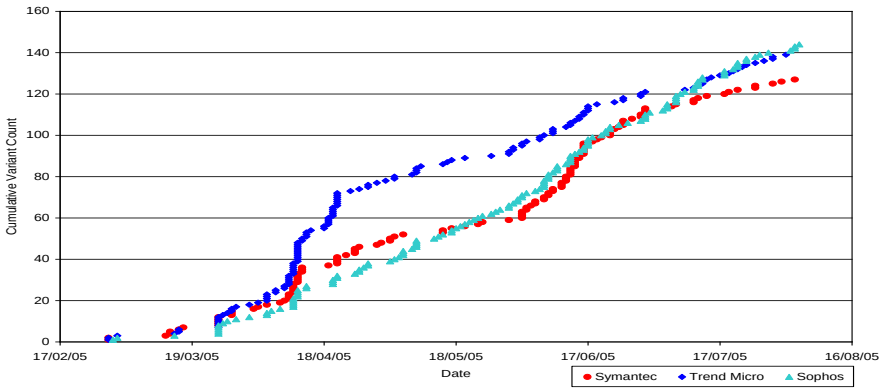

Figure 3: Development of Mytob variants According to Symantec, Trend Micro and Sophos

Since the first outbreak until 3 August 2005, Symantec has discovered 127 variants. The similar trend was also discovered by Trend Micro and Sophos. The rapid succession of variant release reflects that each new version does not show much adjustment from the previous one, but just enough to need new signatures to detect them. This also reflects the bot part in Mytob, which is to steal and relay information from the compromised machine back to its remote attacker and therefore it need to race against AV new signature deployment to ensure it compromised as many machines to make sure the number of compromised machines in the bot network is maintained. The success of this hybrid worm is mainly because of the technique it used, which is mass mailing and vulnerability exploit from Mydoom, and bot functionality from SDbot.

4. Discussion

Having analysed the three worms, it can be seen that the speedy development and the success of these worms proved that mass mailing is still the preferable technique for propagation. However, it is worth noting that using only the mass mailing technique is not adequate, but must be employed with other techniques as well to ensure the widespread of these worms. Success of these worms do not depend on the complexity of codes or technology, but mainly upon the right combination of techniques used to make the worms easily propagate, evade AV software detection and trick users into trusting the incoming e-mail, and open the infectious attachment. Table 1 summarises the characteristics that contribute to these worms' success.

		Beagle	Netsky	Mytob
Propagation technique	Mass mailing	✓	✓	✓
	P2P and shared networks	✓	✓	✓
	Vulnerability exploit		✓	✓
E-mail technique	Built-in SMTP engine	✓	✓	✓
	Social engineering	✓	✓	✓
	Password protected attachment	✓	✓	
Payload	Open backdoor	✓	✓	✓
	Attempt to terminate security-related services	✓	✓	✓
	Perform DoS/ DDoS	✓	✓	✓
	Steal information	✓		✓
	Dropped other malware	✓		✓
Vulnerabilities exploited	DCOM RPC Buffer Overflow			✓
	LSASS Buffer Overflow			✓
	IE Incorrect MIME Header		✓	
	IE XML Page Object Type Validation		✓	

Table 1: Main characteristics of Netsky, Beagle and Mytob

5. Conclusion

Although the result from this research does not provide exhaustive description about Beagle, Netsky and Mytob, it does serve to outline the general trend discovered based upon the analysis of worm variant development. This research also shows how

the functionality of the malware improved and evolved over time. The result derived from this research can be used as stepping stone for continuing this literature and also on other malware analysis. The findings will be useful to predict the evolution of new malware development behaviour. Therefore, if the new malware can be predicted, preventive action could be taken earlier and it the prediction were match, the damage which may caused by this new malware will be minimal as early preventive has been taken. In addition, the findings can be used to aid the quantitative part of malware analysis.

6. References

Microsoft (2004), "The Antivirus Defense-in-Depth Guide", http://www.microsoft.com/technet/security/topics/serversecurity/avdind_2.mspx, (Accessed: 1 September 2005)

SANS (2003), "Bots & Botnet: An Overview", http://www.sans.org/rr/whitepapers/malicious/1299.php, (Accessed: 1 September 2005)

Secunia (2005), "Chronological Virus List", http://secunia.com/chronological_virus_list/, (Accessed: 5 September 2005)

Symantec (2003), "Symantec Internet Threat Report Volume IV", http://enterprisesecurity.symantec.com/content.cfm?articleid=1539, (Accessed: 1 September 2005)

Symantec (2004a), "Symantec Internet Threat Report Volume V", http://enterprisesecurity.symantec.com/content.cfm?articleid=1539, (Accessed: 1 September 2005)

Symantec (2004b), "Symantec Internet Security Threat Report Volume VI", http://enterprisesecurity.symantec.com/content.cfm?articleid=1539, (Accessed: 1 September 2005)

Symantec (2005a), "Symantec Internet Security Threat Report Volume VII", http://enterprisesecurity.symantec.com/content.cfm?articleid=1539, (Accessed: 1 September 2005)

Symantec (2005b), "Search and Latest Virus Threats", http://securityresponse.symantec.com/avcenter/vinfodb.html, (Accessed: 4 September 2005)

Szor, P. (2002), "Blended Attack Exploits, Vulnerabilities and Buffer-Overflow Technique in Computer Viruses", http://peterszor.com/blended.pdf, (Accessed: 3 September 2005)

Szor, P. (2005), *The Art of Computer Virus Research and Defense*, Addison-Wesley, United States, ISBN: 0-321-39454-3.

Trend Micro (2004), "The Trend of Malware Today: Annual Virus Round-up and 2005 Forecast", http://www.trendmicro.com/en/security/white-papers/overview.htm#annualroundup2004, (Accessed: 1 September 2005)

Trend Micro (2005), "Outbreak Incidence and Prevailing Malware Trends: Q1 2005 Virus Roundup", http://www.trendmicro.com/en/security/white-papers/overview.htm#q12005virusroundup, (Accessed: 1 September 2005)