Windows Ransomware detection using KERNEL level features with machine learning

By

Abu Baker

MSIS-16

A thesis submitted to the faculty of Department of Information Security, Military College of Signals, National University of Sciences and Technology, Rawalpindi in partial fulfilment of the requirements for the degree of MS in Computer Software Engineering

September 2021

ABSTRACT

Ransomware is a type of malware which denies access to a user's data by employing locking, deletion especially encryption mechanisms. Due to increasing trends of ransomware in new malwares and disastrous nature of malware, a lot of work has been done to effectively detect and prevent ransomware attacks. Behavior Based detection is carried out by differentiating dynamic behavior of malign and benign applications and creating model to detect malign behavior. Studies conclude that the behavior of ransomware applications from most benign application is very different and easy to detect while some applications like Desktop Encryptors, Compressors and Shredders depict almost same behavior as a ransomware. Dynamic analysis focused on such applications will be helpful in decreasing the false positives of already defined and tested models for ransomware detection.

We have conducted a study to find common and differentiable features on kernel level to identify legitimate full desktop encryptor applications and ransomware by analyzing IRPs using a customized minfilter driver, to improve the ransomware detection model. The functional objective of both type of applications is same since it both are required to make the target data inaccessible for unauthorized personnel without a key. We researched the pattern of encryption for both applications and were able to identify encryptors from ransomware and hence, participated in the improvement of detection capability of existing models.

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by **Mr. Abu Baker**, Registration No. **00000205676**, of **Military College of Signals** has been vetted by undersigned, found complete in all respect as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial, fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the student have also been incorporated in the said thesis.

Signature: _____

Name of Supervisor: Dr. Mir Yasir Umair

Date: _____

Signature (HoD):_____

Date: _____

Signature (Dean/Principal):_____

Date: _____

Declaration

I hereby declare that no portion of work presented in this thesis has been submitted in support of another award or qualification either at this institution or elsewhere.

_____

Signature of Student

Name: Abu Baker

Reg No.: 00000205676

DEDICATION

This thesis is dedicated to my father for giving me unconditional love, for always holding my hand in tough times and for being my safe heaven, to my mother for her prayers and to all teachers in the world, enabling people like me to stand with their heads high.

Acknowledgments

I would like to thank my supervisor, Asst Prof Dr. Mir Yasir Umair, for his help & guidance in making me complete this thesis. He has been a big support throughout the process. I also would like to thank my GEC members Asst. Prof. Dr. Hammad Afzal, Asst. Prof. Mian Muhammad Waseem Iqbal Kakakhel and Asst. Prof. Waleed Bin Shahid for their constant support.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

The advancement of computing technologies has enhanced the human's ability to record, analyze, access and transfer data enormously. The information located at network-connected devices like webservers, emails and databases is supposed to be easily accessible by authorized personnel. The same information is accessible to financially motivated threat actors in a compromised network to alter, steal or simply make it inaccessible for the businesses to hurt business objectives if not paid ransom. This dependency of businesses to keep databases and important applications has motivated threat actors to highly develop tech-savvy capabilities to deploy a virus called, Ransomware.

The quick change in the environments without following well-defined change management protocols, open email servers to accept emails from unknown senders, publicly exposed unpatched servers and remote access are some of the threat vectors used by threat actors to get initial access. The state of internal security of organizations or ability to handle insider threats is almost none which helps threat actors to move laterally and gain privileges to exfiltrate data as well as encrypt that data across the organizations. A short-term solution to this is offline backups which create an enormous human resource and computational expense for the organization as well as they are not able to cope up with GBs of daily updated data.

1.2 Motivation and Problem Statement

Ransomwares are deployed by highly advanced adversaries, normally called Advanced Persistent Threat (APT) who are able to bypass or disable security controls deployed in the

organization. Ransomware detection using dynamic analysis on user level is computationally intensive and prone to bypassing while detectors performing dynamic analysis using KERNEL level features considers valid applications like Desktop Encryptors, compressors and shredders as ransomware and increases the false positive rate to decrease the efficiency of systems.

Through this work, we have tried to reduce the false positive rate and identify the features which ca be used to identify legitimate software like desktop encryptors, compressors and shredders from ransomware.

1.3 Objectives

We targeted to achieve following objectives in our research project.

- Dynamics analysis of ransomware by studying API calls and I/O Request Packets features.
- Extracting features from analysis and selecting usable features for detection to improve the computational efficiency of machine learning algorithm.
- Develop a prototype to detect and prevent execution of ransomware.

1.4 Thesis Contribution

The main contributions of this work are as follows.

- Analysis of Ransomware types, Ransomware groups and families.
- Study of detection techniques and features used and recommended by academia as well as common industry practices.
- Execution of Ransomware in the sandboxed environments
- Recommendation of features from IRPs to reduce the false positive rates.

1.5 Thesis Outline

The thesis is structured as follows:

- Chapter 2 contains the literature reviewed in the thesis.
- Chapter 3 contains the proposed approach used with descriptions of Ransomware and traditional features to detect ransomwares.
- Chapter 4 covers the results obtained after the experiments and analysis.
- Chapter 5 contains the conclusion & future works.

Chapter 2

Literature Review

2.1 Overview

In this chapter, a brief overview of existing methods used to perform the task of ransomware detection and prevention is presented. Our focus is mainly on such methods which are used on detect ransomware using kernel level features; however, literature on user level ransomware detection and prevention is also briefly described. We would also look at the datasets used by different researchers and their attributes.

Considering the disastrous nature of ransomware activities and increasing trend of organizations being compromised and data made inaccessible, there are a lot of researchers committed to researching novel ways to detect and prevent this destructive malware. We have reviewed number of papers presenting different frameworks for detecting and preventing ransomware in enterprise networks and endpoints.

Recent ransomware attacks on colonial pipeline, US and JBS, Brazil has severely affected the critical infrastructure and supply of basic supplies to population of US. The colonial pipeline is responsible for supplying 2.5m barrels of gasoline across the US. The disruption in this supply affected business adversely, delivered effects of ransomware to public and forced colonial pipeline to pay 4.4 million US dollars to get control of their systems and prevent leakage of information from the hackers[1].

JBS is the largest meat processing company and head quartered at Brazil. The biggest 5 plants of this organization were taken offline after a successful ransomware attack which are responsible for processing ¼ of the total beef processing and 1/5 of the total port meat processing in US. The attack was successful to disrupt the meat supply to US vendors and

supermarkets[2].

These actions have forced FBI to compare the ransomware threat with 9/11 or terrorism threat. Government of United States has released an executive order for the study of ransomware attacks and data sharing in the form of indicators of compromise and lessons learn from the compromises[3].

Due to inclusion of new and effective back up strategies by organizations to handle ransomware, it is being observed that the data is being exfiltrated after initial compromise to force victims to pay ransom so that the data could be leaked if victim refuses to meet the demands. In fact, this has become the most motivating factor for organizations to pay ransomware in the current wave[4].

Ransomware has become the most offending threat in cyber security industry with the growing number of victims and ransomware payments. The growing amount of ransomware payments in helping adversaries invest in developing further sophisticated attack and compromise techniques. The ransomware payment has crossed $200 billion in 2020, compared to $11.5 in 2019. More than 50% of the organizations were successful at recovering their data from offline backups. Ransomware is causing a lot of damage to organizations by affecting operations as well as reputations of organization[5]. We are observing heightened level of sensitivity in organization for putting data integrity and confidentiality controls to protect the sanity of data.

