International Journal of Scientific & Engineering Research, Volume 10, Issue 10, October-2019
ISSN 2229-5518

RAPID RESPONSE TO RANSOMWARE

BALOGUN Azeezat Abolaji[1], AJIBOLA Aminat[2], BISALLAH Hashim I[3], EBELOGU Christopher Ubaka[4]

ABSTRACT

Attackers ranging from cyber criminals to state sponsored groups have begun to change their tactics, making use of operating system features, off the shelf tools and cloud services to compromise their victims. Until recently, cyber criminals mainly focus on bank customers, raiding accounts or stealing credit cards. However, a new breed of attackers has emerged and that have bigger ambitions and they are targeting banks themselves, sometimes attempting to steal millions of dollars in a single attack. Ransomware has become one of the most sophisticated cybercrimes as of late. It mode of operations is not limited to individuals but companies and even the government, hospitals and sensitive organizations. Ransomware is a type of malware that infect the targeted system by locking the device or encrypting the files on the system/device and then asking that a ransom be paid before the user can gain access to their device or files, whereby paying such ransom does not guarantee that it been restored. This dissertation looks at establishing a proactive approach towards the detection of ransomware with the use of security tools, configured within a safe and secure environment. The evolutions of ransomware are presented in order to provide a better understanding of ransomware and how it operates. The environment configured within this system uses Windows 10 as the 'victim' machine, since majority of ransomware is known to target Windows platform, because most computer users feel more comfortable at using Windows OS, so this ransomware detection system works for any version of the Windows OS. Therefore a Windows 10 machine running on a virtual box is used to conduct the analysis. Python programming language is used, because it ships with high security modules and it also allows for reverse engineering if further analysis is required. Random forest is the algorithm used to handle the prediction system of the anti-ransomware detection system, while Yara Rules was used in detecting any crypto currency address that might be encrypted in any files. Through the use of controlled environment experiments, it was concluded that the ransomware detection system was successful in detecting characteristics of ransomware. Finally the dissertation concludes by establishing that the ransomware detection system presents a viable approach in tackling ransomware and would prove useful to security professionals in securing their network against ransomware related attacks.

Keywords: *Ransomware, Encryption, Infection, Cybercrime, Cyber Attack, Vulnerability, Malware, Virtual box, Detection System.*

1. INTRODUCTION

The society and the modern world at large continue to embrace internet connected systems at an increasing rate, with a growing number of computer systems being used by individuals and companies. While for many computer users this brings an ease to daily social or working life, it has also brought about a culture of ignorance. Employees are expected to be able to work on a computer and conduct tasks such as accessing the Internet or using email as second nature. Unfortunately, cyber criminals have also embraced the growth in internet connected devices as a global platform to launch cyber-attacks. These acts of cybercrime lead to huge damages in areas of business, healthcare systems, industry sectors, and other fields. Ransomware is considered as a high risk threat, which is designed to hijack the data.

The word Ransomware is a combination of ransom and software; it is a program that is designed to attack a targeted system with the aim of holding the user as a hostage, and restricting users from accessing their devices. It can also be used to encrypt the users' data, forcing the victim to pay the ransom. Generally, ransomware uses malware and Trojan forms to bypass and infect the targeted system. . It can also use cryptographic techniques to lock the system, use the remote command and control channel, and cryptocurrencies [1]. Moreover, ransomware attacks exploit system weaknesses such as Windows SMB (Server Message Block) Remote Code Execution Vulnerability, to get into and lock the system. Ransomware consists of two major types: lockers, which prevent the user having access to the entire system, and crypto ransomware, which only encrypts the user files making the files impossible to use unless decrypted. Removing the ransomware or taking the hard-drive to an unaffected system does not solve the problem as the victim does not have the key to decrypt the data.

Ransomware

Ransomware is computer malware that installs covertly on a victim's computer, executes crypto virology attack that adversely affects it, and demands a ransom payment to decrypt it and if payment is not made threaten to publish such files or delete it. It is a malicious malware program capable of inhibiting a user from accessing their data or computer system. It displays a message to announce what has occurred, shock the user who had no knowledge of an attack taking place. The shock could be more in the message, requesting that payment be made before they can get their files decrypted. The price demanded typically varies by the iteration of ransomware in use and exchange rates of digital currencies at the time of infection, however paying in no way guarantees that the files get decrypted [8].

