A Complete Study on Malware Types and Detecting Ransomware Using API Calls

Conference Paper · September 2021

DOI: 10.1109/ICRITO51393.2021.9596085

5 authors, including:

Gagandeep Kaur
Lovely Professional University
14 PUBLICATIONS **208** CITATIONS
SEE PROFILE

Anshu Vashisth
Lovely Professional University
19 PUBLICATIONS **208** CITATIONS
SEE PROFILE

Rohith Cheerala
Lovely Professional University
3 PUBLICATIONS **35** CITATIONS
SEE PROFILE

2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)
Amity University, Noida, India. Sep 3-4, 2021

2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) | 978-1-6654-1703-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICRITO51393.2021.9596085

A Complete Study on Malware Types and Detecting Ransomware Using API Calls

Nishant Yadav
School of computer science engineering
Lovely professional university,
Phagwara, Punjab.
ynishant435@gmail.com

Gagandeep Kaur*
School of computer science engineering
Lovely professional university
Phagwara, Punjab.
gagandeep.23625@lpu.co.in

Sukhwinder Kaur
School of computer science engineering
Lovely professional university
Phagwara, Punjab.
sukhvinderbudhrayan111@gmail.com

Anshu Vashisth
School of computer science engineering
Lovely professional university
Phagwara, Punjab.
anshu.23500@lpu.co.in

Cheerala Rohith
School of computer science engineering
Lovely professional university
Phagwara, Punjab.
cheerala.11606529@lpu.in

***Abstract*— Due to advancement in technology, cybersecurity attacks are increasing day by day. So, to secure the data and information, Malware analysis is introduced as a new methodology which is the process of understanding the behavior and purpose of a suspicious file or URL. The output of the analysis helps in the detection and mitigation of the potential threat. This novel solution is based on cybersecurity, which can be used in window applications, File scans, URL scans, and link scans. This paper describes different types of malwares and a solution for securing the data by detecting ransomware using API calls.**

Keywords--- Malware, Detection of malware, Cyber-security, Data security, Pre-detection of malware.

I. INTRODUCTION

There is no doubt about the origin of malware that happens because of malicious code. It enters the system hoping to change most of the system's functions or in some cases to bypass all network of your device". So, in simple words, we can say that any code that has aimed to harm device function or activity steals information and data. This type of code is usually created by the one who has wrong intentions or is not ethical about their work in cybersecurity, like black hat hackers. Most of the time, these people are just looking for ransom, either by spreading the malware themselves or selling it to the highest bidder on the Dark Web. However, the first internet worm was written for experiments or pranks. But in today's world, there are other reasons for creating malware (Ex: people are nowadays using it to harm others or starting a Cyberwarfare between two nations) [20]. "Malware is derived from malicious software. It is an instance of malicious code to subvert the system's function and has potential to harm a computer or network."[1]

Thousands of new malware are turning up very soon, G Data and King Soft Laboratory report said. There are various types of this form like Spywares, rootkits, worms or backdoors and botnet, and other types of software with unwanted behavior [1,2]. Global level to steal information to other countries and disrupt their business and growth. They replicate themselves in various ways and join the system. Both via different media and the most common way of getting Downloaded are programmed into the device [18].

Various malware has been further classified into several types, depending on different parameters and their characteristics. In this research paper, a kind of classification, depending on its usage of networks and the internet, is made. The usage of the internet and network depicts the environment of execution of malware. That classification is based on the fact that the application of networks and the internet requires a unique software approach. Like IDS, SQL attack prevention, detection of LAN worms, malware classification in real-time. There are various malware classes:

A. Network-based Malware

Spyware is malicious software that is secretly installed on a user's machine to collect user information. Intentionally, even reputable software vendors such as Microsoft and Google gather data from their users using spyware.

B. Ordinary Malware

Virus is any software code or any program that has capability to replicate itself, during infection, into any other application software or a document.

II. MALWARE TYPES

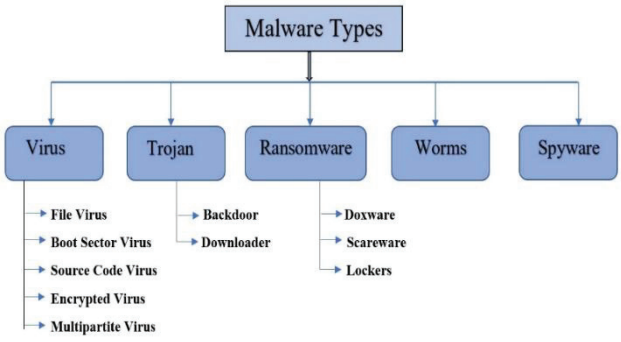

Fig. 1. Malware Types

A. Virus

It is the most common type of malware; it simply attaches itself to the device's files and corrupts them. Exe is the standard file format.