2.2 Existing Detection Techniques

Ransomware has arisen as quite possibly the most troublesome scareware1 to shield from, as it very well may be computationally infeasible to return ransomware's damage. There are two fundamental kinds of ransomware accessible in the wild: the first, storage ransomware, is intended to bolt the casualties' PC, to keep them from utilizing it; the subsequent one, and most regular these days, is crypto-ransomware, which engraves individual documents to make them

difficult to reach to its casualties. In the two cases, clients are compelled to pay a payment to recapture access either to their information (expecting no reinforcement component is set up) or framework. Numerous victims feel their information are imperative to the point that they need to pay the payoff. For instance, when in 2012, Symantec had the option to destroy an Order and Control (C&C) network utilized by the CryptoDefense ransomware family, the follow-on examination showed that 2.9% of casualties, out of 68,000 special contaminations, seemed to have paid the payoff. Blackmail components, consequently, produce huge income for the aggressors. For CryptoWall adaptation 3, measurements represent an expected complete $325 million in harms in the US alone. One study proposes EldeRan, an AI approach to group ransomware dependent on their initial activities. The fundamental presumptions of this work depend on the perception that ransomware contain special powerful highlights and, to stop their spread, it is urgent to distinguish new variations during their first appearance. To this end, EldeRan first and foremost chooses the important highlights that describe the ransomware conduct, and afterward order each recently introduced application on a client PC through an AI calculation as to perform recognition without depending on old style heuristic or mark based procedures. We propose EldeRan, a system to recognize the huge ransomware dynamic highlights, and use them to distinguish ransomware. Through the Shared Data basis, we have recognized the most pertinent unique highlights among a huge arrangement of thought about ones. EldeRan abuses a moderately little arrangement of highlights without lessening the presentation of the AI classifier. By following this methodology, EldeRan is additionally appropriate to distinguish new ransomware families. We contrast the characterization results and those of VirusTotal: EldeRan's normal mistake rate is 2.4% while that of VirusTotal is 5.6%, and EldeRan accomplishes a striking 96.3% identification rate. We additionally tried the capacity of EldeRan to recognize new groups of ransomwares, getting a normal location pace of 93.3%[6].

Recognition of malware utilizing static-based investigation implies breaking down an application's code preceding its execution to decide whether it is prepared to do any vindictive exercises. In the event that the static examination tracks down any malignant code, the executable will be halted from dispatching. The most widely recognized kind of static investigation, which is usually utilized in business infection scanners, is alluded to as signature examination. In the signature investigation, code string designs (marks) are separated from the objective application's code and contrasted with a storehouse of realized malignant code designs. Mark put together identification depends with respect to a colossal store of vindictive code marks. This store should be every now and again refreshed to stay current, which is in no way, shape, or form an insignificant assignment. Business infection scanners regularly have huge groups of online protection specialists that ceaselessly find, research, and concentrate malevolent marks. Dynamic-based examination location involves the live checking of cycles, to decide whether any are carrying on with any malignant expectation. Any noxiously acting interaction will be hailed as hazardous and ended[7].

One can sum up the contrast between static and social examination for recognition in an accompanying way: in the static investigation, induction of the conduct qualities are produced using the double record of an obscure executable. This induced conduct is then utilized by a straightforward coordinating with calculation to relegate a danger level (for example protected or pernicious). In the social investigation, the conduct qualities of the executable are known as it is being seen progressively, and deductions are settled on by an inductive choice calculation on the danger level. The critical distinction between static and conduct-based location is where the deduction is made – static examination derives social qualities from the noticed parallel record, dynamic conduct construes a danger level from noticed conduct. This is a significant distinction as static muddling procedures modify what

conduct is gathered without really meaning for the genuine conduct, hence delivering it repetitive against social-based identification. In social-based recognition, all executables are treated as obscure, where it is up to the executable to demonstrate it is acting in a protected, non-malignant way. In doing as such, the capacity to distinguish zero-day (obscure) assaults is considerably improved. A social way to deal with ransomware discovery requires a dynamic calculation that acknowledges a quantitative conduct hint of a running interaction as a contribution, to yield a basic double choice - yes, it is protected/benevolent or no it is ransomware. Clarified with regards to ransomware recognition, regulated AI is the preparation of a choice calculation to perceive certain social attributes (the info information) of running cycles that ideally separate among ransomware and benign programs[7].

A thorough research was conducted for analysis of ransomware and a framework based on data provided by Promon, a sysinternals tool for monitoring and analyzing process interaction on NT based systems. if In this paper, they present an investigation framework that follows a powerful way to deal with distinguish ransomware assaults and model its conduct. In this methodology, the framework creates a practical, fake client climate where ransomware tests are executed and their associations with the framework climate observed. Close perception of the association of the ransomware with the record framework allows the framework to distinguish cryptographic ransomware conduct. All together for a ransomware assault to succeed, ransomware should get to the client's framework, meddle with the documents and lock the framework leaving it out of reach. numerous ransomware tests are broke down considering the location of ransomware by noticing the record framework. Likewise, this methodology gives experiences on the best way to separate between unmistakable ransomware families like Cerber, CryptoWall, Crysis, and so on by inspecting their record framework calls. This methodology uses the open-source Cuckoo Sandbox which establishes a protected climate for executing untrusted and possibly noxious executable that keep them from spreading and

don't forfeit a client's documents or private data. Cuckoo acknowledges any executable twofold record and creates an itemized report drafting the conduct of the document when executed in a practical climate.

In their exploration, investigated 495 of the most recent malware tests, and this methodology had the option to accurately recognize and distinguish 479 ransomware tests from numerous ransomware families with no bogus positives. The 96.7% identification pace of this methodology recommends that it performs better compared to the current conduct-based examination frameworks as to recognizing and identifying ransomware tests practically. This model can undoubtedly be sent and utilized on any malware examination framework[8].

Since malware regularly enters frameworks through known weaknesses, the best advance to reinforce protections is to forcefully fix systems. By wiping out weaknesses, the malware might not have an approach to get on any of your PCs in any case. In case of an assault, associations can limit harm on the off chance that they can identify the malware early. For starting misuse and infection, a decent protection is to get marks and IOCs into an IDS or other organization gadget. Use danger insight sources to impede or possibly aware of the presence of inconsistencies related with ransomware in your organization traffic once the ransomware has effectively grabbed hold of one gadget, there are steps to contain it locally so that organization records aren't influenced. Having an endpoint insurance framework that can search for the execution and execute the interaction is typically the best methods for regulation. On the off chance that payment product is recognized, once the ransomware episode has been distinguished and has been contained, the subsequent stage is eradicating it from the organization. It is typically suggested that machines be supplanted as opposed to clean. Similarly, as with a malware, it is hard to know whether leftover documents are covered up on the framework and ready to re-contaminate gadgets. Be that as it may, for network areas, for example, post boxes or document shares, in some cases it is more applicable to clean those

areas, eliminate the malicious email message from the letter box, or eliminate the ransomware guidelines from the record share. On the off chance that associations decide to clean instead of supplanting, it is significant that they keep on checking for sig qualities and other IOCs to keep the assault from reappearing. For the last advance recuperation, the main assignment will be reestablishing from back-up. On the off chance that there are acceptable verified back-ups, any ransomware occasion can truly be made into a non-issue by essentially supplanting or cleaning frameworks and recuperating from back-ups[9]. There might be a limited quantity of personal time, however it should not be a major multi-day issue. n most ransomware cases, a full examination concerning what explicit disease vector was utilized against the framework is a significant advance. Was it a phishing email, or was it an online assault unit? On the off chance that it was an electronic assault unit, how did that client get to that page? Knowing how the ransomware went onto your framework can help associations better prime their protection frameworks.