2. AIMS AND OBJECTIVE OF THE STUDY

The primary aim of this project is to establish a rapid approach to ransomware attack and a thorough literature review of various ransomware iterations, in order to establish the best approach to rapidly response and prevention of spreads of ransomware. In order to achieve this, the following objectives must be met:

- ✓ Design an anti-ransomware system along with this dissertation which will be able to detect malicious malwares, and either prevent them from installing or stop it from being able to cause any harm to the system and its file contents.
- ✓ Evaluate (by collecting sample file types and programme type for the computer system to watch) the early symptoms of the presence of a ransomware in a device and rapid response to thwart it from locking the user out of the device.

3. LITERATURE REVIEW

The first crypto ransomware was probably the infamous AIDS Trojan in 1990. It was distributed, on a floppy disk handed out to people who attended an international conference on AIDS disease, and the software file names were encrypted, not the files themselves, and then displayed a demand for payment to a location in Panama. The attacker's motive might have been a desire for vengeance on the organizers of the conference rather than in financial gain, but the attack was ineffective. The exact reason for this wasn't published, but a program for restoring the file names was quickly distributed. For this to have been a successful attack, it must follow some basic principles which would be discussed later in the chapter. The software did a very poor job of hiding the key, and that was the basis for the restoration program. It was needless for anyone to pay the ransom [7].

To turn crypto ransomware into a truly daring attack, there were two more pieces of technology needed, asymmetric cryptography and anonymous payment. Asymmetric cryptography had been invented two decades earlier and was readily available through Pretty Good Privacy (PGP) software and the GNU Multiple Precision (MP) library; it didn't gain much attraction with the malware crowd. This was odd, because in 1996, Adam Young and Moti Yung published a paper describing exactly how to do this. Their method involved generating a unique symmetric encryption key for each infected computer and then encrypting that with a master public key embedded in the virus software [7]. The beauty of their method was that the infected machine didn't need to communicate with the perpetrator until the ransom was paid. At that time, the victim could post the public key encryption of the symmetric key, and the perpetrator could decrypt that and send the symmetric key back to the victim for decryption of the files.

Malware authors didn't pick up on this scheme for about 10 years. Maybe they didn't trust the anonymity or security of the keys, or maybe they were wary of collecting payments. Although scams that took advantage of international banking were common, ransomware faced more difficulties to remain hidden. In an ordinary scam, the victims were unlikely to realize their mistake for several days, but with ransomware, the victims would be calling law enforcement immediately, and the bank account would be tracked or shut down quickly [6]. To reliably evade detection, the perpetrators needed anonymous payment which came to their aid in the form of Bitcoin, which would be further discussed later in the chapter.

Ransomware went out of fashion in the late '90s and begin to return to prominence until 2005. The availability of more complex encrypting schemes helped usher in the new era of ransomware, which has continued to accelerate. As of 2016, it is considered to have become one of the most prevalent forms of attacks against computer systems, requiring limited exposure to vulnerabilities and minimal exploration on target.

A. Characteristics of Ransomware

The computer's files, document drafts, contact lists and so on, all have been transformed into encrypted data and only the encryption keys will undo the damage. From a technology perspective, a successful ransomware must meet some requirements, which are:

- Some resources that is valuable to the user must be made unavailable, i.e., denial of service.
- The denial of resources and the payment instructions must be announced to the user of the infected machine in an inevitable, visible process.
- The ability to restore the valuable resources must depend on a small amount data that is available only to the extortionist and cannot be inferred or calculated by other process at reasonable cost
- The extortionist must be able to verify payment.
- The extortionist must be able to supply information for restoring the resources without identifying himself
- The restoration process must run on the infected computer; it must be reasonably reliable.

B. Targets for Ransomware

The cybercriminals behind ransomware do not particularly care who their victims are, as long as they are willing to pay the ransom. With this in mind, it is easy to see why the cybercriminals tend to take a scatter-gun approach to propagating the ransomware, casting a wide net across targeted regions and types of users. With the cybercriminals hitting millions of users worldwide, if even a small percentage of victims pay the ransom, then it could make the scheme worthwhile. This is why our default recommendation is not to pay the ransom.

