Reducing the number of Ransomware attacks on networks using machine learning pattern analysis

Mishquat Qureshi[1],Nitinkumar chaudhary[2],Dr.Jayant Karanjeker[3]

[1]Student of Computer Science Engineering Department, Wainganga College of Engineering & Management, Nagpur, India.

[2]Asst.Prof. of Computer Science Engineering Department, Wainganga College of Engineering & Management, Nagpur, India

[3] Prof. of Computer Science Engineering Department, Wainganga College of Engineering & Management, Nagpur, India

ABSTRACT

Later overall cybersecurity assaults brought about by Cryptographic Ransomware contaminated frameworks crosswise over nations and associations with a large number of dollars lost in paying blackmail sums. This type of malevolent programming takes client documents prisoner by encoding them and requests a huge payment installment for giving the unscrambling key. Mark based strategies utilized by Antivirus Software are deficient to dodge Ransomware assaults because of code muddling methods and making of new polymorphic variations regular. Conventional Malware Attack vectors are additionally not strong enough for discovery as they don't totally follow the particular personal conduct standards appeared by Cryptographic Ransomware families. This work dependent on examination of a broad dataset of Ransomware families presents RansomWall, a layered safeguard framework for insurance against Cryptographic Ransomware. It pursues a Hybrid methodology of consolidated Static and Dynamic examination to create a novel reduced arrangement of highlights that portrays the Ransomware conduct. Nearness of a Strong Trap Layer helps in early discovery. It uses Machine Learning for uncovering zero-day interruptions. At the point when introductory layers of RansomWall label a procedure for suspicious Ransomware conduct, documents changed by the procedure are upheld in the mood for protecting client information until it is delegated Ransomware or Benign. We will execute RansomWall for Microsoft Windows working framework (the most assaulted OS by Cryptographic Ransomware) and assessed it against numerous examples from various Cryptographic Ransomware families in genuine client situations. The testing of RansomWall with different Machine Learning calculations will give great outcomes with Gradient Tree Boosting Algorithm.

Keywords: *Ransomware, machine learning, gradient tree boosting, cryptography*.

I. INTRODUCTION

In the present carefully associated world, associations over the globe are seeing a huge development in cybercrime. The expanded reliance on computerized innovations has helped economies change the universe of business yet additionally lead to acceleration in the quantity of cyberattacks. Singular clients and corporates keep their significant records, photographs, reports and authoritative information in advanced structure. As of late, massivescale assaults were done utilizing a sort of malware known as Ransomware [1] that denies access to client information documents and requests a payoff for reestablishing it. In an exceptionally brief timeframe, Ransomware has developed exponentially to turn into the most perilous and forceful malware of late occasions. The assaults have been done on different segments [2] including fund, protection, banking, land, restorative, open organization to give some examples. Scareware [3] is an early type of Ransomware which use false dread in the injured individual that his framework is contaminated with an enormous number of infections, spyware and security issues. The client is deceived to purchase a phony antivirus item and thus pay a payoff for expelling contaminations. Client mindfulness and improved security programming have definitely decreased

danger presented by this sort of malware. Storage Ransomware (for example Reveton [4]) denies access to figuring assets by locking framework's UI. It utilizes social building strategies for compromising the client to pay recover. Successful apparatuses and procedures are given by different security sellers which can reestablish the blocked UI for generally variations. Cryptographic Ransomware [5] targets client information documents with explicit expansions that shifts with every family. Access to client information is hindered by scrambling records with cutting edge encryption calculations. A Ransom-note is shown to the client containing compromising message to erase prisoner records for all time if there should arise an occurrence of non-installment. Payoff is mentioned through Bitcoin digital currency. Framework records are not encoded to keep the working framework working. Indeed, even after installment it isn't ensured that the client gets the decoding key to reestablish scrambled documents. Cutting edge Cryptographic Ransomware variations utilize a blend of Symmetric (AES, Triple DES) and Asymmetric (RSA, ECC) Key Cryptographic calculations for encryption. Client records are scrambled utilizing Symmetric Key produced in the injured individual's framework. The Symmetric Key is encoded utilizing Asymmetric Public Key given by the assailant though comparing Asymmetric Private Key is stayed discreet at Command and Control server [6]. Cyberattackers make another Bitcoin wallet for every contamination and send its identifier to the unfortunate casualty for payment installment. Namelessness is given by going Bitcoins through numerous blenders which mixes them among various clients. Tor systems are utilized for concealed correspondence with Command and Control server. Cryptographic Ransomware is the real variation of Ransomware families that has caused destruction worldwide when contrasted with the other two variations - Scareware and Locker Ransomware. Microsoft Windows working framework has turned into the most assaulted OS by Cryptographic Ransomware as of late with immense cyberattacks principally focusing on its vulnerabilities for going into the injured individual's framework [7]. Far reaching utilization of this working framework in different stages crosswise over globe is the fundamental explanation behind developing as the practical objective of these assaults. Because of gigantic coercion sums included, new Cryptographic Ransomware variations are made regular.