1) File Virus

This type of virus attaches itself at the end of file. It changes the start of the file so that the control jumps back to the source code and the control jumps back to the main program after execution.

2) Boot Sector Virus

The system boot sector runs every time the system is booted and loaded before the operating system is loaded. Other bootable media, like floppy disks, are infected. These are also known as memory viruses, as the file system is not infected.

3) Source Code Virus

It looks for and modifies the source code to include the virus and to help spread it.

4) Encrypted Virus

To avoid antivirus detection, an encrypted form of this type of virus exists. It carries with it a decryption algorithm. So, the virus decrypts first, then execute.

5) Multipartite Virus

This form of virus, including the boot field, memory, and data, can infect multiple sections of a machine. This makes identification and containment difficult. Now, suppose we talk about up to which level these viruses can affect our device. In that case, it interferes with the usual functioning of the connected computing device. Interrupts the use of the network device, modifies system settings for setup, Interrupts working of computer networks. Confidential Knowledge Destructs.

B. Trojan

This kind of malware is disguised as legitimate software or is concealed in legitimate, tampered software. This type tends to act differently and distract users from creating a backdoor to give way to other malicious codes [19]. It can compromise our security in many ways, like detecting data, copying data, deleting, or modifying data. It also disturbs the performance of the system. There are different types of Malwares like Backdoor Trojan, Downloader Trojan, Fake AV Trojan.

1) Backdoor Trojan

On your machine, this Trojan will build a "backdoor." It allows your device to be accessed by an intruder and managed. Your data can be downloaded and stolen by a third party. Or they can upload more malware to your computer.

2) Downloader Trojan

Targets infected machine with this Trojan. It downloads new versions of malicious programs and installs them. Trojans and adware may provide these. Fake AV Trojan: This Trojan works like antivirus software, but to detect and delete threats, whether they are real or false, it needs money from you.

C. Ransomware

This type is the same as the real-life scenario of asking for ransom against any valuable thing. In this case, this piece of code locks all the device files and displays a ransom message or by any other means. So, until and unless a ransom is not paid, you will not get your data back. It is used in money laundering frauds via emails, instant messages [3]. It has the power to lock a computer screen or password-encrypt important, predetermined data. At the same time, the idea behind ransomware is simple. Fighting it can be more complicated if you suffer a malicious ransomware attack. And you may be unable to regain access to your data or device if the attackers do not give you the decryption key.

1) Doxware

Doxware, usually referred to as leakware or extortion ware, if you do not pay the ransom, threatens to post the stolen information online. People store confidential data and personal images on their devices. Understandably, some victims have to pay when their files and documents have been hijacked.

2) Scareware:

Scareware is kind of fake software that pretends as an antivirus. Scareware also claims to have discovered problems on your computer, requesting cash to repair the issues.

Lock your computer with certain kinds of scareware. Others flood the computer with bothersome reminders and pop-up messages.

3) Lockers

To fully lock you out of your computer or computers, Locker-ransomware is known to corrupt your operating system. Making it difficult to access any of your files or applications. Most frequently, this type of ransomware is Android-based.

D. Adware

The other side of this concept is that some adverting software also adds some malicious code that can breach your security and make an entrance for other malware. To target ads that seem targeted to your interests, adware uses the browser to gather your web browsing history. Adware, for instance, barrages you with pop-up advertising that can make your Internet experience substantially slower and more labor-intensive. There are many ways by which we can tell that the system is having adware like crashing system, slow internet, Bombarding with ads and many more.

E. Botnet

Attackers control this type, and it is a whole infrastructure of infected devices. From one central point, the attacking party may order every device on its botnet to conduct a coordinated criminal activity simultaneously. The size of a botnet allows the attacker to perform large-scale acts that were previously impossible with malware. Since botnets remain under the control of a remote attacker, infected machines can obtain updates, and their behavior changes on the fly. As a result, for substantial financial benefit, bot-herders can also rent access to portions of their botnet on the black market. Now, let's talk about different types of Botnets. Popular botnet-executed tasks include using your computer's strength to help shut down websites in distributed denial-of-service attacks, replacing banner ads directly aimed at you on your web browser. Pop-up ads are intended to make you pay for a fake anti-spyware kit for the elimination of the botnet. As mentioned above, each malware is harmful to our systems in one or another way. Some can jam your whole network and can corrupt. You whole files, and yes, some are just annoying, like pop-up messages. So, there should be a way to stop or at least minimize them. In our project we are working on this, we will design a PEDA to detect this malware before entering your system

