A Review of the Effectiveness of Malware Signature Databases against Metamorphic Malwares

Adesegun Oreoluwa A.
Department of Computer Science, Babcock Univeristy
adeseguno@babcock.edu.ng

Abstract

Known obfuscation techniques and other methods discovered by other researches such as Desai and Stamp (2010), Mohan & Hamlem(2012) have made detection of malware more difficult. This research is positioned to reviewing the current practices in the antivirus industry and determining if malware signature databases are adequate in detecting metamorphic malwares.

1.0 Introduction

The word Malware comes from two words, malicious and software, put together would be malicious software. A malicious software can be defined as an application or code that installs itself in stealthy way without permission to steal data, make services unavailable to legitimate users among others in the computer System. (Elhadi A.A.E. , Maroof M.A. & Barry B.I.A. , 2013; Jain & Bajaj, 2014; Mohan & Hamlem, 2012; Sharma & Sahay, 2014).

The origin of malwares can be traced back to Jon von Neumann’s studies in the 1940’s, when he was studying self-replicating mathematical model known as an automaton. (Kamarudin, Md Sharif, & Herawan, 2013). However the first malware on the windows platform was first discovered in the 1986. It was a virus developed by two brothers from Pakistan as a proof of concept that the pc platform was unsafe. (Kamarudin, Md Sharif, & Herawan, 2013; Milošević, 2013).

Since the 1986 when the two brother came up with their proof of concept, there has been an upsurge in the number of viruses to more 1,000,000 different computer virus strains (Kumar, Kumar, & Kumar, 2014). Some researchers have observed that at least one computer is infected with a malware every 39 seconds as found in (Mohan & Hamlem, 2012).

Today, there are several reason other than proof of concept why malwares continue to exist. Some of these reasons include but are not limited to the following:

- Financial gain both on the side of the cyber criminals and the anti-virus companies (Mohan & Hamlem, 2012)
- To spy on people’s internet activities (Jain & Bajaj, 2014)
- To deny legitimate user’s from accessing legitimate web services (Kumar, Kumar, & Kumar, 2014).

2.0 Current Antivirus Methodology

Malware analyst/reverse engineers work at antivirus companies surfing the web to determine applications with malicious intent and applications that have legitimate uses. Usually once a malicious application is found, a string is extracted from the malware to enable detection. The extracted string is called a signature. These signatures are then sent to the databases of the users of the various antivirus signature databases. Antivirus applications can then scan other applications and files to find matches against their databases and if found such files are quarantined or deleted.

3.0 Limitations of a Signature Database System

Polymorphic/Metamorphic Malware: These are malwares that change the structure of their code from the parent malware to child malware in successive iterations. This is achieved through the use of obfuscation techniques used by metamorphic engines. These kind of malwares are able to maintain the same semantics, that is, they act like the previous generation of the same malware but look different in terms of code structure (Berkat, 2011; Desai & Stamp, 2010). Metamorphic engines have several obfuscating or beclouding methods that help malwares evade detection from the signature database based scanning existing today. Some studies have shown some of these beclouding/obfuscation techniques are dead code insertion, register assignment, subroutine reordering and instruction substitution (You & Yim, 2010).

Hence, the problem is that metamorphic malwares have metamorphic engines that use varying obfuscation techniques to change their code structure from one form to another making detection difficult by antivirus software.

Time: There is a time gap between when a malware is created and when it is categorised as a malicious application. This implies that during this time gap the malware is free to operate on infected hosts. There could also be a delay between when a malware signature database is updated leading to a vulnerable system.

4.0 Possible Solution/Future Study

More work has to be done in the area of automation of antivirus software in detecting malicious software by their behavioural characteristic than continually updating the virus signature database, as this has proven ineffective to the detection of metamorphic malwares. If antivirus software are able to adequately analyse the behaviour of a program and act appropriately then there wouldn't be any lag in time as there would be no need for a malware signature database.

References

Desai, P., & Stamp, m. (2010). A Highly Metamorphic Virus Generator. *Int. J. Multimedia Intelligence and Security*, 402-427.

Elhadi A.A.E. , Maroof M.A. , & Barry B.I.A. . (2013). Improving the Detection of Malware Behaviour Using simplified Data Dependent API call Graph. *International Journal of Security and its Application 7(5) 29-42, 7*(5), 29-42.

Jain, M., & Bajaj, P. (2014). Techniques in Detection and Analyzing Malware Executables: A Review. *International Journal of Computer Science and Mobile Computing, 3*(5), 930-933.

Kamarudin, I., Md Sharif, A., & Herawan, T. (2013). On Analysis and Effectiveness of Signature Based in Detecting Metamorphic Virus. *International Journal of Security and Its Applications, 7*(4), 375-384.

Kumar, D., Kumar, N., & Kumar, A. (2014). Computer Viruses and Challenges for Anti-virus Industry. *International Journal of Engineering and Computer Science, 3*(2), 3869-3872.

Milošević, N. (2013). History of Malware. *Computer Security*, 1-11.

Mohan, V., & Hamlem, K. (2012). Frankenstein: Stiching Malware from Benign Binaries. *In WOOT (*, 77-84.

Sharma, A., & Sahay, S. (2014). Evolution and Detection of Polymorphic and Metamorphic Malwares: A Survey. *International Journal of Computer Applications, 90*(2), 7-11.

You, I., & Yim, K. (2010). Malware Obfuscation Techniques: A Brief Survey. *International Conference on Broadband, Wireless Communication and Applications* (pp. 297-300). IEEE Computer Society.