Malicious software, referred to as malware, is a consistently developing security danger and the recognition of malware stays a significant space of exploration. The initial phase in discovery is analysis. This includes either static or dynamic investigation of known malware and is typically performed disconnected with master human info. Aftereffects of the examination are refined into a "signature". One method for malware identification is the utilization of static marks to look at programs after they are stacked and before execution. Shockingly, this can be crushed by malware that utilizations obscurity. Because of this, unique conduct-based identification has been proposed These strategies screen the conduct of the framework utilizing working framework or hypervisor-based instrumentation to recognize malevolent conduct. Static and dynamic marks can be inferred utilizing either deterministic or measurable methods.

Typical malware analysis techniques – both static and dynamic – are executed in programming. In this paper, we contend against unadulterated programming executions both because their overhead and the resulting confided in registering base (TCB) bulge. Programming for malware discovery is defenseless to the very weaknesses that malware abuses during contamination and thusly can be, and frequently is, handicapped by malware. Equipment helped recognition components are not defenseless against such impairing and have been proposed for explicit classes of malware Hardware-based control-stream respectability (CFI) approaches remain rather than these outside-the-processor plans. CFI is an in-processor screen gathering equipment execution counters accessible in contemporary processors to recognize control-stream deviations. address the code-reuse assault by upholding in reverse edge CFI with equipment support. These techniques depend on master information on the executable twofold and its memory formats. In this paper, author propose an alternate malware identification situation. Rather than looking for a solitary model that recognizes all pernicious and favorable applications, we learn one model for every application that isolates its malware-contaminated executions from real executions. The model is prepared on both the malignant and considerate conduct of the application. At the point when the program is stacked, its related conduct model is stacked, and its execution is observed. On the off chance that the cycle executes in a way that makes the related model banner its conduct as dubious, a product special case is raised. The actual program is a genuine application. Maybe than arranging the program as pernicious/considerate as in the "conventional" situation, our finder recognizes runs where the contamination is set off from ones where it is not and raises a special case when noxious conduct is identified[10].

A significant test in checking memory gets to be the sheer volume of the information. Our system tends to this by isolating gets to into ages, summing up the memory access examples of every age into highlights which are then taken care of to an AI classifier. Investigations show

this system is powerful in distinguishing assorted classes of malware. The commitments of this paper we focused on application-run malwares and presented a structure of malware identification based on virtual memory access patterns. The paper presented novel syscall memory extraction techniques leading to low false positive and true negative rates[10].

Ransomware, as different classes of malware, utilizes various methodologies to avoid discovery, spread, and assault clients. For instance, it can perform multi-contamination or cycle infusion, ex filtrate the client's data to an outsider, encode documents, and build up secure correspondence with C&C workers. Our identification approach accepts that ransomware tests can and will utilize the entirety of the strategies that other malware tests may utilize. Moreover, our framework accepts that effective ransomware assaults perform at least one of the accompanying exercises.

Determined work area message. After effectively playing out a ransomware contamination, the vindictive program ordinarily shows a message to the person in question. This "recover note" advises the clients that their PC has been "bolted" and gives guidelines on the most proficient method to make a payment installment to reestablish access. This payoff message can be created in an unexpected way. A well-known procedure is to call devoted Programming interface capacities (e.g., Make Work area) to make another work area and make it the default arrangement to keep the casualty out of the undermined framework. Malware essayists can likewise utilize HTML or make different types of determined windows to show this message. Showing a tenacious work area message is an exemplary activity in numerous ransomware assaults [11]. Unpredictable encryption and cancellation of the client's private records. A crypto-style ransomware assault records the casualty's documents and forcefully encodes any private documents it finds. Access is confined by retaining the unscrambling key. Encryption keys can be produced locally by the malware on the casualty's PC, or distantly on C&C workers, and afterward conveyed to the undermined PC. An aggressor can utilize altered

ruinous capacities, or Windows Programming interface capacities to erase the first client's documents. The assailant can likewise overwrite records with the encoded form or utilize secure cancellation by means of the Windows Secure Erasure Programming interface.

Specific encryption and erasure of the client's private documents dependent on specific credits (e.g., size, date got to, augmentation). To stay away from location, countless ransomware tests scramble a client's private records specifically. In the most straightforward structure, the ransomware test can list the documents dependent on the entrance date. In more refined situations, the malware could likewise open an application (e.g., word.exe) and list as of late got to records. The example can likewise infuse vindictive code into any Windows application to acquire this sort of data (e.g., straightforwardly perusing measure memory). Then again in this paper, they talk about the Ensuring malware examination conditions against fingerprinting methods is non-paltry in a certifiable sending. Modern malware creators misuse static highlights inside investigation frameworks (e.g., name of a PC) and dispatch observation based assaults to unique finger impression both public and private malware examination frameworks[11]. Naturally, a potential way to deal with address such surveillance assaults is to fabricate the client climate so that the client information is substantial, genuine, and non-deterministic in each malware run. These consequently produced client conditions fill in as an "tempting objective" to urge ransomware to assault the client's information while simultaneously forestalling the chance of being perceived by enemies.

Ransomware keeps on being quite possibly the main security dangers on the Web. While ransomware is definitely not another idea (such assaults have been in the wild since the most recent decade), the developing number of prominent ransomware assaults has brought about expanding worries on the best way to protect against this class of malware Nonetheless, the developing number of paying casualties proposes that unsophisticated clients – who are the fundamental objective of these assaults don't follow these proposals, and effectively become a

paying survivor of ransomware. Our assessment exhibits that Recovery can guarantee zero information misfortune against current ransomware families without bringing down the client experience or actuating alert weakness. What is more, we show that Recovery causes humble overhead, averaging 2.6% for sensible jobs.

In this article, likewise t expanding the working framework with a bunch of lightweight and nonexclusive strategies, which we all in all call redemption it is feasible to stop present day ransomware assaults without changing the semantics of the basic document framework's usefulness or performing huge changes in the engineering of the working framework. Our examinations on 29 contemporary ransomware families show that our methodology can be effectively applied in an application-straightforward way and can fundamentally upgrade the current assurance abilities against ransomware (accomplishing a genuine positive [TP] pace of 100% at 0.8% bogus positives [FPs]). At last, we show that this objective can be accomplished without a detectable presentation sway, or different changes to the manner in which clients associate with standard working frameworks. To sum up, we make the accompanying commitments. – An overall way to deal with safeguarding against obscure ransomware assaults in a straightforward way. In this methodology, admittance to client records is intervened, and favored solicitations are diverted to an ensured territory, keeping up the reliable condition of client information. – It shows that proficient ransomware security with zero information misfortune is conceivable. – A model execution for Windows and assess it with genuine clients to show that the framework can secure client records during an obscure ransomware assault while forcing no noticeable presentation overhead. Considering the expanding ransomware danger, clients are regularly informed to make reinforcements with respect to their basic information. Absolutely, having a dependable information reinforcement strategy. They need to apply this redemption configuration is adequately broad to be applied to any operating system that is a likely objective for ransomware. Nonetheless, the fabricated model for the

Windows climate which is the principle focus of current ransomware assaults today. It is carried out this module as a client mode administration. This was a cognizant plan decision like the plan of most enemy of malware arrangements. Note that Microsoft formally upholds the idea of secured administrations, called Early Dispatch against Malware (ELAM), to permit hostile to malware client mode administrations to be dispatched as ensured administrations. furthermore, Recovery should mediate on completely advantaged gets to touchy documents. The execution of the framework depends on the Windows Part Advancement structure with no changes on the hidden document framework semantics. To this end, it gets the job done on Windows to screen the compose or erase demands from the I/O framework to the base record framework driver [12].

