IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2320-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 9, Issue 12, December 2020

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.512

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

||Volume 9, Issue 12, December 2020||

Study on Detection and Analysis of Metamorphic Malware

Sanskruti Sarfare

Academic Research Student, Department of Information Technology, B.K.Birla College of Arts, Science and Commerce (Autonomous) Kalyan, Maharashtra, India

ABSTRACT: Malware is becoming more aggressive and new techniques are evolving to allow malicious code to avoid detection by antivirus. The habits of downloading unknown files from the internet may be the key reason for infecting computer systems with malware such as metamorphic malware and others. The main feature that distinguishes the metamorphic viruses from other malware is the usage of the code obfuscation techniques. Metamorphic malware is a kind of virus that changes its form at each infection. The purpose of this research paper is to have a look at the detection methods and analysis of Metamorphic Malware.

KEYWORDS: Metamorphic Malware, Opcode, Obfuscation, Static analysis, Behavior-based detection.

I. INTRODUCTION

Malware is software designed to damage computers. It is the root cause of many Internet security problems. Malware can steal sensitive data or infect files one after another. It is very difficult to detect because it changes its form and can also recompile itself by using the same compiler making it more difficult to identify. These malware use code obfuscation to change and have many forms. Code obfuscation technique is done by garbage instruction, code reordering, changing the syntax, etc. making it more complex and difficult to analyze. Metamorphic malware detection can also be done based on Obfuscation features analysis. There are three major techniques of detection: Heuristic, signature-based, and behavior-based. Signature-based malware detection is the most common method and efficient to detect known viruses. Behavior-based detection can evaluate an object before it can perform any action. Behavioral or heuristic methods are more effective than signature-based methods. Heuristic analysis can employ various techniques that include decompiling or suspecting a program and examining its source code. The Hidden Markov technique is the most efficient technique for malware detection. Hidden Markov Model is an efficient model to detect and improve the classification of malware. As malware detectors use more advanced detection techniques and malware writers use better evasion techniques to avoid detection. Malware analysis is important as in today's time it is very difficult to detect or analyze the malware to have a detailed study on it so that we can prevent systems, devices from malware attacks. Static analysis methods also include deobfuscating and disassembling which is basically restoring or cleaning the program which has been already obfuscated. This paper consists of the various types of detection methods and also analysis techniques but mainly emphasis some of them based on the hypothesis below.

II.OBJECTIVES

- ➢ To find out that behavior-based malware detection is more accurate compared to signature-based malware detection.
- ➢ To find out static analysis is more effective than dynamic analysis.

The above objectives would be fulfilled by proving the following hypothesis through survey analysis:

H1)" Behavior-based malware detection evaluates an object based on its suspicious or abnormal actions before it can execute."
H2)" Static malware analysis analyses the malware without going through each phase or executing it by observing its behavior."