TABLE I. COMPARISON OF MALWARE FAMILY BASED ON VARIOUS TECHNIQUES

Factors of Comparison \ Malware Family		Spyware	Adware	Trojan horse	Botnet	Worm	Virus
Creation Techniques	Pattern	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
	Obfuscated	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
	Polymorphic	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
	Toolkit	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Execution environment	Network	**Yes**	**Yes**	**No**	**Yes**	**Yes**	**No**
	Remote execution through a web	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**No**
	PC	**No**	**No**	**No**	**No**	**Yes**	**Yes**
Propagation Media	Network	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
	Removable disk	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
	Internet download	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
	Breaching confidentiality	**Yes**	**No**	**No**	**No**	**No**	**No**
Negative Impacts	Inconveniencing users	**No**	**Yes**	**No**	**No**	**No**	**No**
	Denying service	**No**	**No**	**No**	**No**	**Yes**	**Yes**
	Data corruption	**No**	**No**	**No**	**Yes**	**No**	**Yes**

TABLE II. COMPARISON OF VARIOUS RESEARCH ARTICLES BASED ON SECURITY PARAMETER AND THEIR TECHNIQUES

REFERENCE	ACCURACY	SECURITY	APPROACH TO PROVIDE SECURITY	Finding
Saja Alqurashi, Omar Batarfi (2016)	Yes (Using HMM as malware detection. it is based on fixed probabilities).	No	We train an HMM using the observation sequences to represent a set of data.	Hidden Markov models (HMMS) are generally used for statistical pattern analysis. HMM is used for software piracy detection.
Nwokedi Idika Aditya P. Mathur (2017)	YES (Anomaly-based, specification-based, and signature-based).	Yes	To make more straightforward and more efficient detection, they follow a three-way approach: anomaly-based, specification-based, and signature-based.	It provides three ways detection anomaly-based, specification-based, and signature-based. That further divides into three subdivisions dynamic, static, hybrid. The hybrid method is more accurate.
Rabia Tahir Published: 08 March 2018	Yes (Static analysis performs better in multipath malware. Also, their accuracy level is high compared to dynamic analysis.)	No	New methods are being used in the detection, which combines the existing techniques with machine learning and data mining methods.	Discussed about Heuristic and specification-based detection methods. These methods detect new and unknown malware. But the level of false-positive is high and needs more resources.
Yanfeng Ye, Tao Li, Donald Adjeroh, S. Sitharama Iyengar (June 2017)	Yes (Provide a comprehensive investigation on both the feature extraction and the classification/clustering techniques.)	No	The process of detection is usually divided into two stages: feature extraction and classification/clustering.	The performance of such intelligent malware detection approaches critically depends on the extracted features and the methods for classification/clustering.
Sana Aurangzeb - Muhammad Aleem - Muhammad Azhar Iqbal - Muhammad Arshad Islam	Yes (it performs header-based detestation.	No	The problems related to the traditional signature-based detection method are also highlighted. The rate of new malware destroying systems worldwide is increasing at an alarming rate.	The challenges that malware authors pose by developing complicated malware that frequently changes their signature to evade detection. It will release more sophisticated versions of malware that use new complication techniques.

F. Worms

This type of malware can infect an individual's whole machine in a brief period. When it gets attached to one file, it replicates itself to others and so on and infects one's whole network. Worms can alter and remove files, and additional malicious software can even be inserted into a computer. The aim of a computer worm is often just to create copies of itself over and over, depleting machine resources by overloading a shared network, such as hard drive space or bandwidth. Worms can also steal data, install a backdoor, and allow a hacker to gain control over a computer and its system settings, in addition to wreaking havoc on the resources of a computer. The worm can enter your system through email, messages, or via the internet, file sharing [4].

G. Spyware

Now it so simple to understand it's working from its name only. It is created to spy on someone's device activity. It does not let you know that it is present in the background of your device. However, it will record all your activities like typing, which applications you use most frequently, even your passwords and OTPs.

The information collected by spyware can predict about your web browsing nature or habits and taste for purchases online. It can compromise our system in many ways, like collecting confidential information like login credentials like username, passwords, account pins, credit card pins, tracking browser habits [5].

