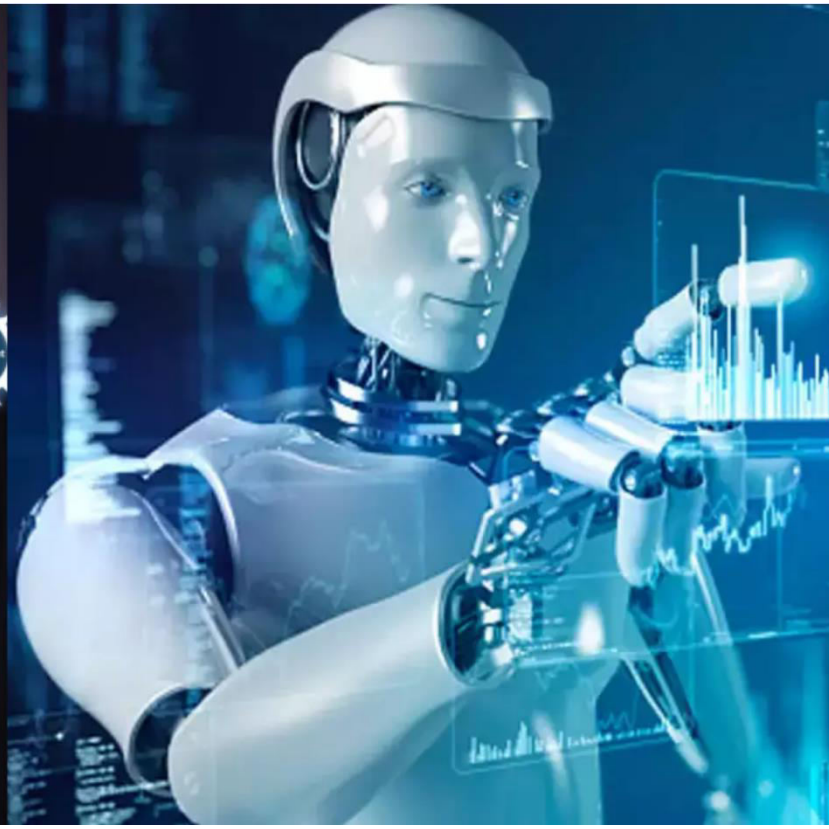
IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

ISSN(O): 2319-8753
ISSN(P): 2347-6710

International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 2, February 2025

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 2, February 2025

|DOI: 10.15680/IJIRSET.2025.1402074|

A Survey on Behavioural-Based Classification and Identification of Ransomware Variants using Machine Learning

Jitti Annie Abraham[1], Femeena Fousiya S.R.[2]

Assistant Professor on Contract, Department of Mathematics and Computer Science, St. Cyril's College, Adoor, Kerala, India[1]

Assistant Professor on Contract, Department of Mathematics and Computer Science, St. Cyril's College, Adoor, Kerala, India[2]

ABSTRACT: Due to the changing behaviour of ransomware, traditional classification and detection techniques do not accurately detect new variants of ransomware. Attackers use polymorphic and metamorphic techniques to avoid detection of signature-based systems. Hereuse machine learning classification to identify modified variants of ransomware based on their behaviour. This paper provides a survey on behavioural based classification and identification of ransomware variants using machine learning. The main goal and contribution of this review paper is to present the overview of machine learning, ransomware attacks, and provides machine-learning techniques ad also ransomware prevention and detection techniques Two main parts of this study are identification of the behavioural attributes which can be used for optimal classification accuracy and classification of ransomware using machine learning algorithms. Herehave evaluated classification accuracy of machine learning classification algorithms.

KEYWORDS: behavioural analysis, malware classification, machine learning, ransomware, malware detection, ransomware prevention.

I. INTRODUCTION

Ransomware attacks are becoming a serious cyber threat to organizations and individuals around the global. Ransomware is a type of malicious software from cryptovirology that threatens to publish the victim's data or perpetually block access to it unless a ransom is paid. While some simple ransomware may lock the system in a way which is not difficult for a knowledgeable person to reverse, more advanced malware uses a technique called cryptoviral extortion, in which it encrypts the victim's files, making them inaccessible, and demands a ransom payment to decrypt them. In a properly implemented cryptoviral extortion attack, recovering the files without the decryption key is an intractable problem and difficult to trace digital currencies such as Ukash and cryptocurrency are used for the ransoms, making tracing and prosecuting the perpetrators difficult.