- ***Home users***

Ransomware is perhaps the most effective against individuals who are not fluent with computers or are not familiar with ransomware and how it works. The most common group that we see impacted by ransomware is the home user, who often has the least amount of access to technical assistance. The lack of support may leave the user feeling isolated and helpless, further increasing the pressure to pay.

Home users often have sensitive information, files, and documents that are personally valuable stored on the computer, such as college projects, photos, and video game save files. Despite these things being of value to users, home users are still unlikely to have an effective back up strategy in place to successfully recover from events such as a fire or theft, let alone a crypto ransomware attack [1].

- ***Businesses***

For many businesses, information and the technology to use it is their life blood, without which the act of conducting day-to-day business is impossible. Consider a retailer running a computerized point-of-sale (POS) system. If the POS system was unavailable due to a ransomware infection, the retailer would not be in a position to transact sales. Business computers are also more likely to contain sensitive data and documents of critical importance, such as customer databases, business plans, proposals, reports, source code, forms, and tax compliance documents.

- ***Public agencies***

Public agencies such as educational institutes and even law enforcement entities are not excluded from the attention of these cybercriminals and in some cases, they may be specifically targeted.

- ***Personal computers***

The vast majority of ransomware threats today are designed to target personal computers running the Windows operating system. This is unsurprising, as Windows-based computers make up around 89 percent the OS market share for desktop computers, with Mac OS X and Linux making up the rest. Given that ransomware is a commercial activity for cybercriminals, it makes sense for them to maximize potential returns on their investments [8]

- ***Mobile devices***

The next most targeted types of devices are tablets and mobile phones. These devices have become ubiquitous worldwide, with studies showing that users are spending more time on mobile devices than ever before. Ever since the advent of the iPhone back in 2007 and Android in 2008, smartphone and tablet device ownership has been on a steep upward trajectory. Today, there are basically just two main players in the mobile OS market: Android and iOS.

Android has a massive global footprint, with a share of over 80 percent of the mobile market, representing billions of smartphone and tablets worldwide. In terms of the malware landscape, there is a world of difference between the Android and iOS world.

C. How Ransomware Spreads

- **Phishing / Spam Email**

Ransomware is most commonly spread via email, usually either in the form of mass spam messages, or more carefully crafted messages tailored to the recipient (also known as phishing). The email messages include a file attachment, which is most commonly a Microsoft Office document, a zip file or an executable program. If the user opens the attachment, it will run code that installs and launches the ransomware on the machine.

- **Exploit Kit**

Exploit kits are another common distribution method. These are toolkits that are planted by attackers on a website. In some cases, the site is deliberately created for malicious use, while in others; the site is a legitimate one that has been compromised. Once installed, the exploit kit probes the devices of each website visitor for any vulnerability that can be targeted. If a flaw is found, the kit exploits it to download ransomware onto the device

- **Drive-By Downloads**

This occurs when a system automatically downloads a piece of malware or spyware without the knowledge of the end user.

4. METHODOLOGY

A ransomware model detector was created (utilizing random forest in Machine Learning), in an endeavour to portray all variations of every group of ransomware into one model. The procedure included the development of a classifier (to parse, classify and output graphs detailing the behavioural constructs of a ransomware), as well as creating a safe environment to analyse the ransomware samples.

The implementation and experiments that this project considered follow the methodology that is outlined.

- Identification and Definition of the Problem
- Accumulation of Relevant Data
- Formulation of Hypothesis
- Experimental Conduct
- Summary and Conclusion

A. Design

Tools/Technologies Used

- A solid computing processor
- Python Programming Language
- PyQt
- Virtual Box
- Random Forest (For predicting)
- Yara Rules

In this section, data collection, experimental protocols, evaluation measures will be discussed.

B. Experimental Setup

In the designing of the model, Virtual Box was used as the guest analysis machine and Windows as the virtual Operating System (OS). Before launching the Windows OS, it was pre-installed with all necessary software, such as Python, PyQt, Microsoft Office, etc. This gives the OS controlled access to the Internet via NAT adapter. Outbound traffic to any other machine in the local environment was blocked so that other machines in the network are not affected. The Windows firewall was also disabled, and no other processes were running while executing the ransomware.

