A human-based approach for detecting modern malware

Asanga Samaranayake
Faculty of Graduate Studies and Research
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
asangais@outlook.com

***Abstract*—Malware which was developed only to slowdown a PC or hide a folder has become a thing of the past. Modern malware can collapse huge networks and steal huge amount of secret data within minutes of operation. The journey travelled by the malware developers were also travelled by the malware analysts. Successful detection of a malware has resulted in effectively diffusing its operation in most of the cases. It is always a never-ending battle between malware developers and malware analysts.**

To conceal the inner operation of a malware, the developers use many techniques. Many of these can be broken down by effective detection using traditional techniques. However, knowing this fact, modern malware uses previously unseen techniques, so that the analysts will either be incapable or slower in identifying the malware, thereby the malware itself has more time to cause its intended damage.

The above-mentioned new trends require new techniques to identify and detect these modern malware. This paper starts by reviewing existing literature to understand the current status of these developments. Thereafter it identifies the limitations of currently available solutions in detecting modern malware. Finally, the paper suggests a novel human-based approach as a countermeasure against rapidly evolving modern malware.

***Index Terms*—malware detection, modern malware, human-based**

I. Introduction

This paper reviews existing literature to identify the reasons why current techniques for malware detection are limited in capability against modern malware and thereafter presents currently available solutions. Thereafter a more proactive human based approach is suggested as a countermeasure against rapidly evolving modern malware.

II. RESEARCH STATEMENT / OBJECTIVES

This paper reviews the existing literature related to modern malware detection and proposes a human based approach as a suitable countermeasure.

III. LITERATURE REVIEW

When conducting the literature review, 8 published articles in IEEE Xplore database were used all of which are recent articles published from 2016 to 2021 (including both years).

A. Why current techniques are inadequate

When mentioned about malware detection, signature-based systems come to the mind first. While they were effective during past times, with the advent of obfuscation and encryption techniques, malware developers have defeated them [1][2]. Traditional signature-based systems and even behaviour-based systems are not effective against latest trends such as zero-day malware. New malware mutates and completely changes their appearance using the mutation engine which is in the malware body [3].

As per Cisco 2017 Annual Cybersecurity report, 95% of analyzed malware were below 24 hours of age [1]. Therefore malware evolve faster than research [1]. When malware code is published, numerous variants of malware are created with different behaviours. This makes detecting each variant difficult. Founders of Norton Antivirus, a popular antivirus solution, has mentioned that their AV can only detect and prevent 45% of malware attacks on a system [3].

Some of the common dual use tools such as PS, Macros are used as fileless malware, effectively reducing their detection [4]. All file based antivirus techniques fail in detecting fileless malware as no file is saved in the file system [4]. At start the fileless malware was triggered by a file-based malware but it will be deleted later to ensure persistence in the system [4]. Despite having no file stored on the victim system, the power of fileless malware should not be underestimated as they infect first and open up space for file-based malware attacks [5].

There is a challenge in selecting which detection technique to select. Static detection is for previously seen malware and dynamic detection is for previously unseen malware such polymorphic malware. Usually, a combination of these will be used which is a called hybrid technique. When it comes to IDS, selection is a tough tradeoff between static (low detection rate) and dynamic (low timeliness) [6].

Another challenge is application behaviour. For example, new malware gives expected result of the program while malware operations are carried out underlying. Therefore users do not suspect they are infected as the program gives out the standard output to them [1].

Antivirus software are complex, requires elevated or administrator privileges and are resource intensive [3]. Even tolerating the above requirements, users expect that their wellrespected antivirus software will effectively safeguard them from malware attacks. However, it was found that detection rate of two popular antivirus programs, Kaspersky and McAfee only achieved 71% and 67% respectively [3]. Risk is further aggravated as users may not run long resource intensive scans as it slows the entire system down which prevents them from doing day-to-day work.

