RansomDef: A Modular Approach in Defence against Cryptographic Ransomware Attacks using Machine Learning

[1]Prof. Vishal Shinde, [2] Mr.Upanishadh Prabhakar, [3] Mr.Sahil Patel, [4] Miss.Khyati Selani,

[1]Asst.Professor, [2,3,4]UG Student, [1,2,3,4]Department of Computer Engineering, Shivajirao S. Jondhle College of Engineering & Technology, Asangaon, Maharashtra, India.

[1]mailme.vishalshinde@gmail.com, [2]upa1408@gmail.com, [3]sahil.sp770@gmail.com, [4]khyati.selani03@gmail.com.

Abstract- Cryptographic ransomwares have been wreaking havoc across the extended cyberspace throughout the world leading to millions of dollars lost either in paying ransoms or in recovering the data lost in the attack by other means. In such cases the malicious software encrypts various user files and uses them as a leverage to extort ridiculous amounts of money to decrypt the files. Ransomware use various hiding techniques ranging from code obfuscation to creating new polymorphic variants which aid them in evading signature based methods deployed by Antivirus software. The idea presented here works on basis of analysis of an extensive dataset of Ransomware families. The result is simple yet effective software - RansomDef, a modular defence system against cryptographic ransomware. It follows an intuitive approach, a conjunction of combined static and real-time analysis to generate an exhaustive set of traits that characterize Ransomware behavior. Early detection is aided because of implementation of a robust trap layer. [1] Zero day intrusions are exposed by machine learning. When initial layers of RansomDef detect a process for suspicious Ransomware behaviour, the corresponding system process is killed by RansomDef before encryption activities are executed on any of the user files present in the system..

Keywords- Machine Learning, Ransomware, Encryption.

I. INTRODUCTION

Cybercrime is one of the most commonly occurring forms of crime in today's era. The evolution of technology has increased the occurrence of cybercrime manifold has its effects in modern day corporate and personal environments is ginormous. An upcoming & fast spreading form of Cyber-attacks is in the form of Ransomware. Ransomware is a subset of malware in which the data on a victim's computer is held captive, usually by encryption, and payment is extorted before the ransomed data is decrypted and possession is returned to the victim. The intent for ransomware attacks is nearly always of a fiscal nature, and not like other types of attacks, and the victims are usually informed that an exploit has occurred and are given imperatives regarding how to recover from the attack. Payment is often demanded in via virtual currency, such as bit coin, so that the cybercriminal's identity isn't known. Ransomware has grown exponentially to become the most dangerous and threatening from of malware in recent times. The Ransomware attacks are not limited to a specific sector but harass people from many dynamics including personal, Corporate, finance, banking & real estate to name a few. The solution to Ransomware attacks provided here is **RansomDef**. [1]RansomDef is a layered defense system for protection against Cryptographic Ransomware. Each layer in RansomDef has a predefined function. The organization of layers is in accordance to the computation of the features that are generated during a sample test execution.

II. AIMS AND OBJECTIVES

a) Aim

RansomDef aims at providing a novel and comprehensive set of features required in detecting and stopping ransomware attacks in computers. It also aims to improve ransomware detection through machine learning and secure the user data from daily mutating natures of ransomwares.

b) Objective

1. Identify a novel comprehensive set of features that identify with those of Cryptographic ransomware behaviour.

2. Create a Strong Trap Layer that helps in early detection

3. Use Machine Learning for exposing zero-day intrusions

III. LITERATURE SURVEY

Cryptographic ransomware, is a niche ransomware which doesn't delete or hide user or system data unlike many other mainstream viruses and/or malwares. Rather it is a malware which deprives the user access to his files in some way, and forces the user to pay some kind of remuneration in order to lift the restrictions put upon the files. RansomDef has been pre-trained to detect some of the well-known ransomwares in today's cyber world. Some of them are analyzed and explained in the following paragraphs.