II. RELATED WORK

The related works can be extensively grouped into two classifications: an) Approaches that treat Ransomware as a subset of the general malware network and apply conventional malware pointers for their recognition. b) Methods planned explicitly for Ransomware dependent on their trademark properties. Antivirus Signature Based Detection Techniques [9] are viable against known dangers whose marks are as of now present in their databases yet feeble against polymorphic and zero-day assaults. Gathering Policies and Application Whitelisting [10] are generally utilized inside confined corporate systems however are not reasonable for people and open associations. There is a reliance on right upkeep of whitelisted applications list. In addition, it is as yet feasible for malware to misuse vulnerabilities in whitelisted programming. Static Analysis Detection Techniques dependent on Control Flow Graphs [11], Data Flow Graphs [12] and System API Calls [13] are inclined to code muddling, polymorphic and changeable procedures. Ongoing Virtual Environment Analysis dependent on following data stream [14] has the restriction that numerous Ransomware variations hang tight for quite a while dependent on clocks, tally of framework restarts and so forth before beginning malevolent action. Subsequently, it isn't constantly doable to seclude the example and run it in a virtual situation during continuous execution. NetworkBased Intrusion Detection Systems [15] discover peculiarities in system designs. Ransomware trade predetermined number of encoded messages with Command and Control server which are hard to separate from typical traffic. Kharraz et al. [16] performed development based investigation of Locker and Cryptographic Ransomware. They recommended procedures dependent on observing of Master File Table and filesystem exercises however these strategies were not assessed by the creators. Andronio et al. [17] built up a method for distinguishing Android portable Ransomware utilizing examination of undermining payoff messages. Recognition of payoff note on Windows stage once it is shown to the unfortunate casualty isn't valuable continuously examination as client information is now scrambled at this stage. Kharraz et al. [18] actualized procedures joining auxiliary likeness of screen captures when test execution for distinguishing Locker Ransomware and observing of document framework exercises for Cryptographic variations. These strategies don't follow vindictive activities, executed by Ransomware to break framework barriers, for early recognition nor give any reinforcement component to safeguarding client information documents during investigation process. Mercaldo et al. [19] used formal strategies for recognizing Android Ransomware utilizing Bytecode portrayals.

III. PROBLEM DEFINATION

Each RansomWall layer depends on a particular usefulness. The layers are sorted out in calculation request of the highlights that are produced during the example's execution. It is actualized for Microsoft Windows working framework. Commitments of this work are as per the following: • Identify a novel minimized arrangement of highlights that describes the Cryptographic Ransomware conduct: Based on examination of a broad dataset of Ransomware families, a novel conservative list of capabilities is recognized which catch designs basic crosswise over various Cryptographic Ransomware variations. The Layered Architecture of RansomWall pursues a Hybrid methodology of joined Static and Dynamic examination to figure estimations of the chose list of capabilities.

• Create a Strong Trap Layer that aides in early discovery: This layer tracks pernicious exercises performed by Cryptographic Ransomware to break guards of the injured individual's framework and screens record tasks performed widely for encoding client information documents. These exercises are basic parts of a Cryptographic Ransomware assault.

• Use Machine Learning for uncovering zero-day interruptions: Cryptographic Ransomware broadly utilize polymorphic, transformative and obscurity methods to dodge signature-based discovery instruments utilized by Antivirus Software. Best interruptions are zero-day assaults. AI is utilized to build up a summed up model for the conservative list of capabilities that can identify zero-day tests. Execution of following directed learning calculations is assessed: Logistic Regression, Support Vector Machines (Gaussian-Kernel), Artificial Neural Networks, Random Forests and Gradient Tree Boosting.