B. What are the solutions currently available

Use of cloud antivirus found to be less resource intensive than their computer-installed counterparts [3]. A novel solution was identified by using a cloud engine and a lightweight local agent. It uses publicly available APIs of 15 different AVs, CTI (Cyber Threat Intelligence) and OpenIOC (Open Indicators of Compromise) format [3]. An advantage of this approach is that the OpenIOC format can be used to enhance detection of the device (IDS, IPS, Firewall etc.) and the outputs and logs of these devices can be used to enhance the OpenIOC documents creating a continuous improvement process cycle [3]. Disadvantages of this approach include limitation to Windows only environments and ability to only detect executed files, not the stored files. Due to this a host which is used as a carrier to transport malware without executing it will go unnoticed with this technique [3].

An interesting malware detection method is vision-based analysis where malware files are converted into grayscale images [2]. Advantages of this approach include the ability to detect modified malware, no in-depth analysis is required, no disassembly or execution is required and possessing a 98% detection rate of novel, zero-day and even obfuscated malware [2].

Another study proposes malware detection and family classification using dynamic analysis to achieve a decent detection rate. During this study only 4 out of 1650 tested malware went undetected for a detection rate of above 99% which is a key advantage of the technique [7]. Disadvantages of the technique include requiring a more uniform sample set and requiring all malware families are learnt, which is practically difficult. Also due to this difficulty, the technique operates in limited functionality and cannot be used as a standalone solution and need to depend on other solutions such as signatures to arrive at a solid prediction [7]. In another study, hardware assisted malware detection using explainable AI was used [8].

In another study, a hybrid tool was used which has three buckets to categorize files entering a network. IDS categorizes each file as very benign, very malicious or needs further analysis. Very benign files are allowed to pass while very malicious files are blocked. Only remaining files (under needs further analysis category) are further analyzed. This is a considerable improvement over traditional methods where all files are either statically or dynamically analyzed. During this method, researchers could get 9x reduction of malware entering a network, 27.5x reduction in computer resource utilization and 10.5x reduction in alert generation [6].

IV. PROPOSED SOLUTION

After analyzing results of aforementioned studies, author could come up with below conclusions.
A user:

- cannot depend on the AV anymore (due to mutation engine operation, numerous variants of viruses in short period of time, low detection rates etc.)
- cannot trust an app just because it provides the expected output (infected program can still provide expected output)
- cannot trust an app based on its standard usage (malware uses standard apps such as PS, Macros etc.)
- cannot rely on file analysis anymore (many of new variants are fileless with nothing on file system)

While the above conclusions seem somewhat alarming, not all hope is lost. Out of the techniques mentioned, vision analysis is most effective against file-based malware which has 98% detection rate, minimal processing and even detects latest and previously unseen malware. However, this is ineffective against systems which are already infected with fileless malware. The reason for this is the fileless malware is originated by a file-based malware at initial stage and then the file is deleted from the system for persistence. Now the only available malware in the system is the fileless instance, hence file-based vision analysis technique may fail to detect this.

Summarizing all above, author identifies that users are either a) waiting for someone else to make a solution to find malware (e.g. – AV) or b) waiting for research to be conducted and they are provided with findings to effectively identify the malware. Author identifies that both these approaches are reactive in nature, taking a longer time to identify a threat and thereby causing users to be exposed to malware for a longer time. All these are suboptimal results and author suggests a proactive, more human based approach towards malware detection.

Author believes that it is time for users to start thinking about malware detection on their own. Every malware uses a weakness in humans (be it system administrator, end user, application developer or web developer etc.) such as lack of skills or knowledge to get into the systems in an initial stage. Even skilled, knowledgeable and experienced users fall victim to malware due to lack of attention on that given moment. In reality, malware utilizes a moment in which we are not fully present and focused as the entry point to our systems. If we could be more mindful with all our actions then it will be possible for us to prevent many malware attacks in future. This is a more human based approach than a technical solution which might take some time for users to adapt. However, in long term, this will surpass even the most expensive high-end technical solution. Let us take malware detection to our own hands from now. Knowing the fact that nobody or nothing is there to fully protect us and we are on our own will be a good starting point.

V. RESEARCH GAPS AND FUTURE RESEARCH

There are limitations in currently available solutions for modern malware detection. For example, a new and efficient defense system was proposed but it is limited in functionality for a single system and cannot be operating in an enterprise network [3]. It also uses a continuous improvement process for threat intelligence but limited to Windows environment and to executed files only [3].