Locky: Locky is a ransomware that spreads through malicious documents. These documents are mostly .doc, .xls or zip files to benign looking spam emails.[2] These files contain macros which look like scrambled or unreadable text. Once Locky has started execution it one by one infects files on your computer. Infected files stop having user access and get encrypted with a unique 16 character long alpha-numeric combination with varied extensions.[3]

Diablo and Lukitus: Diablo and Lukitus are variants of the Locky ransomware with some minor changes which help them fly undetected in the anti-virus's radar. They spread through the Necurs botnet, one of the most widely used botnets today.

Cerber: Cerber's reach is far, as this is one of the most prominent ransomware-as-a-service platforms. Cerber has many variants. All variants have some subtle difference between them but their general modus operandi remains the same. The ransomware spreads via email with a JavaScript extension.[4] When executed, the script establishes a connection to the internet and downloads the actual payload that encrypts the data. The ransomware then attacks the files and changes the desktop wallpaper to a ransom note.

CryptoWall: Cryptowall uses a Java vulnerability to launch itself in the victim's PC. The threat typically arrives on the affected computer through spam emails, exploit kits hosted through malicious sites or misleading click-baits, or other malware. [5]Once the Trojan is executed on the compromised computer, it creates multiple registry entries that store the file paths to the encrypted files which run each time the computer is restarted. Another way CryptoWall spreads itself is through exploit kits hosted on various websites scattered throughout the internet.[6]

IV. COMPARATIVE STUDY

Table 4.1: Comparative Analysis of several well-known ransomwares.

Sr no	Name	Origin	Symptoms of attack	Encryption Type	Attack Vector	Decryption Status
1.	Cerber	Unknown but research indicates a Russian origin.	A notepad file gets stored with instructions to unlock the data in all the infected folders.	RSA-2048 key (AES CBC 256-bit encryption)	Via emails through a security hole in JavaScript execution	Decryptors for some early iterations are available but newer versions are still encrypted.
2.	Locky	Best guess by analysts- Dridex Gang	All files are renamed with ".locky" extensions and computer's wallpaper is changed to ransom note.	RSA-2048 + AES-128 cipher with ECB mode used to encrypt files.	Usually via .doc, .xls, or .zip files though other attack vectors have been found.	Original locky ransomware has been decrypted by the trend micro decryptor.
3.	Diablo	Unknown	Infected files have ".diablo" or ".diablo6" extension and a "diablo.htm" file is found in all infected directories.	RSA-2048 + AES-128 cipher with ECB mode used to encrypt files	Necurs Botnet	Decryption not yet possible.
4.	Lukitus	Unknown	Infected files have ".Lukitus" extension and a "lukitus.htm" file is found in all infected directories. Computer's wallpaper is also changed.	RSA-2048 + AES-128 cipher used to encrypt files	Necurs Botnet	Decryption not yet possible.

5.	Crypto wall	Created by developers of Cryptodefense ransomware	Computer’s wallpaper is also changed to a weird text background stating that the data has been encrypted and instructions to download decryptor. “HELP_DECRYPT” file is created in the computer.	Strong 2048-bit RSA public key encryption	Spreads via exploiting a vulnerability in java or by many exploit kits spread across the internet like Rig exploit kit and Nuclear exploit kit.	Decryption not yet possible without private and public keys.

V. PROBLEM STATEMENT

Ransomwares are special case malwares which don’t destroy user data but rather make it inaccessible to the user via encryption techniques. Such special cases require a different approach at safeguarding systems from ransomware attacks. The techniques in deployment are static and are unable to adapt to the lofty security standards cyber protection applications these days must uphold in order to provide a robust and trustworthy defense mechanism The methods used these days mostly tackle viruses and Trojans that destroy or corrupt user data altogether rather than holding the data hostage and demanding a ransom from the user – potentially creating a cyber-kidnapping situation.[1] Static methods are prone to code obfuscation, code wrapping, polymorphic nature of software and many other loopholes that are being exploited by cyber criminals to attack computer systems. In addition to this, anti-viruses are not able to defend against upcoming ransomwares due to code and/or signature spoofing. These shortcomings are kept in mind while designing RansomDef so that the resultant software offers an intelligent, evolving, dynamic and robust line of defense against potential ransomware attacks.