The server runs simultaneously with the ransomware execution and retrieves all the changes done inside the guest machine in form of report files. Multiple report files are collected for each binary. The output data are stored in a specific directory format. Figure 4.1 shows the directory structure for storing all the output files.

The different types of report files generated are described as below:

- ***analysis.log***

This is created only by the analyzer when execution runs inside the guest machine. It gives feedback of creating processes, files and any errors that might have happened while execution took place.

- ***dump.pcap***

This is where we have the entire file that contains all the information related to network traffic. This file is created by a network sniffer and was used for further analysis.

- ***memory.dmp***

This file contains the memory dump of the guest analysis machine (Virtual Box).

- ***Logs/***

This directory contains all the logs generated.

- ***reports/***

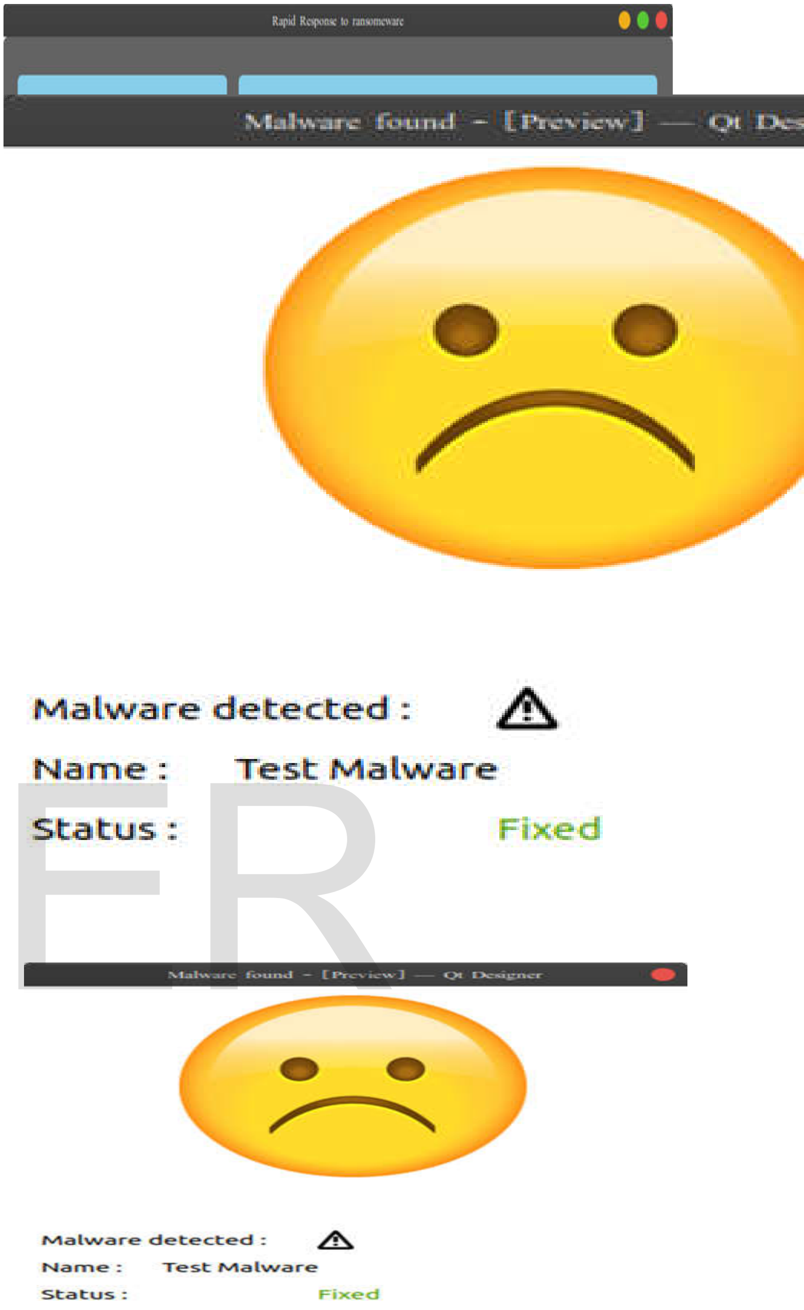
Based on the configurations, this directory contains all the reports.