In another study, effective technique for detecting and categorizing malware was discussed its practical implementation is difficult and not possible to be used as a standalone solution [7].

Vision based analysis a highly effective technique for file based malware but may not be effective against fileless malware [2]. Similarly, hybrid tool which uses both static and dynamic analyses were found to be effective against file based malware but again not effective against fileless malware [6].

Therefore, the current techniques are found to be limited to either an OS such as Windows, a single system environment or inability to cover entire spectrum of malware. Further research has to be conducted in finding a more versatile solution which can be universally applied.

VI. Conclusion

Traditional malware detection techniques have been protecting users from malware for many years. However, they are no longer able to keep up with the rapid rate of modern malware development. Most of the detection techniques developed against novel malware are either limited in capabilities or cannot be universally applied.

Most of the techniques are developed assuming that the malware is resident in the system as a file but which is not always the case with novel variants such as fileless malware. Malware gets developed at a faster rate than antivirus solutions and research evolve. Users are in a critical position where they no longer can depend on their antivirus solution, cannot trust an app based on its expected output or its standard usage. This calls for a more proactive approach in detecting modern malware.

Technical solutions are developed to find malware based on their specific composition or behaviour, therefore cannot be used universally against all variants. However, malware enters or gets initial foothold to a system by exploiting a human mistake or weakness such as lack of knowledge, skills or experience. Even skilled and knowledgeable users fall victim for malware because of lack of attention in the moment. In this context author suggest users maintain focus on the present moment when using systems thereby enabling them to fully utilize their knowledge, skills and experience in detecting these entry points of malware. While it may take some time to adapt to this approach, it will be a universal solution and its long-term results will surpass even the most advanced technical solution.

Acknowledgments

Author would like to thank Dr. Lakmal Rupasinghe for his valuable guidance and advice.

References

[1] O. P. Samantray, S. N. Tripathy, and S. K. Das, “A study to understand malware behavior through malware analysis,” 2019 IEEE Int. Conf. Syst. Comput. Autom. Networking, ICSCAN 2019, 2019, doi: 10.1109/ICSCAN.2019.8878680.

[2] S. A. Roseline, S. Geetha, S. Kadry, and Y. Nam, “Intelligent Vision-Based Malware Detection and Classification Using Deep Random Forest Paradigm,” IEEE Access, vol. 8, pp. 206303–206324, 2020, doi: 10.1109/ACCESS.2020.3036491.

[3] Q. K. A. Mirza, G. Mohi-Ud-Din, and I. Awan, “A cloud-based energy efficient system for enhancing the detection and prevention of modern malware,” in Proceedings - International Conference on Advanced Information Networking and Applications, AINA, May 2016, vol. 2016-May, pp. 754–761, doi: 10.1109/AINA.2016.133.

[4] A. Afreen, M. Aslam, and S. Ahmed, “Analysis of Fileless Malware and its Evasive Behavior,” Oct. 2020, doi: 10.1109/ICCWS48432.2020.9292376.

[5] L. Caviglione et al., “Tight Arms Race: Overview of Current Malware Threats and Trends in Their Detection,” IEEE Access, vol. 9, pp. 5371–5396, 2021, doi: 10.1109/ACCESS.2020.3048319.

[6] D. Kim, D. Mirsky, A. Majlesi-Kupaei, and R. Barua, “A Hybrid Static Tool to Increase the Usability and Scalability of Dynamic Detection of Malware,” in MALWARE 2018 - Proceedings of the 2018 13th International Conference on Malicious and Unwanted Software, Mar. 2019, pp. 115–123, doi: 10.1109/MALWARE.2018.8659373.

[7] S. S. Hansen, T. M. T. Larsen, M. Stevanovic, and J. M. Pedersen, “An approach for detection and family classification of malware based on behavioral analysis,” Mar. 2016, doi: 10.1109/ICCNC.2016.7440587.

[8] Z. Pan, J. Sheldon, and P. Mishra, “Hardware-Assisted Malware Detection using Explainable Machine Learning,” in Proceedings - IEEE International Conference on Computer Design: VLSI in Computers and Processors, Oct. 2020, vol. 2020-Octob, pp. 663–666, doi: 10.1109/ICCD50377.2020.00113.