VI. PROPOSED SYSTEM

RansomDef is a layered defense mechanism against ransomware. Each layer in RansomDef has a specific predefined function which it performs. The layers are organized in a strategic manner with the objective that ransomwares will be detected as early as possible. Detecting Ransomware’s malicious behavior early on will lead to detecting the presence of Ransomware in the system at the earliest. The various layers of RansomDef are described below in the same order as they occur in the actual software:

1.Honeypot Files and Trap Layer: A thorough behavioural analysis of Cryptographic Ransomware reveal that they perform various activities to break defences of the victim system and encrypting the user data. The function of this layer is to set traps for the ransomware by tracking occurrence of their malicious activities. Cryptographic Ransomware performs encryption of user data files with specific extensions. Honey files are like baits for ransomwares to attack. These files have attributes similar to the ones mostly attacked by ransomwares. Such honey files and directories are placed in strategic user data folders. These locations are selected such that the files stored in that location are not expected to be modified during ordinary day to day operation of the system. Modification of these files are a pretty substantial indication that some malicious activity or a ransomware attack in this use case is taking place.

2.Dynamic Analysis Engine: Static features are insufficient in detecting the ever-evolving threat of ransomware. Code obfuscation,varying encryption techniques and packing are some of the reasons making detecting ransomwares on basis of static analysis alone is difficult. Processes that enter the system are scanned in real time, during execution by this layer. [7] Cryptographic Ransomware performs extensive encryption of user data files. This layer monitors the file system operation that execute in the system. It also looks out for the massive entropy level fluctuations that occur in the system during extensive encryption process.

3.Machine Learning Engine: This layer is used to build a baseline model for detecting zero-day ransomware attacks. [8] It takes outputs of the trap layer and dynamic analysis engine as the input to create a set of indicators which help in classifying a sample executable as malicious or benign. The machine learning layer is trained to identify ransomwares using supervised, offline training algorithms. The data used to train the AI is bifurcated using harmful and benign tags. This data consists of feature sets determined from the data collected from the previous layers. This learned model used by the Machine Learning Engine to classify executables in real-time based on their input feature sets.

VII. ALGORITHM

```
Step 1: Set Target file for Encryption.
FileInputStream inputStream ← new
FileInputStream(inputFile);
byte[] inputBytes ← new byte[(int) inputFile.length()];
inputStream.read(inputBytes);
Step 2: Initiate encryption using AES.
String TRANSFORMATION ← "AES";
void encrypt(String key, File inputFile, File outputFile);
doCrypto(Cipher.ENCRYPT_MODE, key, inputFile,
outputFile);
Step 3: Generate Encryption Key
Key secretKey ← new SecretKeySpec(key.getBytes(),
ALGORITHM);Cipher cipher ←
```

```
Cipher.getInstance(TRANSFORMATION);cipher.init(cipherMode, secretKey);
Step 4: Output encrypted file.
FileOutputStream outputStream ← new FileOutputStream(outputFile);
outputStream.write(outputBytes);
```

Working:

The basic modus operandi here is regarding how a potential ransomware would encrypt data. It will identify a file to encrypt. Next the AES algorithm will accept the file as an input and apply its encryption standards on it. As a final step, the algorithm will give an encrypted file as the output to the attacked victim and it will forward the decryption key to the attacker to use as leverage for extortion.

Detection:

```
Generating honey files to use as bait.
HoneyCombFileNameGenerator {
getAlphaNumericString(int n);
generate AlphaNumericString ←
"ABCDEFGHIJKLMNOPQRSTUVWXYZ"+
"0123456789"+ "abcdefghijklmnopqrstuvxyz";
l← len(AlphaNumericString);
Generate buffer sb sizeof(l);
Fill buffer with random string name.
For each item i in buffer do generate a random number between
0 to l
int index ← (int)(l * Math.random());
append Character one by one in end of sb
 return sb;
 }
```