2.3 Analysis

Our detailed literature review brings up following facts:

- Ransomware is a threat to be acknowledged to enterprise's business objectives and availability of operations.
- Ransomware is being used to bully enterprises to collect ransom and then those monetary gains are being used to fund research on developing advanced offensive techniques as well as terrorism and a range of crimes.
- An increased trend of data exfiltration before the encrypting the data is being observed causing organizations to pay ransom even if an efficient back up policy is in place
- The service of encrypting organizations is being sold as ransomware as a service on darkweb where initial access to organizations is being bought from a separate group focused on initial access operations and lateral movements which has provided

ransomware operators to just focus on making their encryption process more efficient as well as stealthy.

- Multiple frameworks and detection techniques are being developed to fight against ransomware focusing user-mode and kernel-mode features.
- Kernel mode features provide an enhanced rate of detection than user-mode features, but user-mode detection mechanisms are computationally efficient than kernel-mode detection mechanisms.
- The most efficient way to detect ransomware on kernel-level has been proposed and tested by Amin Kharraz et.al but the model is showing 100% false positive rate in case of legitimate file handling tools such as compressors, desktop encryptors and shredders.
- New kernel level features need to be identified to distinguish between ransomware and legitimate file handling applications.

Chapter 3

Proposed Approach

We recommend an approach to monitor interaction of processes with the filesystem drivers by inserting a minifilter driver between service manager and file system driver to monitor the I/O request packets. The next step is build a baseline for applications interacting with file system similar to what A. Kharraz and et. al. suggested in their work. The final step would be to identify legitimate file handling applications from the ransomware applications.

3.1 Windows I/O system and operation

The windows I/O systems was created based on several executive components which communicate the data inserted by user or a program on application level to Hard drive through Kernel Manager, and Device drivers using I/O request packets known as IRPs [13].

The core components of I/O system are:

- I/O Manager provides infrastructure for Device drivers for connecting application and system components to devices.
- Device Drivers are responsible for forwarding read / write / operate instructions from I/O manager to the devices they manage. They also send status information of device manger request to I/O manager and messages other drivers involved in device management through I/O manager.
- PnP manager works with installation and removal of hardware components and related drivers.
- The registry provides database for device connectivity, status and description of connected devices which is populated during deriver installation by INF files.

- Windows Management instrumentation (WDM) provides an interface to application and programs in user-mode, showing itself as hardware service provider.
- The hardware abstraction layer (HAL) provides APIs for compatibility and provide support for the devices for which drivers are not present.

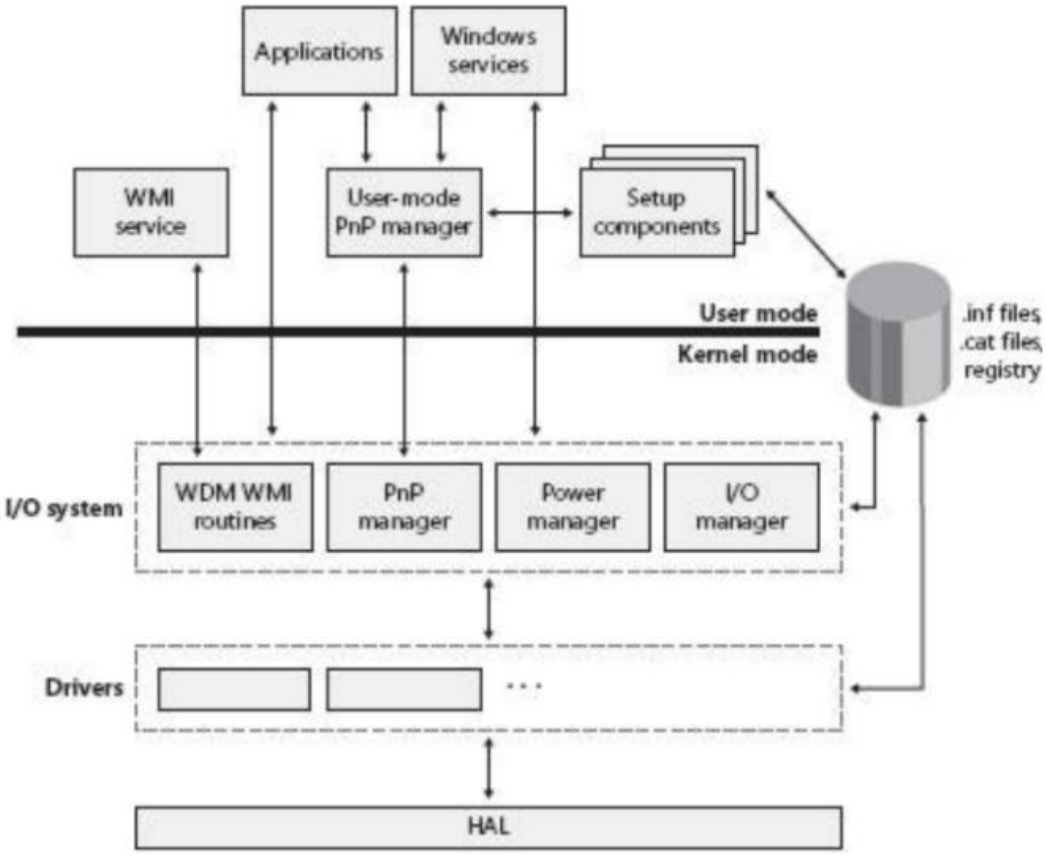

Figure 3. 1 Windows I/O System
Source: Windows Internals, 5th Edition, Chapter 7

I/O manager uses a packet driver approach to handle communicate with device drivers and perform transactions requested by applications to hardware devices. Each packet is called Input/Output Request Packet (IRP) and all the operations on device drivers and hardware devices are performed through these packets excluding some exceptions [13].

The windows Input Output system is designed in a compartmentalized and modular manner where each module is responsible for handling the requests received and forwarded to and from other modules. This design approach helps to add new features, backward compatibility, and easy upgrades. A user mode API is responsible for handling incoming read / write requests from the user or an application such a notepad.exe. The R/W request is forwarded to I/O system Services API which translates user request is to I/O Service request. The I/O System API is then sent to I/O manger which maps the service request to a hardware device and relevant device drivers. The I/O manager create a packet known as I/O Request packet which contains the timestamp of the request, the operation to perform, the target object and offset and address of the read write buffer as per the nature of request. This IRP is then forwarded to device drivers which are responsible for controlling actual operations on the on hardware devices through Hardware Abstraction Layer.

A typical I/O operation with involved components is present in the figure below.

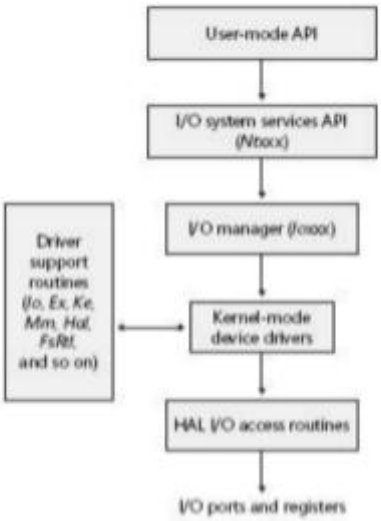

Figure 3. 2: Windows Typical I/O Operation
Source: Windows Internals, 5th Edition, Chapter 7

3.2 Filter drivers

File system drivers receive I/O requests to files and then issue their own explicit requests to the disk drivers with logical address of the drive to store data. These operations are handled via FIO and IRPs [13].