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

||Volume 9, Issue 12, December 2020||

III. LITERATURE REVIEW

In a study published by Choudhary,S.P, et al [1] in 2015, proposed how to analyze dynamic reports and how we can prepare models which would help make detection decisions of Metamorphic Malware. In a study published by Stamp, et al [2] in 2011, contains elementary code obfuscation techniques used in Metamorphic generators and discusses the use of HMMs for Metamorphic detection. In a study published by Pomorova,O. , et al [3] in 2016 , presents a new technique for Metamorphic virus detection using modified emulators, placed in the hosts of the network. The proposed technique provides the classification of the Metamorphic virus in classes with the usage of fuzzy logic. In a study published by Markowsky, G. , et al [4] in 2018 , presents an approach for the Metamorphic viruses detection based on its obfuscation features analysis. In a study published by Canfora, G. , et al [5] in 2014 , contains a technique for detecting Metamorphic viruses based on identifying specific features of the assembly code, such as the instructions that change the contents of the registers. In a study published by Savenko, O. , et al [6] in 2017 , presents a new technique for unknown Metamorphic viruses detection .It is based on analysis of the potentially suspicious behaviour of the programs on the host. In a study published by Sasidharan, S., et al [7] in 2018 , describes the advantages of the stochastic modeling method to detect Metamorphic malware which evades the normal detection method. In a study published by Irshad, M., et al [8] in 2018 , the paper was analyzed based on their approach,limitation,empirical evidence ,and key parameters such as dataset, detection rate, and false-positive rate. In a study published by Sahay , S., et al [9] in 2019 , detects the malware which quietly takes confidential data from computational devices and cripples the infrastructure. In a study published by Asari , E., et al [10] in 2020 , explains the process of modelling and training an advanced PHMM using sequences obtained from the extraction of each malware and MSA production. In a study published by Han , W., et a [11] in 2019 , accurate identification of disguised malware, a malware detection framework is proposed named "MalInsight".Due to the continuous upgradation of malware , a systematic profiling technique is introduced. In a study published by Alam , S., et al [12] in 2015 , contains the change in obfuscations by the malware writers to hide detection. To overcome this new framework named "Metamorphic Malware Analysis and Real-Time Detection(MARD)" is introduced. In a study published by Mohamed , G. , et al [13] in 2017 , contains signature and behavior-based representation techniques for detecting Metamorphic malware.It is a combination of signature and behavior-based techniques. In a study published by Alqurashi , S., et al [14] in 2016 , presents the hidden Markov Model technique as a malware detection tool as it is the most efficient way to detect Metamorphic malware. In a study published by Irshad , M., et al [15] in 2018 , presents detailed knowledge about the detection of Metamorphic malware. In a study published by Ling , Y., et al [16] in 2017 , states the types of malware, detection and obfuscation techniques and malware analysis. In a study published by Qu, Y., et al [17] in 2013 , states that behavioural based signature techniques are more efficient than other signature techniques. In a study published by Galal, H., et al [18] in 2015 , a behavior based features model is proposed to describe the malicious actions by the malware during the runtime. In an article published by John Cloonan , et al [19] in 2019 , states the difference between signature based and behavior based malware detection methods. In a study published by Mujumdar, A., et al [20] in 2013 , analysis of signature based and behavior based anti-malware approaches is done. In a study published by Bhojani, N., et al [21] in 2018 , consists of the information about static analysis and also its limitations with some trends in malware and de-obfuscating malware.

IV. METHODOLOGY

A.*Participants*

A study was conducted to know people's opinions and experiences with malware(viruses) in their devices. The study was conducted through Google forms that were circulated randomly because it didn't need any specific population rather a population that uses devices like mobiles, laptops, desktops, etc on a daily basis. A total population of 30 were involved in this survey (19 females and 11 males). The survey was conducted within the limits of Mumbai.

B.*Material*

The source of the conducted survey is Google Forms. The population of the conducted survey mostly consists of students who use devices like mobiles, laptops, desktops, etc more often and are used to it or maybe experienced any malware attack sometime.

C.*Procedure*

A Google form was chosen as a medium to conduct a survey by using a simple sampling technique that includes the population for a fair opportunity to test the hypothesis. The Google form was circulated for a week or two which resulted in 30 responses.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