VIII. MATHEMATICAL MODEL

Suppose similar signatures are used as a metric for classifying the numerous ransomware attacks that occur on a system. The classifications for ransomware attack are susceptible (S), or infected (I). Once the malicious objects enter into the network, the files become susceptible (S) and after a certain time delay the files become infected (E) and then it gets infectious (I). After it gets infectious, anti-malicious software is run which helps the files to recover (R) temporarily from the attack and provide temporary immunity to the files in the network.[9]

$dS/dt = A- \beta SI – dS + \varepsilon R$

$dE/dt = \beta SI – dE – \alpha E$ (1)

$dI/dt = \alpha E – (d+\delta) I – \gamma I$

$dR/dt = \gamma I – dR – \varepsilon R$

Where $N (t) = S (t) + E (t) + I (t) + R (t)$

IX. SYSTEM ARCHITECTURE

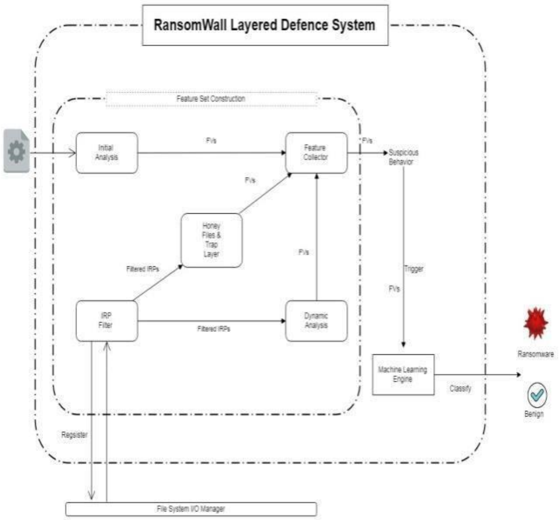


Fig.9.1: System Architecture

Description: The overall system can be divided into three functioning blocks:

1. Feature Set Collection
2. Machine Learning Engine
3. File System I/O Manager

Feature Set Collection:

The main block is to create datasets consisting of feature values to be sent to the machine learning layer. [10] The processes entering the system are monitored in regards to their data requests, CPU cycle usage etc. The read and write requests made by the processes are tracked with help of Input Output Request Packets (IRP) which gives information if any process has set off the honey files trap. This data is collected and consolidated into a feature set or feature values. In addition to the honey pots the IRPs' are also sent to the dynamic analysis engine which also contributes to the data sets.

Machine Learning Engine: The feature values which are fed to this engine are used to determine if a process is malicious or benign. [8] This block uses a generalized learned model to determine the nature of a process.

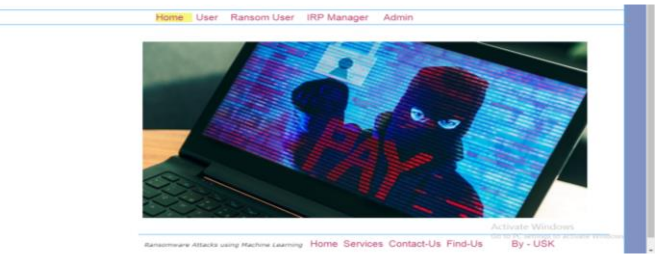
File System I/O Manager: File System I/O Manager is mainly used to feed IRPs' to the Feature Set Collection module to supply information about a process's disk read/write access requests.

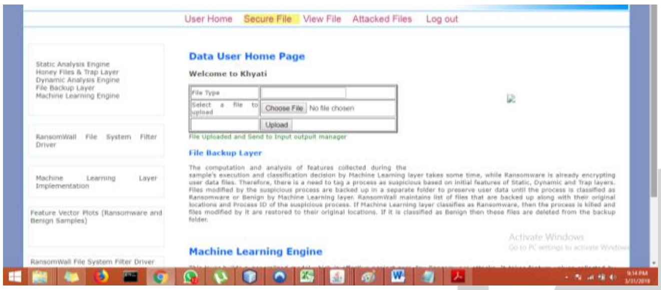
X. ADVANATGES

1. RansomDef can be deployed in a variety of environments to detect and prevent ransomware attacks.
2. In personal spaces and private computers RansomDef can be installed to prevent user data from being attacked by a ransomware.
3. In commercial and corporate sectors banking sectors, finance sectors and virtually every area which makes use of data consolidation and utilization, RansomDef can be

deployed to secure the data from Ransomware attacks which can cause catastrophic losses on an operational and financial level to corporate sectors.