Generally, when a process requests to read / write or access resources on disk drive, the request is sent to I/O Manger which is then converted into an I/O request packet. This IRP is then sent to file system drivers which convert these to disk drive requests for accessing data at logical addresses [14].

Fast I/O is a special mechanism to bypass IRP creation and enter the request directly into file system driver stack.

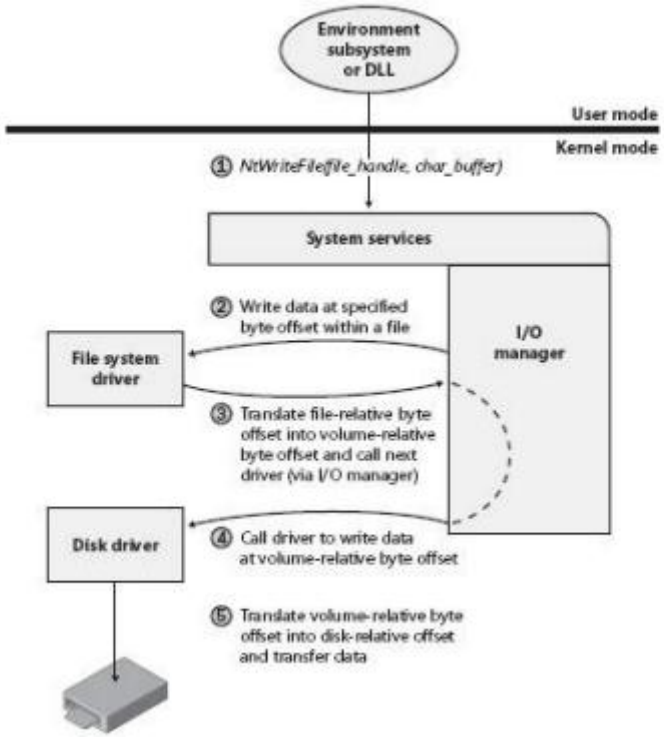

Figure 3. 3: File Write Scenario
Source: Windows Internals, 5th Edition, Chapter 7

A minifilter driver is user to filter disk drive access request including IRP, FIO and FSFilter callbacks. These call backs are put into a file based on the selection criteria. Minispy is a minifilter driver developed by Microsoft to observed I/O requests and transactions being conducted in the system [15].

We have developed a customized version of mini filter driver to focus on the relevant data and extract features which are required. The relevant data features are IRP function codes being conducted (IRP_MJ_CREATE, IRP_MJ_READ, IRP_MJ_WRITE, IRP_MJ_SET_INFORMATION), the file and directory being accessed, the offset in the hardware address on the disk and the cardinality of read/write buffers for the corresponding IRP. The operation and collection of minifilter for intercepting and logging IRPs is shown in the figure 3.4. A filter similar to Amin Kharraz's Redemption filter was written to extract IRPs to identify the access patterns for ransomware applications. We have tried to develop the filter to be used in cascading configuration with the Amin Kharraz's end point detection system. The target would be to identify the legitimate file handling applications from ransomware.

We are monitoring following IRP major codes to monitor file system activities

- IRP_MJ_CREATE is being monitor for creation of handles and creation of new files on the hardrive. IRP_MJ_CREATE operation code is called when a handle to an existing or non-existing file is requested by user level API call.
- IRP_MJ_READ follows the IRP_MJ_CREATE code for reading data from a handle. The device driver is supposed to transfer data to Read buffer from the disk on reception of this code.
- IRP_MJ_WRITE is being monitored to monitor the write operation on disk. The operation of this code is exactly reverse to that of IRP_MJ_READ code where the data is transferred from Write Buffer in memory to the I/O Device / Disk.

- IRP_MJ_SET_INFORMATION is being monitored to observe metadata changes to a files or I/O Objects. When a file's metadata is changed i-e the file is created or deleted, the IRP_MJ_SET_INFORMATION is called.

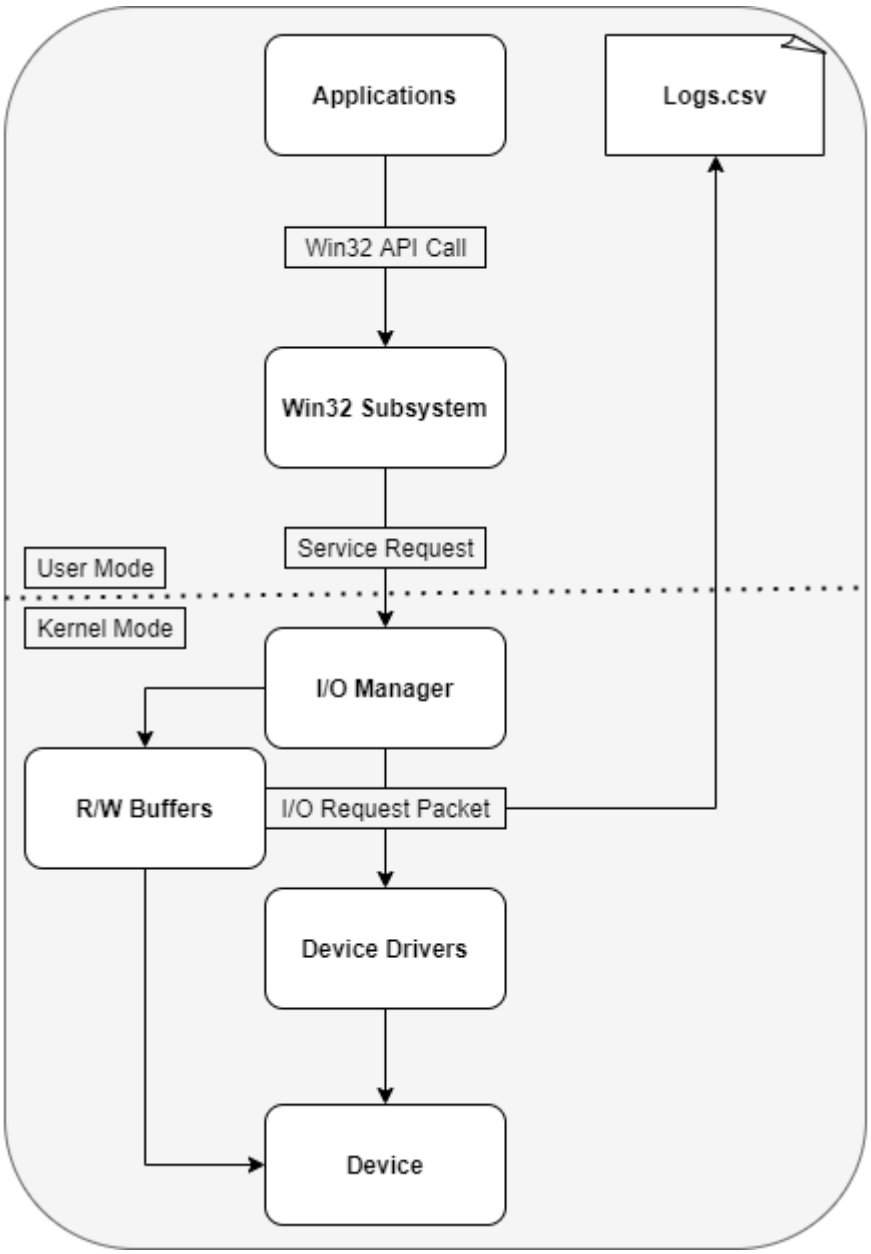

Figure 3.4: IRP Interception

3.3 Dataset

We downloaded a repository virushare of 7.8 GB of 36,000 samples, multiple samples from ransomware repository zoo and collected some crypto ransomware samples found in incident response engagements in multiple organizations. Then we used virustotal to identify and categorize crypto ransomware and associated them with their families.