- When we discuss spyware types, we generally classify Spywares into categories such as, Via Trojan malware, which delivers the spyware software, Trojan spyware enters computers.
- Adware can track you to sell advertisers' information or serve tricky malicious ads. A website will implant cookie tracking files to follow you through the network.
- Device Control detects all computer operations, recording confidential data such as keystrokes, all the web visited pages, online transactions, emails, and more. Mainly, key loggers fall into this category.

III. METHODOLOGY AND RESULTS

The first program invented to work as an anti-malware was Flushot Plus by Ross Greenberg in 1987 [1]. Software companies develop detection systems products at laboratories, keep track of new programs, analyze them, and put the valid software in whitelist and malicious software on the blacklist. For the undecidable software, so-called grey list, the scanners operate them in a controlled environment for more classification. When analyzing a program in the grey list results in new malware, the company releases online updates for new malicious software. Then users can update their product databases by using remote access through an Internet connection. This research mainly has three objectives as mentioned below:

- Firstly, the behavior of ransomware is to be studied through a system call, or the Application Program Interface (API), of a Windows platform. The second objective is the selection of the Windows Operating System as the most used platform and most attacked one. API is the medium used by the program to interact with the operating system. For better understanding of the ransomware behavior, it is important to examine the API generated by ransomware.
- As this paper focused to develop an effective early detection system which can prevent from a ransomware attack or decrease the vandalization from the ransomware attack. Here, particularly the crypto ransomware is targeted in research which uses an encryption algorithm. With the help of encryption algorithm, the victim's data and files will be encrypted. The selection of crypto-ransomware is because the damage caused by this type is irreversible most of the time. Especially true, when the victim's data and files are encrypted with a robust encryption algorithm, such as the

  hybrid encryption. In pre-encryption stage, no encryption has been performed by the ransomware attack.
- The third objective is to produce a ransomware signature database. All the collected ransomware during literature having critical information is stored in this database. For generating the hash value as Ransomware identifier Secure Hashing Algorithm SHA-256 is used because it tends to generate a concise and unique identifier.
- This research aims to produce a ransomware dataset that allows machine learning methods to predict a ransomware attack. And last two points are more important according to the future perspective of research. To make more straightforward and more efficient detection, we also follow a three-way approach: anomaly-based, specification-based, and signature-based.

The valuable endowment of this research is the development of PEDA pre-encryption detection of crypto ransomware. We have taken APIs to compare to Header. Suppose that is found in the restored API's database. In that case, that will be an alert for malware. Additionally, we are also adding some extra detection techniques to block and detect all types of malwares.

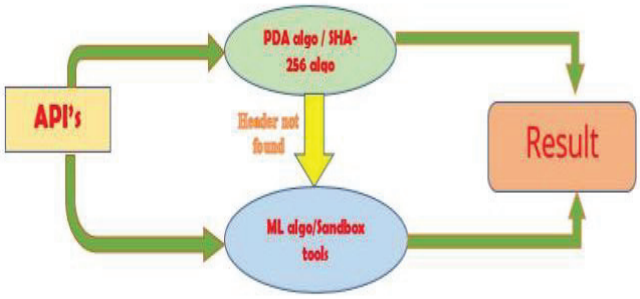

Fig. 2. Detecting Ransomware using API calls

We follow all steps of objective and methodology. In the end, we get the final output. This detector is written in python, and it works properly. It can detect Malware of URL, link, and file (window-based app).

Fig. 3. Comparing with API calls

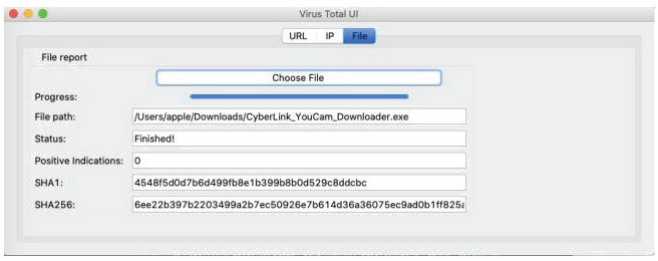

Fig. 4. Detecting Ransomware with API calls

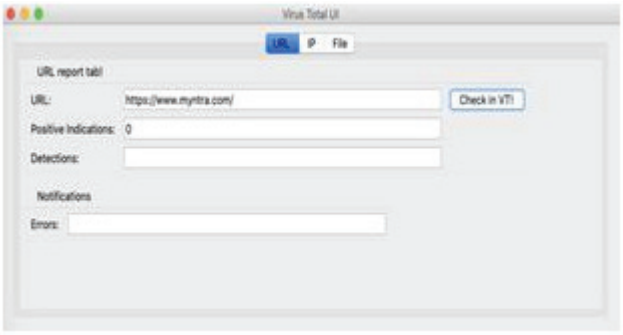

Fig. 5. Detecting URL malware

IV. CONCLUSION

As security is the major concern in this era. A systematic review of malware, different malware types, and detection techniques. Also, the comparative study for most of the malware families based on certain parameters are discussed. The developing processes of malware and its detection systems are rapidly growing. This study can be considered an essential reference for the developers in the field. The proposed model detects ransomware using API Calls.