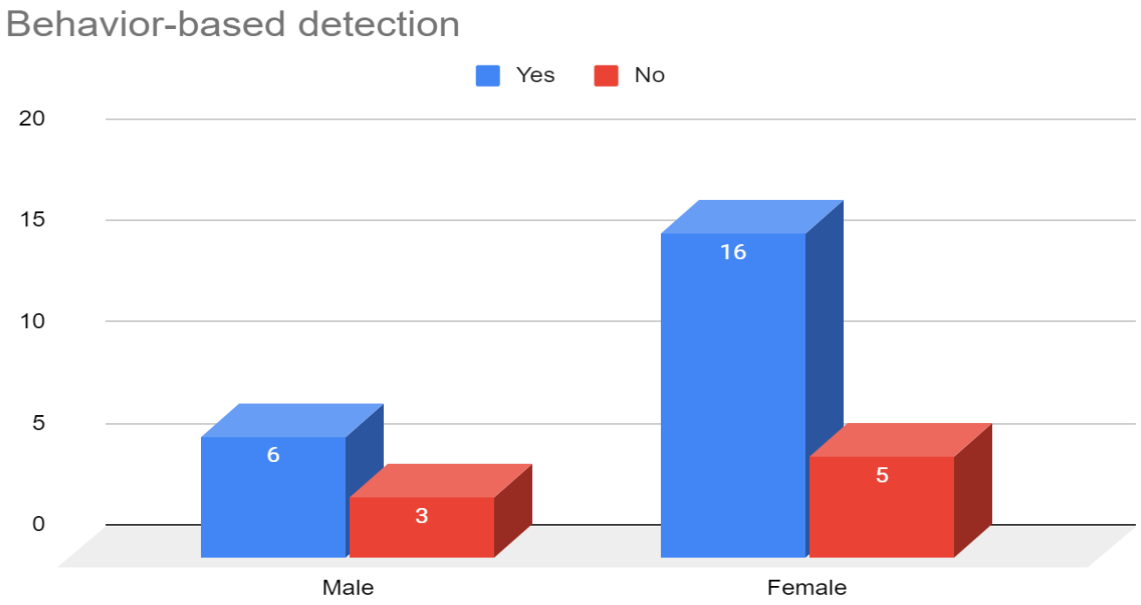
||Volume 9, Issue 12, December 2020||

V. EXPERIMENT

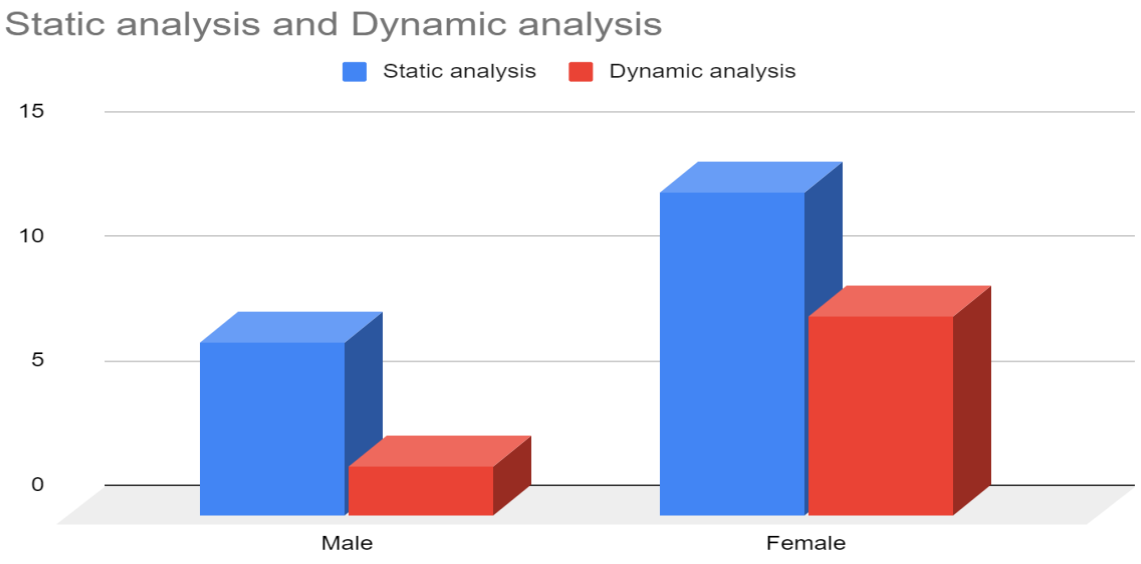
The test scores of the hypothesis-based survey were taken into consideration by performing the Chi-Square Test with a confidence level of 95 percent. The questions that were appropriate were asked in the survey which was supposed to prove the hypothesis stated above. A factor to distinguish the survey questions is gender to notice the difference in the response values to perform the test.