Since we were not able to identify a valid approach towards handling this problem with conventional machine learning techniques and code, we created a set of scripts to extract parameters which would be interesting to us. Due to this limitation, we used the data set in such a way that maximum number possible behaviors could be noted and included in the analysis. Unfortunately, we were not able to apply deep learning and machine learning methods to results due to limited skills in coding and use of machine learning software, we settled for execution of unique samples of malware and legitimate applications. Here is the summary of software's we used during execution.

Sr	Name	Category
1	VeraCrypt	Encrypter
2	BitLocker	Encrypter
3	DiskCryptor	Encrypter
4	7-Zip	Complessor / Encrypter
5	AxCrypt	Encrypter
6	Winzip	Compressor
7	Eraser	Shredder
8	Sdelete	Shredder

Table 3.1 List of Legitimate File Handling Applications

Sr	Ransomware Family	Hash
1	CrypoLocker	ddb9b850fa0eee2f62463728b07bffc11eaa9b241d215029eaddf1de4ec54936
2		43c5f2e7aacbc9a3439a810e3768087b7c8bea191ef84d71b2aa8686befed073
3		e908dca957b9cb7759feeabef0f2921e3cb236368acc5e124e87af0492308b14
4	Cerber	e67834d1e8b38ec5864cfa101b140aeaba8f1900a6e269e6a94c90fcbfe56678
5	Cyptowall	60d574055ae164cc32df9e5c9402deefa9d07e5034328d7b41457d35b7312a0e
6	JigSaw	86a391fe7a237f4f17846c53d71e45820411d1a9a6e0c16f22a11ebc491ff9ff
7	Locky	0537fa38b88755f39df1cd774b907ec759dacab2388dc0109f4db9f0e9d191a0
8	Mamba	2ecc525177ed52c74ddaaacd47ad513450e85c01f2616bf179be5b576164bf63
9	Matsnu	f73027dd665772cc94dbe22b15938260be61cbaad753efdccb61c4fa464645e0
10	Petrwrap	cf01329c0463865422caa595de325e5fe3f7fba44aabebaae11a6adfeb78b91c
11	Petya	33ca487a65d38bad82dccfa0d076bad071466e4183562d0b1ad1a2e954667fe9
12	Radamant	3e1813da2d561157df7667cde0117fdddd883c5b1272f76d1ae85ad889c38220
Sr	**Ransomware Family**	**Hash**
13	RedBoot	7fa2bf61405ac573a21334e34bf713dcb5d1fc0c72674e6cebc48d33a4a14d44
14	Rex	32856e998ff1a8b89e30c9658721595d403ff0eece70dc803a36d1939e429f8d
15	Satana	4785c134b128df624760c02ad23c7e345a234a99828c3fecf58fbd6d5449897f
16	TeslaCrypt	3b246faa7e4b2a8550aa619f4da893db83721aacf62b46e5863644a5249aa87e
17	Thanos	cd0f55dd00111251cd580c7e7cc1d17448faf27e4ef39818d75ce330628c7787
18	Vipasana	1733b199a7063443c167e3caeae7dda2315f590341ea2152a9b132e1ad8e94a8

19	WannaCry	ed01ebfbc9eb5bbea545af4d01bf5f1071661840480439c6e5babe8e080e41aa
20	WannaCry_Plus	2027a053de21bd5c783c3f823ed1d36966780ed4
21	RYUK	f471126e5b750ef9ea6d965b5e6d2145054fcbae29d051934350153f07455c3e

Table 3.2 List of Malware Samples included in Analysis

The detection approach contains a minifilter driver which would send IRP logs to a user level file. Ideally a single set of code should extract the features from the file but in our case, we have manually read the contents of IRP messages and extracted some information manually, some with grep commands and rest of the information is being extracted using a set of scripts. The extracted feature values are then used in a model to create link of ransomware executables with the values of formula.

Chapter 4

Results

4.1 Analysis

Following behaviors of ransomware were observed by literature review stage and malware execution stage. These behaviors are linked to read / write operations and objective of ransomware, which is to encrypt the victim's files. We have proposed a system to be used in cascading configuration with Redemption, as proposed by A Kharraz. The features can be classified as Content-based features and Behavior-based features.

4.1.1 Content-based Features

Content based features refer to indicators which show / quantify the read / write behaviors of a ransomware and legitimate file handling operations.

4.1.1.1 Entropy Ratio of Data Blocks

The purpose of ransomware is to encrypt the files with using some strong encryption algorithm. This is generally implemented by reading data blocks from the target files, sending data blocks to block encryption algorithms, and then writing the encrypted data back to hard drive. The strength of an encryption algorithm is measured based on entropy (randomness) of cipher text[16]. This provides the theoretical basis for the observation that the entropy of write buffer at file offset is always greater than the entropy of read buffer at same file offset because of the increased entropy in ciphered version.

4.1.1.2 File Content Overwrite

Another behavior observed was that ransomware overwrites the plain text with cipher text so that there are no traces left of how the actual data looked like. This reduces the possibility of

file recovery. The process overwriting a file is observed and the percentage of the file content changed is considered directly proportional to ransomware like behavior.

4.1.1.3 Delete Operation on user files

Another observed behavior of ransomware was to read the file, encrypt that file and then write that cipher text to a new file. Again, the purpose of this activity is to make data recovery difficult for the victim. The important data needed for business continuity is normally stored in user owned files. The system continuity data is stored in system owned files and these systems can be recovered quickly considering the modern infrastructure technologies being used by organizations. The number of delete operations on user files is directly proportional to ransomware like behavior of a process.

4.1.2 Behavior-based Features

4.1.2.1 Directory Traversal

Ransomware encryption is a system-wide operation where a process reads the file inventory in directories and encrypt the most files as possible. The count of unique file paths accessed by malicious process is directly proportional to ransomware like behavior.

4.1.2.2 Access to multiple File Types

Ransomware accesses multiple type of files which it is programmed to access which are generally Images, Videos, database backups and a lot of other file types. The access to multiple file extensions is not generally considered malicious but access to high number of file extensions could be considered as an indicator of ransomware operation.

4.1.2.3 Converting to a Specific File Type

The ransomware changes the extension of plaint text files of victims to something which may or may not be specifically attached to specific threat actor like (*.locked, *.lkd, *.enc). IRP_MJ_SET_INFORMATION is the IRP needed to change the extension or rename the file..

4.1.2.4 Access Frequency

Ransomware executables are just designed to read data, encrypt the data and then write the cipher text in the file. Ransomware performs this activity in very fast manner to avoid detection before it finishes up messing with all the data. The time difference between access requests to multiple files is inversely propositional to ransomware-like behavior.

4.1.2.5 Network and Admin Share Access

Admin shares are not accessed by compressors, shredders and encryptors do not access admin shares by default but we have observed that ransomwares always try to access files present in network shares and admin shares.

4.1.2.6 Multithreading

We have observed that ransomware does not use muti-threading to improve speed and efficiency. Instead, ransomware writers go for creating multiple processes to improve efficiency and increased difficulty in blocking the encryption process for the defenders.