XI .DESIGN DETAILS





Data User Home Page

Welcome to Khyati

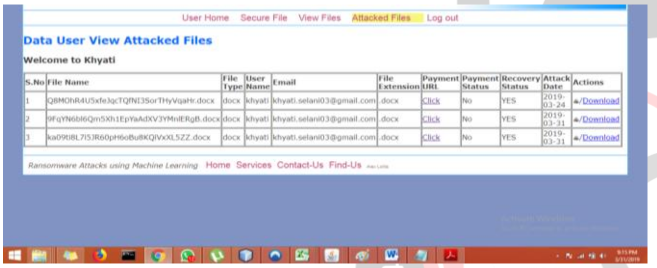
File Type	
Select a file to upload	Choose File No file chosen
	Upload

File Uploaded and Send to Input output manager

File Backup Layer

The computation and analysis of features collected during the sample's execution and classification decision by Machine Learning layer takes some time, while Ransomware is already encrypting user data files. Therefore, there is a need to tag a process as suspicious based on initial features of Static, Dynamic and Trap layers. Files modified by the suspicious process are backed up in a separate folder to preserve user data until the process is classified as Ransomware or Benign by Machine Learning layer. RansomWall maintains list of files that are backed up along with their original locations and Process ID of the suspicious process. If Machine Learning layer classifies as Ransomware, then the process is killed and files modified by it are restored to their original locations. If it is classified as Benign then these files are deleted from the backup folder.

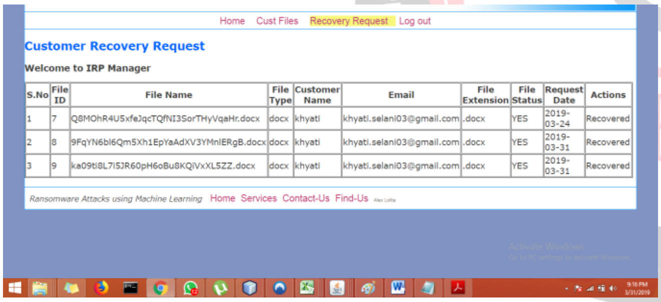
Machine Learning Engine



Data User View Attacked Files

Welcome to Khyati

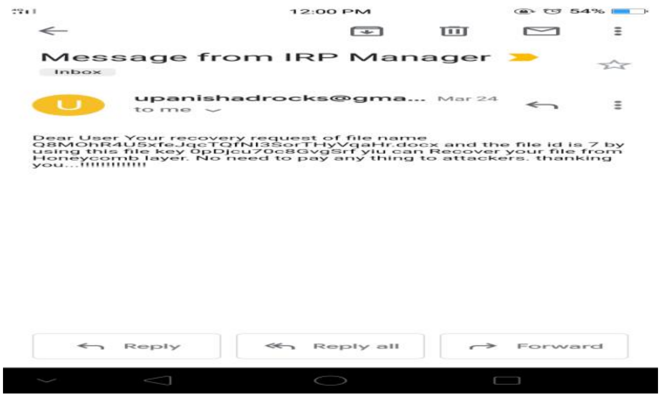
S.No	File Name	File Type	User Name	Email	File Extension	Payment URL	Payment Status	Recovery Status	Attack Date	Actions
1	Q8MOhR4U5xfeJqcTQfNI3SorTHyVqaHr.docx	docx	khyati	khyati.selani03@gmail.com	.docx	Click	No	YES	2019-03-24	/Download
2	9FqYN6bl6Qm5Xh1EpYaAdXV3YMnlERgB.docx	docx	khyati	khyati.selani03@gmail.com	.docx	Click	No	YES	2019-03-31	/Download
3	ka09ti8L7i5JR60pH6oBu8KQiVxXL5ZZ.docx	docx	khyati	khyati.selani03@gmail.com	.docx	Click	No	YES	2019-03-31	/Download