The survey questions which were used to prove the hypothesis are : (1) Would you prefer your software to detect the abnormal behavior of the malware before it can execute? Provided the calculated value of 0.1909.

(2) Which analysis would you choose? Options were:(a) Static analysis analyzes the malware without even executing it.(b) Dynamic analysis analyzes the malware after running the code. Provided the calculated value of 0.6664. These values are calculated by applying formulas for observed value and expected value. The values were checked with a tabulated value of 3.84(having a confidence level of 95 percent).



Fig(1): The comparison of test values of Behavior-based detection.



Fig(2): The comparison of Static analysis and Dynamic analysis.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

||Volume 9, Issue 12, December 2020||

VI. RESULTS

The test scores of the survey which was analyzed on the basis of certain parameters resulted in an inclination towards behavior-based detection rather than any other detection methods as it detects the malware before its execution hence proving H1.
The test scores of the survey in the case of H2 proved that static analysis of malware is more convenient than dynamic analysis because static malware analysis analyses the malware without going through each phase or executing it by observing its behavior.

VII. LIMITATIONS AND FUTURE SCOPE

The survey conducted had some limitations of population and time. If provided more time and specific population would surely help in more accurate results in analyzing the survey and a better understanding of the results. In this survey conducted the population was asked about the advancements they wish to have from anti-malware software to prevent malware attacks. Some of them are proper malware detection in first priority, warnings must be provided, quick scanning and user-friendly software, and immediate actions on malware. The software should be performing behavior-based detection so that it can satisfy the needs of the population by immediate scanning and taking actions on the same and also static analysis so that it would analyze it by its behavior without going through each phase resulting in quick output.

VIII. CONCLUSION

Behavior-based detection in malware detection software would be beneficial because it detects the malware by its behavior before it executes, preventing the malware attack on the device. It protects some new and unimagined types of malware, identifies what malware is doing in a specific file or folder when opened. This technology first identifies some suspicious behavior or malicious activity which when observed together let us know about the damaging behavior.
Static analysis is a method in which the analysis is done without running the code and on the other hand in Dynamic analysis the code needs to be run to study the components of malware. Moreover, basic static analysis is done by hashing or detecting the packed or obfuscated programs. Deobfuscating or disassembling is also done to analyze the malware. Static malware analysis is more capable of providing complete information about the characteristics of the malware.

IX. ACKNOWLEDGEMENT

Special thanks and gratitude to Prof. Swapna Augustine Nikale, Department of Information Technology, B.K.Birla College of Arts, Science and Commerce, Kalyan for guiding and providing needful help in this work of paper.

REFERENCES

[1] Choudhary, S. P., & Vidyarthi, M. D. (2015). A Simple Method for Detection of Metamorphic Malware using Dynamic Analysis and Text Mining. *Procedia Computer Science*, *54*, 265–270. https://doi.org/10.1016/j.procs.2015.06.031
[2] Stamp, Mark & Venkatachalam, S.. (2011). Detecting Undetectable Metamorphic Viruses.
[3] Pomorova, O. (2016). Metamorphic Viruses Detection Technique Based on the Modified Emulators. *Metamorphic Viruses Detection Technique Based on the Modified Emulators*, 376–383. http://ceur-ws.org/Vol-1614/paper_106.pdf
[4] Markowsky, G. (2018). The Technique for Metamorphic Viruses' Detection Based on its Obfuscation Features Analysis. *The Technique for Metamorphic Viruses' Detection Based on Its Obfuscation Features Analysis*, 1–8. http://ceur-ws.org/Vol-2104/paper_231.pdf
[5] Canfora, G., Mercaldo, F., Visaggio, C. A., & Di Notte, P. (2014). Metamorphic Malware Detection Using Code Metrics. *Information Security Journal: A Global Perspective*, *23*(3), 57–67. https://doi.org/10.1080/19393555.2014.931487
[6] Savenko, O., Lysenko, S., Nicheporuk, A., & Savenko, B. (2017). Approach for the unknown metamorphic virus detection. 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). doi:10.1109/idaacs.2017.8095052