4.2 Features

4.2.1 Entropy Change

(r1): The entropy change can be a high weighted parameter in identification of encryption. r1 is calculated an additive inverse of the ratio of the entropy of read buffer to entropy of write buffer for same offset in a file. The range of entropy is between 0 and 1. If the entropy of the write buffer is greater than entropy of the read buffer, the value will be close to 1 and vice versa.

Entropy calculation formula as explained in [17]

$$H(d) = -\sum_{i=1}^{n} \frac{\log_2 n}{n}.$$

4.2.2 High number of Data blocks modified

(r2): If a process changes high number of datablocks in a file, it will receive higher value of r2 which is close to 1. The value is calculated as ratio of number of data blocks modified to total number of data blocks in the file.

Let sA be the number of modified blocks and vA be the total number of data blocks in file, the r2 will be calculated as

R2= vA : sA

A mean value of r2 for all files accessed is used in the final threshold calculation.

4.2.3 Delete Operation

(r3): if a process requests file deletion through IRP_MJ_SET_INFORMATION to more than 100 files, the r3 value is becomes 1.

4.2.4 Privileged File Access

(r4): Privileged access to files, specifically writing on data blocks of file is also an indicator of file overwriting activity matching with ransomware behavior.

Let's say X process is being monitored. We will need to monitor fi which is the ith write request on file f. Then additive inverse of the unique file access is calculated to get a result residing in 0-1 range. Higher the number of write requests to unique files, closer will be the value of r4 to 1.

4.2.5 Access to Different type of Documents

(r5): File accesses to different types of files is observed and monitored based on file extensions to identify cross-document-type file accesses. If a process accesses multiple type of file, specifically greater than 5 unique extensions, r5 is assigned value 1.

4.2.6 Access Frequency

r6): The system calculates the time interval between write access of multiple files. This helps us calculate the access frequency 1/δ. Slower the access to new files, lower will be value of r6.

The overall malice score of a process at time t by applying the weights of individual features:

$$MSC(r) = \frac{\sum_{i=1}^{k} w_i \times r_i}{\sum_{i=1}^{k} w_i}$$

Where if malice score function generates a value greater than 0.12, we determine that the process is malicious.

The weight assigned to different features explained above are as follows,

w1 = 0.9, w2 = 1.0, w3 = 0.6, w4 = 1.0, w5 = 0.7, w6 = 1.0.

Then, we added two features based on our own studies which are to be used in cascading configuration to identify the differentiate between legitimate file handling applications and ransomware based on following criteria

4.2.7 Network and Admin Share Access

If the process has accessed admin and network shares with write privilege, the process is illegitimate or suspicious.

4.2.8 Multithreading

If the process has created multiple child threads, the process is a legitimate file handling application.

Below are some of the screenshots of ransomware notes and notifications obtained during the experimentation

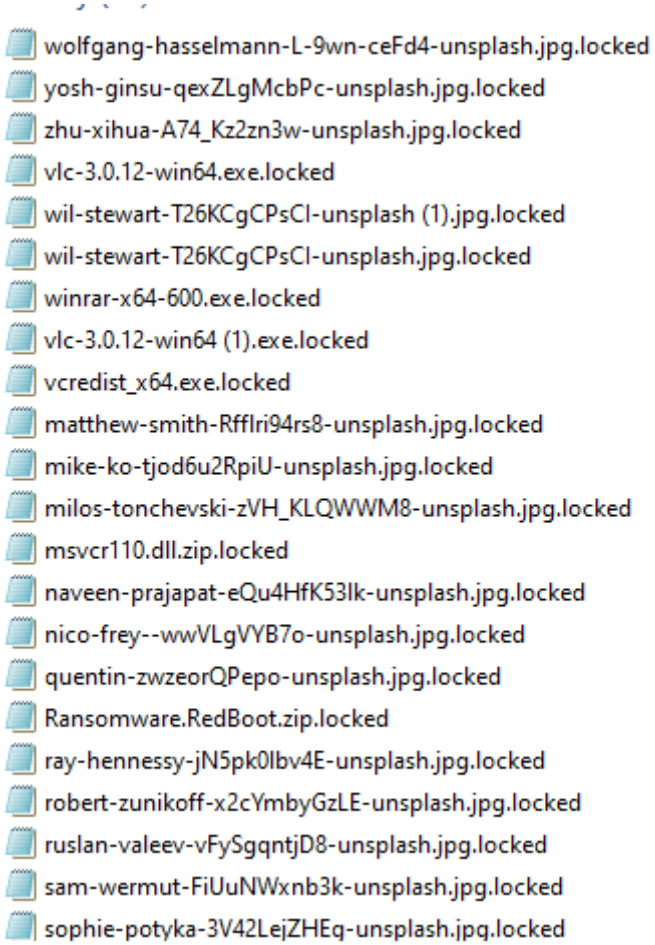

Figure 4. 1: Redboot Encrypted Files

Figure 4. 2: Redboot Ransom Note

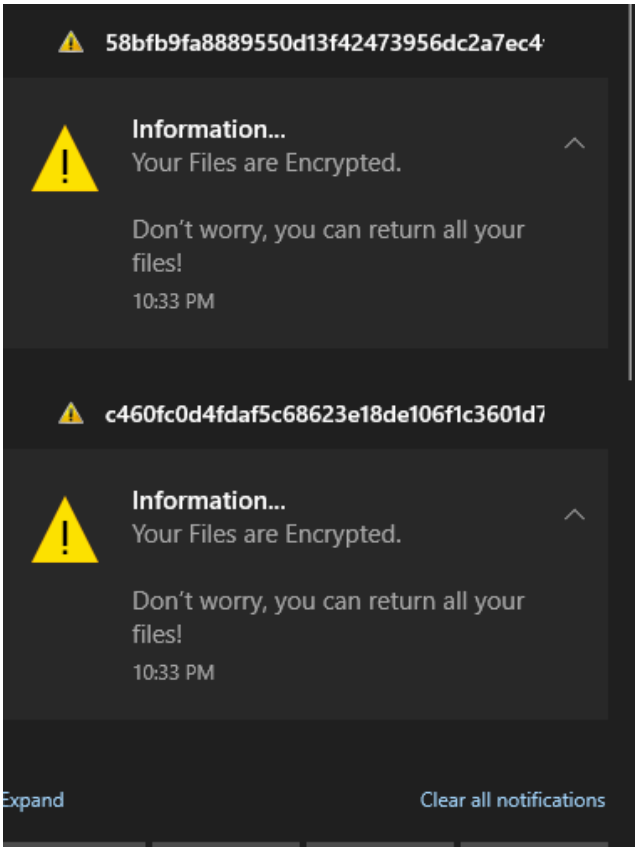

Figure 4. 3: Thanos Ransom Notification

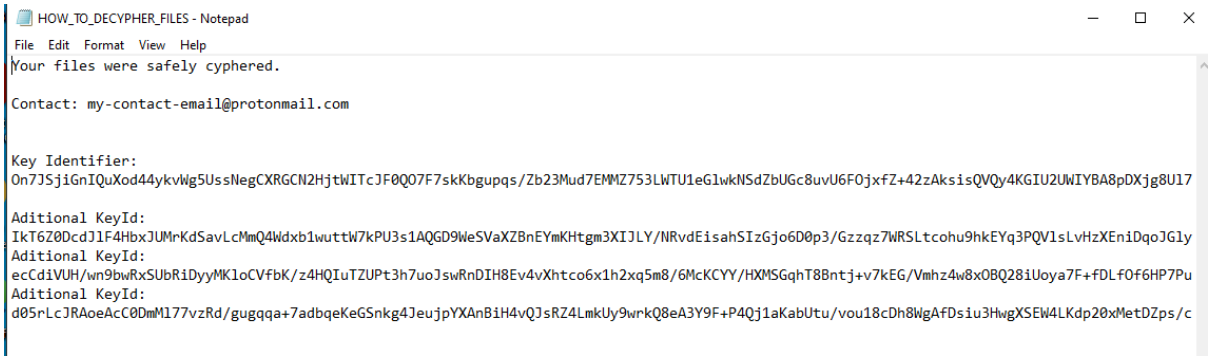

Figure 4. 4: Thanos Ransomware Note

During the investigation, we observed the following pattern of presence of assumed features in the ransomware execution behavior