5. SYSTEM EVALUATION AND ACTION FOR RANSOMWARE

The purpose of this is to obtain ransomware samples, run them within the environment and establish behaviour that can be transferred to the detection system. It has been established that ransomware can take varying times to complete its operations. Hence, there will be need for continuous update and upgrade of tools used for this development.

The first recommended action during ransomware attack is the most important, which is that the system must be turned off in order to be disconnected from the hackers' server. This action will prevent the ransomware to pass to other connected devices and networks. After the system has been turned off, the safe mode option must be used in rebooting the system. This option allows only default programs of the operating system to be operated which will fix critical problems in the system. While at this, it is highly recommended not to delete the ransomware files in the system before it is recognized, because taking this action by a non-expert might cause damage to the system files, and potential data loss due to interrupting the connection with attackers. Therefore, this procedure has to be taken very cautiously.

Furthermore, once the ransomware program has been terminated, the next step would be dealing with system file recovery to get back the infected files. This can be performed manually with the windows recovery system. The recovery programs are used in the recovery of any deleted files by the ransomware. In many cases the recovered files are encrypted with ransomware infection. Therefore, it is required that the decryption key is revealed. Ransomware attacks use different encryption key, and technique to infect the data.


Rapid Response to ransomware
Malware found – [Preview] — Qt Des
Malware found – [Preview] — Qt Designer

Malware detected : ⚠
Name : Test Malware
Status : Fixed

Malware detected : ⚠
Name : Test Malware
Status : Fixed

Figure 5.2: Successful run of model with the expected prompt message

6. RESULTS, ANALYSIS AND DISCUSSION

Dataset Information

After the extraction of the features from all pcap files, the dataset for training the machine learning classifier was prepared. The data that were extracted from the pcap files using the scripts are in text format. Each row in the data is a flow represented by the tuple followed by corresponding feature values. The dataset has 15,524 rows of data and 30 columns, out of which 13,058 samples are normal and 2,466 samples, are ransomware. This means the percentages of normal and ransomware data are approximately 85% and 15%, respectively. Table 1 shows the counts of flows for each individual family.

Family	Flows Count
Normal	13058
TeslaCrypt	1388
Locky	459
Cerber	292
CTBLocker	164
Win32.Blocker	48
Spora	20
Sage	15
Crysis	14
Unlock26	11
CryptoShield	10
GlobeImposter	8
Mole	7
Jaff	6
Xorist	4
Petya	4
Satan	3
Striked	3
Zeta	3
WannaCry	2
Flawed	2

Table 6.1: Dataset

A. Performance Metrics

The following performance metrics was used in the evaluation of the detection model.

$$Accuracy = (tp + tn)/(tp + tn + fp + fn)$$

$$Detection\ rate = (tp)/ (tp + fn)$$

$$False\ Potive\ Rate = (fp)/ (fp + tn)$$

Where fp is false positive, tp is true positive, tn is true negative and fn is false negative.

False negatives correspond to the ransomware samples that are classified by the detection system. False positives are the normal samples that are classified as ransomware. Since the thesis is about rapid response to ransomware detection, detection rate (DR) and false positive rate (FPR) are very important. The performance of the model can be expressed by using the Receiver Operating Characteristic (ROC) curve. The ROC curve is generated by plotting the False Positive Rate (FPR) on x-axis versus True Positive Rate (TPR) (also known as Detection Rate) on y-axis. ROC curve is plotted in the range of [0, 1] for both axes. Any machine learning classifier has a function whose values decide the class outcome. The output of the function can be viewed as probability outcome in range of [0, 1]. The ROC curve plots the FPR and TPR values for varying the value of threshold between [0, 1].

B. Results

The classification algorithms used in the current thesis is Random Forest (RF). Random forest is a collection of decision trees that can help limit over-fitting. Table 4.2 shows the accuracy results for Random forest with and without SMOTE for the 'train_test_split' set. Similarly, Table 4.3 shows the accuracy for 10-fold cross validation.