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.512|

||Volume 9, Issue 12, December 2020||

[7] Sasidharan, S. kumar, & Thomas, C. (2018). A Survey on Metamorphic Malware Detection based on Hidden Markov Model. 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI).
[8] Irshad, M., Khateeb, H. M. A., Mansour, A., Ashawa, M., & Hamisu, M. (2018). *Effective methods to detect metamorphic malware: a systematic review. International Journal of Electronic Security and Digital Forensics, 10(2), 138.* doi:10.1504/ijesdf.2018.090948
[9] Sahay, S. K., Sharma, A., & Rathore, H. (2019). *Evolution of Malware and Its Detection Techniques. Advances in Intelligent Systems and Computing, 139–150.* doi:10.1007/978-981-13-7166-0_14
[10] Ansari, E. (2020). An advanced profile hidden Markov model for malware detection. *An Advanced Profile Hidden Markov Model for Malware Detection*, 759–778. https://doi.org/10.3233/IDA-194639
[11] Han, W., Xue, J., Wang, Y., Liu, Z., & Kong, Z. (2019). MalInsight: A systematic profiling based malware detection framework. Journal of Network and Computer Applications, 125, 236–250. https://doi.org/10.1016/j.jnca.2018.10.022
[12] Alam, S., Horspool, R. N., Traore, I., & Sogukpinar, I. (2015). A framework for metamorphic malware analysis and real-time detection. Computers & Security, 48, 212–233. doi:10.1016/j.cose.2014.10.011
[13] Mohamed, G. A. N., & Ithnin, N. B. (2017). SBRT: API Signature Behaviour Based Representation Technique for Improving Metamorphic Malware Detection. Lecture Notes on Data Engineering and Communications Technologies, 767–777. doi:10.1007/978-3-319-59427-9_79
[14] Alqurashi, S., & Batarfi, O. (2016). A Comparison of Malware Detection Techniques Based on Hidden Markov Model. *Journal of Information Security*, *07*(03), 215–223. https://doi.org/10.4236/jis.2016.73017
[15] Irshad, Mustafa & Al-Khateeb, Haider & Mansour, Ali & Ashawa, Moses & Hamisu, Muhammad. (2018). Effective methods to detect metamorphic malware: A systematic review. International Journal of Electronic Security and Digital Forensics. 10. 138. 10.1504/IJESDF.2018.090948.
[16] Ling, Y. T., & Sani, N. F. M. (2017). Short Review on Metamorphic Malware Detection in Hidden Markov Models. *International Journal of Advanced Research in Computer Science and Software Engineering*, *7*(2), 62–69. https://doi.org/10.23956/ijarcsse/v7i2/01218
[17] Qu, Yanzhen & Hughes, Kelly. (2013). Detecting metamorphic malware by using behavior-based aggregated signature. 2013 World Congress on Internet Security, WorldCIS 2013. 13-18. 10.1109/WorldCIS.2013.6751010.
[18] Galal, Hisham. (2015). Behavior-based features model for malware detection. Journal of Computer Virology and Hacking Techniques. 12.10.1007/s11416-015-0244-0.https://www.cyberdefensemagazine.com/advanced-malware-detection/
[19] Mujumdar, A., Masiwal, G., & Meshram, D. B. B. (2013). Analysis of Signature-Based and Behavior-Based Anti-Malware Approaches. *Analysis of Signature-Based and Behavior-Based Anti-Malware Approaches*, *2*(6), 2037–2039.
[20] http://ijarcet.org/wp-content/uploads/VOLUME-2-ISSUE-6-2037-2039.pdf
[21] Bhojani, N. (2018). Malware Analysis. *Malware Analysis*, 1–4. https://www.researchgate.net/profile/Nirav_Bhojani/publication/267777154_Malware_Analysis/links/545a59e10cf2c46f66427285.pdf

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.512

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details