Executable	Entropy Change	Data blocks modified	Delete Operation	Privileged file access	Access to different type of documents	Access Frequency	Admin share access	Multi-threading	Overwriting Shadow copies
CrypoLocker	P	P	P	P	P	P	P		P*
Cerber	P	P	P	P	P	P	P		P
Cyptowall	P	P	P	P	P	P			P*
JigSaw	P	P	P	P	P	P	P		
Locky	P	P	P	P	P	P			P
Mamba	P	P	P	P	P	P	P		
Matsnu	P	P	P	P	P	P			P*
Petrwrap	P	P	P	P	P	P	P		P
Petya	P	P	P	P	P	P			
Radamant	P	P	P	P	P	P	P		P
RedBoot	P	P	P	P	P	P	P	P	P*
Rex	P	P	P	P	P	P	P		P
Satana	P	P	P	P	P	P	P		
TeslaCrypt	P	P	P	P	P	P	P		P*
Thanos	P	P	P	P	P	P	P		P
Vipasana	P	P	P	P	P	P	P		
WannaCry	P	P	P	P	P	P	P		
WannaCry_Plus	P	P	P	P	P	P	P		P
RYUK	P	P	P	P	P	P	P	P	P*
VeraCrypt	P	P		P	P	P		P	
BitLocker	P	P		P	P	P		P	
DiskCryptor	P	P		P	P	P		P	
7-Zip		P		P	P	P		P	
AxCrypt	P	P		P	P	P		P	
Winzip		P		P	P	P		P	
Eraser	P	P	P	P	P	P		P	
Sdelete	P	P	P	P	P	P		P	

Table 4. 1: Feature Analysis on Data Set

Chapter 5

Conclusion and Future Work

We have conducted a study to find common and differentiable features on kernel level to identify legitimate full desktop encryptor applications and ransomware by analyzing IRPs using a customized minfilter driver, to improve the ransomware detection model developed by Amin Kharraz et.al. The functional objective of both type of applications is same since it both are required to make the target data inaccessible for unauthorized personnel without a key. We researched the pattern of encryption for both applications and were able to identify encryptors from ransomware and hence, participated in the improvement of detection capability of existing models.

From manual and automated analysis of the data, we were able to extract some features which would get high weight while comparing ransomware and legitimate file handling applications.

We have observed that ransomware tried to access shadow copies of files while legitimate file handling applications doesn't not do so. We have also observed that the number of unique files encrypted by ransomware in a directory is lower than the number of unique file types encrypted by compressors or full desktop encryptors, the reason being the ransomware only selects low some specific file which they are programmed to encrypt or supposedly contain data while legitimate applications conduct their operations on all sort of files. It was also noted that most of the malware do not use multi-threading to keep themselves stealthy and create as less noise as they can while compressors or desktop encryptors do not have to hide themselves. They use multi-threading to improve efficiency of their products.

Following streams could be used as future guidelines for work to extend the functionality and improve the results and detection.

- Apply Deep learning to extract unknown features to identify ransomware from legitimate file handling applications
- Enhance the dataset to include more ransomware samples
- Due to increasing trend of data exfiltration before the encryption, the features discussing outbound data transfer should also be investigated.

BIBLIOGRAPHY

[1]*Colonial pipeline boss confirms $4.4m ransom payment*. (2021, May 19). BBC News. https://www.bbc.co.uk/news/business-57178503

[2]JBS: Cyber-attack hits world's largest meat supplier. (2021, June 2). BBC News. https://www.bbc.com/news/world-us-canada-57318965

[3]F.B.I. Director compares danger of ransomware to 9/11 terror threat. (2021, June 5). The New York Times - Breaking News, US News, World News and Videos. https://www.nytimes.com/2021/06/04/us/politics/ransomware-cyberattacks-sept-11-fbi.html

[4]Ransomware: The data Exfiltration and double extortion trends. (2021, April 22). *CIS*. https://www.cisecurity.org/blog/ransomware-the-data-exfiltration-and-double-extortion-trends/

[5] Duc, H. N. (2021, March 5). Ransomware statistics, trends and facts for 2020 and beyond. *Pentestmag*. https://pentestmag.com/ransomware-statistics-trends-and-facts-for-2020-and-beyond/

[6]Sgandurra, Daniele & Muñoz-González, Luis & Mohsen, Rabih & Lupu, Emil. (2016). Automated Dynamic Analysis of Ransomware: Benefits, Limitations and use for Detection.

[7]Nieuwenhuizen, D. (2017). A behavioural-based approach to ransomware detection. Retrieved from MWR Labs website: https://labs.mwrinfosecurity.com/publications/a-behavioural-based-approach-to-ransomware-detection/

[8]Kardile, A. B. (2017). *Crypto Ransomware Analysis And Detection Using Process Monitor* [Unpublished master's thesis]. Tth University Of Texas At Arlington.

[9]Brewer, R. (2016). Ransomware attacks: detection, prevention and cure. Retrieved from Log Rhytm website: https://www.researchgate.net/publication/308736523_Ransomware_attacks_detection_prevention_and_cure

[10]Z. Xu, S. Ray, P. Subramanyan and S. Malik, "Malware detection using machine learning based analysis of virtual memory access patterns," Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017, 2017, pp. 169-174, doi: 10.23919/DATE.2017.7926977.

[11]Kirda, E. (2017). UNVEIL: A large-scale, automated approach to detecting ransomware (keynote). *2017 IEEE 24th International Conference on Software Analysis, Evolution and Reengineering (SANER)*. https://doi.org/10.1109/saner.2017.7884603

[12]Kharraz, A., & Kirda, E. (2017). Redemption: Real-Time Protection Against Ransomware at End-Hosts. Research in Attacks, Intrusions, and Defenses, 98-119. doi:10.1007/978-3-319-66332-6_5

[13] Solomon, D. A., Russinovich, M. E., & Ionescu, A. (2009). *Windows internals* (5th ed.). Microsoft Press.

[14] Lorihollasch. (n.d.). Filter manager concepts. *Developer tools, technical documentation and coding examples | Microsoft Docs*. https://docs.microsoft.com/en-us/windows-hardware/drivers/ifs/filter-manager-concepts

[15] VSC-Service-Account. (n.d.). Minispy file system Minifilter driver. *Developer tools, technical documentation and coding examples | Microsoft Docs*. https://docs.microsoft.com/en-us/samples/microsoft/windows-driver-samples/minispy-file-system-minifilter-driver/

[16] Comparative analysis of different modified advanced encryption standard algorithms over conventional advanced encryption standard algorithm. (2017). *International Journal of Current Research and Review*. https://doi.org/10.7324/ijcrr.2017.9227

[17] Lin, J. Divergence measures based on the shannon entropy. IEEE Transactions on Information theory 37 (1991), 145–151.

[18] Arif, M. (2017). Ransomware forensics (Master's thesis, MCS-NUST, Rawalpindi, Pakistan).