Classifier (%)	Accuracy%	
	Without SMOTE	With SMOTE
Random Forest	100	100

Table 6.2: Accuracy score for the classifiers for 'train_test_split' set (with and without SMOTE)

Classifier (%)	Average Accuracy%	
	Without SMOTE	With SMOTE
Random Forest	98	100

Table 6.3: Accuracy score for the classifiers for 10-fold cross validation set (with and without SMOTE)

Classifier	Without SMOTE		SMOTE		SMOTE+ Hyperparameter tuning	
	FPR (%)	DR (%)	FPR (%)	DR (%)	FPR (%)	DR (%)
Random Forest.	0.2	99.8		99.9	0	99.9

Table 6.4: Detection rate (DR) and False Positive Rate (FPR) for the classifiers over test dataset.

7. SUMMARY AND CONCLUSION

The main aim of this research was to develop a greater understanding of ransomware through literature, provide a rapid response to ransomware attack and develop suitable methods detection using open source security tools that are currently available. The detection system setup was inspired having reviewed the thesis of Sivasubramanian Nambivelu, with the vision of improving upon his proposed methods [5]. The work has been updated to take into account the newest ransomware families currently active.

Developing systems like the one proposed within this paper should be something for security professionals to consider, giving as it is better to be proactive with securing data, rather than reactive in trying to recover from a ransomware attack. Thus, the designs are replicable, extensible and also scalable depending on what system the system drives are connected to, moreover, the systems meet the aim of the project.

In this research, a ransomware detection framework leveraging techniques of reverse engineering, static analysis, and machine learning was proposed. The ransomware detection rate is significantly higher for the dataset with combined feature dataset maximum of 97.95% for the Random Forest. The contributions of this work can be summarized as follows:

- Designing of an automatic static analysis framework to improve the detection accuracy of ransomware.
- The average ransomware detection accuracy was 92.11% while considering both individual and combined feature dataset for all given supervised machine learning classifiers. In the future, there is going to be a plan to have more number of ransomware samples and perform the experiment with other machine learning algorithms including deep learning.

8. RECOMMENDATION

Nowadays, cyber-attacks are going global which affects variety of organizations and endpoint users. While, hackers use different approaches and tools including ransomware threats to take over the targeted systems which might leads to cause a huge damage such as in business, healthcare system, industry sectors, and other fields. While ransomware isn't going away any time soon (if ever), you can defend your organization, if you're properly prepared." Under the light of this idea, this paper worked to understand the common form of ransomware threats due to identifying the root and types of ransomware and diagnosing effective types over different platforms. Also, how ransomware works in the systems and possible changes which can be made by ransomware. In addition, ransomware uses combined algorithms to make more sophisticated threats which make it harder to decrypt until the ransom will be paid.

In contrast this research shows common solutions and approaches to remove ransomware codes from the entire system and networks. In conclusion, in order to stop ransomware attacks there are some certain steps that should be taken such as having regular back up, patching software, and some other outstanding techniques illustrated to prevent ransomware attacks.

REFERENCE

[1] Abrams, L. (2016) "Emsisoft releases Decrypter for the LeChiffre Ransomware," Bleeping Computer, 25 January 2016 [Online].

[2] McAfee Labs, 2015a. McAfee Labs Threats Report, (November). Available at: www.mcafee.com/us/mcafee-labs.aspx.

[3] McAfee Labs, 2015b. Meet "Tox": Ransomware for the Rest of Us - McAfee. Available at: https://blogs.mcafee.com/mcafee-labs/meet-tox-ransomware-for-the-rest-of-us/ [Accessed February 7, 2016].

[4] McAfee Labs, 2016. Understanding Ransomware &amp; Strategies to Defeat it.

[5] Nambivelu, S., 2015. Analysis and Detection of Ransomware. Edinburgh Napier University O'Brien, D., 2016. Special Report : Ransomware and Businesses 2016. Symantec Corp, pp.1–30.

[6] Sikorski, M. and Honig, A. (2012). Practical Malware Analysis: The Hands-On Guide to Dissecting Malicious Software. No Starch Press Series. No Starch Press.

[7] Young, A. & Yung, M., 1996. Cryptovirology: extortion-based security threats and countermeasures. Proceedings 1996 IEEE Symposium on Security and Privacy, 5111(C), pp.129–140. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=502676.

[8] Zavarsky and D. Lindskog, "Experimental Analysis of Ransomware on Windows and Android Platforms: Evolution and Charac