• Develop a Backup component for safeguarding client information during identification process: The calculation and examination of minimized list of capabilities esteems by RansomWall layers and order choice by Machine Learning Engine takes some time, while Ransomware is as of now encoding the client information documents. Accordingly, there is a need to label a procedure as suspicious dependent on introductory highlights of Static, Dynamic and Trap layers. The records altered by the suspicious procedure are upheld up in a different envelope to protect client information until the procedure is named Ransomware or Benign by Machine Learning Layer.

• Evaluate RansomWall against 574 examples from 12 Cryptographic Ransomware families and 442 examples of Benign Software in genuine client situations: The Performance Metrices of Machine Learning Layer show best outcomes with Gradient Tree Boosting Algorithm. With this learning model

• Compare RansomWall's ability to distinguish zero-day interruption tests with 60 Security Engines connected to VirusTotal: 30 zero-day interruption tests having under 10% location rate by 60 Security Engines connected to VirusTotal [8] are gathered

IV. PROCESS DIAGRAM

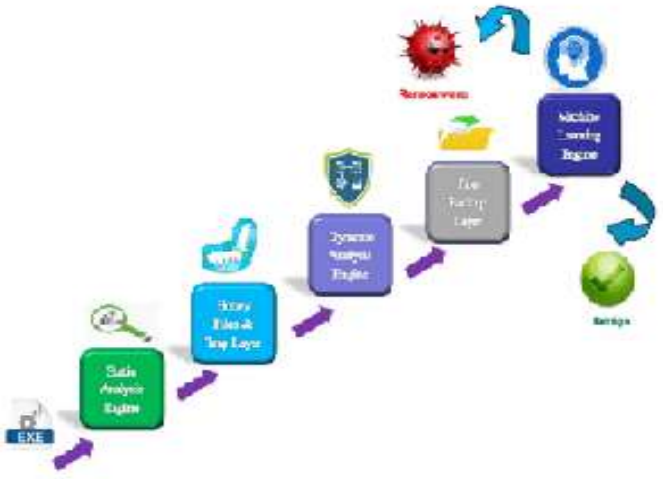

Figure 1. Flow diagram of the proposed RansomWall system

The different layers present in RansomWall Architecture are portrayed underneath.

1) Static Analysis Engine: It gives helpful data from double code of the example. This is the principal layer of RansomWall Architecture as highlights required for investigation can be gotten before executing the example. Static highlights considered during the trial investigation for finding their viability against Ransomware families are: PE (Portable Executable) header subtleties, implanted assets, location of packers/cryptors, test's Entropy, PE Digital Signature, installed strings and fluffy hashes.

2) Honey Files and Trap Layer: The personal conduct standards normal for Cryptographic Ransomware include performing noxious exercises to break resistances of the unfortunate casualty's framework and scrambling client information records. This layer sets trap by following event of these pernicious exercises. Cryptographic Ransomware perform encryption of client information records with explicit expansions. Nectar Files (with client information record expansions generally assaulted by Ransomware) and Honey Directories are conveyed in basic client information organizers. These are trap records/registries which are not expected to be adjusted by the client during typical activity. Alteration of these records/indexes by a procedure gives a sign of suspicious conduct.

3) Dynamic Analysis Engine: Static highlights alone are not adequate because of code confusion, pressing and encryption strategies utilized by Ransomware. Dynamic examination screens conduct of the example during real execution. Cryptographic Ransomware performs broad encryption of client information records. This layer screens document framework tasks and entropy adjustments for following huge encryption exercises.

4) File Backup Layer: The calculation and investigation of highlights gathered during the example's execution and arrangement choice by Machine Learning layer takes some time, while Ransomware is as of now scrambling client information records. Along these lines, there is a need to label a procedure as suspicious dependent on starting highlights of Static, Dynamic and Trap layers. Records adjusted by the suspicious procedure are supported up in a different envelope to safeguard client information until the procedure is named Ransomware or Benign by Machine Learning layer. RansomWall keeps up rundown of documents that are supported up alongside their unique areas and Process ID of the suspicious procedure. On the off chance that Machine Learning layer groups as Ransomware, at that point the procedure is murdered and records changed by it are reestablished to their unique areas. On the off chance that it is named Benign, at that point these documents are erased from the reinforcement organizer.