REFERENCES

[1] S. H. Kok, A. Abdullah, and N. Z. Jhanjhi, "Early detection of crypto-ransomware using pre-encryption detection algorithm," J. King Saud Univ. - Comput. Inf. Sci., no., 2020.

[2] P. V. Shijo and A. Salim, "Integrated static and dynamic analysis for malware detection," Procedia Comput. Sci., vol. 46, no. Icict 2014, pp. 804–811, 2015.

[3] A. Arabo, R. Dijoux, T. Poulain, and G. Chevalier, "Detecting ransomware using process behavior analysis," Procedia Comput. Sci., vol. 168, no. 2019, pp. 289–296, 2020.

[4] N. Idika and A. P. Mathur, "A Survey of Malware Detection Techniques," SERC Tech. Reports, 2007.

[5] K. Rieck, T. Holz, C. Willems, P. Düssel, and P. Laskov, "Learning and classification of malware behavior," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 5137 LNCS, pp. 108–125, 2008.

[6] I. Santos and J. Devesa, "OPEM: A Static-Dynamic Approach for Machine-Learning-Based Malware Detection," pp. 271–272, 2013.

[7] H. V Nath and B. M. Mehtre, "Static Malware Analysis," pp. 440–450, 2014.

[8] M. A. Jerlin and K. Marimuthu, "A New Malware Detection System Using Machine Learning Techniques for API Call Sequences," J. Appl. Secur. Res., vol. 13, no. 1, pp. 45–62, 2018.

[9] A. Shabtai, E. Menahem, and Y. Elovici, "Automatic Malware Signature Generation," October, vol. PP, no. October, pp. 1–15, 2010.

[10] S. Alqurashi and O. Batarfi, "A Comparison of Malware Detection Techniques Based on Hidden Markov Model," J. Inf. Secur., vol. 07, no. 03, pp. 215–223, 2016.

[11] S. Gupta, H. Sharma, and S. Kaur, "Malware characterization using Windows API call sequences," J. Cyber Secur. Mobil., vol. 7, no. 4, pp. 363–378, 2018.

[12] N. Naik, P. Jenkins, N. Savage, and L. Yang, "Cyber threat Hunting - Part 2: Tracking Ransomware Threat Actors using Fuzzy Hashing and Fuzzy C-Means Clustering," 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2019, pp. 1-6, DOI: 10.1109/FUZZ-IEEE.2019.8858825.

[13] K. Raman, "Selecting Features to Classify Malware," InfoSec Southwest 2012, pp. 1–5, 2012.

[14] S. Gadhiya, K. Bhavsar, and P. D. Student, "Techniques for Malware Analysis," Int. J. Adv. Res. Comput. Sci. Softw. Eng., vol. 3, no. 4, pp. 2277–128, 2013.

[15] R. Tahir, "A Study on Malware and Malware Detection Techniques," Int. J. Educ. Manag. Eng., vol. 8, no. 2, pp. 20–30, 2018.

[16] S. Najari, "Malware Detection Using Data Mining Techniques," Int. J. Intell. Inf. Syst., vol. 3, no. 6, p. 33, 2014.

[17] S. Aurangzeb, M. Aleem, M. A. Iqbal, and M. A. Islam, "Ransomware: A Survey and Trends," J. Inf. Assur. Secur., vol. 12, no. January 2020, 2017.

[18] A. Vashisth and R. S. Batth, "An Overview, Survey, and Challenges in UAVs Communication Network," 2020 International Conference on Intelligent Engineering and Management (ICIEM), 2020, pp. 342-347, DOI: 10.1109/ICIEM48762.2020.9160197.

[19] A. Vashisth, R. Singh Batth, and R. Ward, "Existing Path Planning Techniques in Unmanned Aerial Vehicles (UAVs): A Systematic Review," 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), 2021, pp. 366-372, DOI: 10.1109/ICCIKE51210.2021.9410787.

[20] C Rohith and G. Kaur, "A Comprehensive Study on Malware Detection and Prevention Techniques used by Anti-Virus," 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), 2021, pp. 429-434. DOI: 10.1109/ICIEM51511.2021.9445322.