Customer Recovery Request

Welcome to IRP Manager

S.No	File ID	File Name	File Type	Customer Name	Email	File Extension	File Status	Request Date	Actions
1	7	Q8MOhR4U5xfeJqcTQfNI3SorTHyVqaHr.docx	docx	khyati	khyati.selani03@gmail.com	.docx	YES	2019-03-24	Recovered
2	8	9FqYN6bl6Qm5Xh1EpYaAdXV3YMnlERgB.docx	docx	khyati	khyati.selani03@gmail.com	.docx	YES	2019-03-31	Recovered
3	9	ka09ti8L7i5JR60pH6oBu8KQiVxXL5ZZ.docx	docx	khyati	khyati.selani03@gmail.com	.docx	YES	2019-03-31	Recovered



Message from IRP Manager

upanishadrocks@gma... Mar 24
to me

Dear User Your recovery request of file name Q8MOhR4U5xfeJqcTQfNI3SorTHyVqaHr.docx and the file id is 7 by using this file key 0pDjcu70c8GvgSrf yiu can Recover your file from Honeycomb layer. No need to pay any thing to attackers. thanking you...!!!!!!!!!!!!!

XII. CONCLUSION

We have tried to implement **"RansomWall: A Layered Defense System against Cryptographic Ransomware Attacks using Machine Learning".** The result is an intuitive way to provide effective way to secure data against ransomwares in the form of RansomWall, using a multi layered approach where each layer tackles various problems created by a cryptographic ransomware attack. The proposed method can be extended to all sorts of ransomwares by training it to tackle the new adversaries that emerge in real time. The system can and will be ameliorated as it gets used over and over in multiple deployment machines as it will gather new data sets to broaden in knowledge base. The system is currently being run on a windows environment. It can be modified to run on other computer operating systems as well as extended to create a mobile version which can be deployed on smart-phones and other mobile gadgets. Moreover a backup layer can be implemented to track and backup the infected files in case of an attack and later restore the said files after dealing with the attack. RansomWall can serve as a model on which new age ransomware detection and protection software can be based on.

REFERENCE

[1]Saiyed Kashif Shaukat, Vinay J. Ribeiro,"RansomWall: A Layered Defense System against Cryptographic Ransomware Attacks using Machine Learning," in **10th International Conference on Communication Systems & Networks,(COMSNETS),IEEE, 2018, pp. 356–363.**

[2]Avast Antivirus, "Locky Ransomware,"[Online].Available:https://www.avast.com/c-locky

[3]MalwareBytes, "Look into Locky," [Online].Available:https://blog.malwarebyte.com/threat-analysis/2016/03/look-into- locky/

[4]Comodo AV, "What is Cerber ransomware and how to remove it?," [Online]. Available: https://antivirus.comodo.com/blog/how -to/cerber-ransomware/

[5]Knowbe4, "Cryptowall," [Online]. Available:https://www.knowbe4.com/cryptowall

[6]Symantec, "Rasom.cryptowall," [Online]. Available:https://www.symantec.com/securitycenter/ writeup/2014-061923-282499

[7]Kara, I., & Aydos, M., "Static and Dynamic Analysis of Third Generation Cerber Ransomware", 2018, **International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT)**

[8]Daku H., Zavarsky P. & Malik, Y., "Behavioral-Based Classification and Identification of Ransomware Variants Using Machine Learning", 2018, 12th IEEE International Conference in Big Data Science and Engineering

[9]Prasant Kumar, Bimal Kumar Mishra, "e-Sier Epidemic Model for Spread of Malicious Objects, in Computer Network", 2013, in **Mathematical Sciences International Research Journal Volume 2 Issue 2**

[10]Gonzalez D., & Hayajneh, T., "Detection and prevention of crypto-ransomware", 2017 **IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)**