5) Machine Learning Engine: This layer assembles a summed up model which is powerful against zero-day Ransomware assaults. It takes highlight esteems gathered by Static, Dynamic and Trap layers as info and characterizes the executable as Ransomware or Benign. The Machine Learning Engine is prepared disconnected utilizing Supervised calculations. Preparing information comprises of highlight esteems with Ransomware and Benign names. Prepared Machine Learning Engine utilize the educated model to order executables progressively dependent on info highlight esteems. Execution of following Supervised Machine Learning calculations is assessed: Logistic Regression, Support Vector Machines (Gaussian-Kernel), Artificial Neural Networks, Random Forests and Gradient Tree Boosting.

V. CONCLUSION AND FUTURE WORK

Recent worldwide cybersecurity attacks caused by Cryptographic Ransomware massively crippled organizations across the globe. Based on the analysis of an extensive Ransomware dataset, this work presents a layered defense mechanism with monitoring of a novel compact feature set that characterizes Ransomware behavior. Strong Trap layer (early detection), Machine Learning layer (zero-day intrusions) and File Backup layer (preserving user data) helps RansomWall to attain a high detection rate with near-zero false positives using Gradient Tree Boosting Algorithm. We will be evaluating RansomWall on large-scale real setups as a future work.

FUTURE SCOPE:

1. To enhance the security more, a mechanism to secure the keys in security cloud can be a area of research.
2. To reduce the overhead of network traffic can be another area of research.

3. To provide a framework to supply a secure cloud database that will guarantee to prevent security risks facing the cloud computing community.

VI. REFERENCES

[1] Barkly, "WannaCry Ransomware Statistics: The Numbers Behind the Outbreak," May 2017. [Online]. Available: https://blog.barkly.com/ wannacry-ransomware-statistics-2017
[2] CNN Tech, "Ransomware attack: Who's been hit," May 2017. [Online]. Available: http://money.cnn.com/2017/05/15/technology/ ransomware-whos-been-hit/index.html
[3] TechTarget, "Scareware," Aug 2010. [Online]. Available: http://whatis. techtarget.com/definition/scareware
[4] F-Secure, "Trojan: W32/Reveton: Threat description," 2017. [Online]. Available: https://www.f-secure.com/v-descs/trojan_w32_reveton.shtml
[5] Sophos, "The current state of ransomware: CTB-Locker," 2015. [Online]. Available: https://news.sophos.com/en-us/2015/12/31/ the-current-state-of-ransomware-ctb-locker
[6] Panda Security, "CryptoLocker: What Is and How to Avoid it," 2015. [Online]. Available: http://www.pandasecurity.com/mediacenter/ malware/cryptolocker
[7] SecureList, "WannaCry ransomware used in widespread attacks all over the world," May 2017. [Online]. Available: https://securelist.com/ wannacry-ransomware-used-in-widespread-attacks-all-over-the-world/ 78351
[8] VirusTotal, "Free Online Virus, Malware and URL Scanner," 2017. [Online]. Available: https://www.virustotal.com
[9] Comodo, "How Antivirus Works," 2017. [Online]. Available: https: //antivirus.comodo.com/how-antivirus-software-works.php
[10] SentinelOne, "The Truth About Whitelisting," Dec 2014. [Online]. Available: https://sentinelone.com/2014/12/07/ the-truth-about-whitelisting
[11] P. Faruki, V. Laxmi, M. S. Gaur, and P. Vinod, "Mining control flow graph as api call-grams to detect portable executable malware," in Proceedings of the Fifth International Conference on Security of Information and Networks. ACM, 2012, pp. 130–137.
[12] T. Wüchner, M. Ochoa, and A. Pretschner, "Robust and effective malware detection through quantitative data flow graph metrics," in International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment. Springer, 2015, pp. 98–118.
[13] Y. Ye, D. Wang, T. Li, D. Ye, and Q. Jiang, "An intelligent pe-malware detection system based on association mining," Journal in computer virology, vol. 4, no. 4, pp. 323–334, 2008.
[14] H. Yin, D. Song, M. Egele, C. Kruegel, and E. Kirda, "Panorama: capturing system-wide information flow for malware detection and analysis," in Proceedings of the 14th ACM conference on Computer and communications security. ACM, 2007, pp. 116–127.
[15] N. Das and T. Sarkar, "Survey on host and network based intrusion detection system," International Journal of Advanced Networking and Applications, vol. 6, no. 2, p. 2266, 2014.