An ongoing report, reports a monstrous increment and relentless development in new ransomware tests and attacks. Also, the capacity of the attackers to make new variations of existing malware utilizing metamorphic, polymorphic, obscurity and other concealing methods makes it harder to adequately identify and characterize these attacks utilizing static location frameworks, since they may have distinctive substance or mark from their past renditions. Regularly, new variations of malware are not recognized from their ancestors because of the impediments of order frameworks depending just on static investigation. Accordingly, methods like static based, signature-based and design coordinating methods for malware investigation are ending up less viable to identify and order new variations of ransomware and give knowledge data about the risk, objectives and practices of ransomware.

Efforts have been made to develop behaviour-based classification techniques. Classification of malware samples based on their behaviour requires implementation of algorithms that are capable to produce models and learn through the classification process. The ability of machine learning to learn with data during the process of classification, makes them attractive and effective for malware classification. Using machine learning classification algorithms, ransomware samples can be identified with different behaviours from other samples that are part of the same family. The reason

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 2, February 2025

|DOI: 10.15680/IJIRSET.2025.1402074|

behind this study is to identify new modified variants of ransomware based on their behaviour using machine learning algorithms. Two main parts of this study are identification of the behavioural attributes which can be used for optimal classification accuracy and classification of ransomware. [1]

Paper is organized as follows. Section II describes the survey that done on the paper "Behavioural Based Classification and Identification of Ransomware Variants Using Machine Learning". This section goes through more than ten various papers giving an overview on machine learning techniques and on ransomware attack analysis, detection and prevention. Also provides the importance of classification of ransomwares based on their behaviour. Section III presents conclusion.

II. LITERATURE SURVEY

Artificial Intelligence is everywhere. Possibility is that we are using it in one way or the other and you don't even know about it. One of the popular applications of AI is Machine Learning, in which computers, software, and devices perform via cognition which is very similar to human brain. Machine learning a field of artificial intelligence that uses statistical techniques to give computer systems the ability to "learn" from data, without being explicitly programmed. Some of the trending applications of machine learning includes: virtual personal assistants, predictions while commuting, social media services, email spam and malware filtering, search engine result refining and product recommendations. Machine learning is probably going to be a standout amongst the most transformative innovations of the 21st century. [2]

Machine learning is multi displinary field in artificial intelligence, likelihood insights data hypothesis, reasoning, human science, and neurobiology. Machine learning tackles this present reality issues by building a model that is great and valuable estimation to the information. The study on machine learning taking in has developed from the endeavours of investigating regardless of whether computer could figure out how to imitate the human mind, furthermore, a field of measurements to a wide control that has created central factual computational speculations of learning forms. The fundamental objective and commitment of this audit paper is to display the diagram of machine learning and gives machine-learning procedures. Additionally, paper surveys the benefits and limitations of different machine learning algorithm in diverse methodologies. [3]

Ransomware attacks are quickly developing in fame with cybercriminals what's more, all things considered – it's assessed that this sort of attack gains lawbreakers a large number of pounds multi month. Over the past three years, ransomware has become one of the biggest cyberscams to hit businesses. Ransomware is a malicious software that enables a hacker to confine access to a person's or organizations indispensable data here and there and afterward request some type of payment to lift the limitation. The five phases of ransomware can be pointed out as exploitation & infection, delivery & execution, back-up spoliation, file encryption and user notification & clean-up. The steps for handling a ransomware attack including preparation, detection, containment, eradication and recovery. [4]

Ransomware is a class of malware whose goal is to extort money. At the point when introduced on a framework, a ransomware encodes documents or squares functionalities and when the activity is done it requests a recover. In this paper we survey current barrier strategies for ransomware, talking about their solid and powerless focuses. Here we describe existing techniques to mitigate ransomware and discuss their limitations. The existing ransomware mitigation systems are built upon the analysis of collected samples that is they excluding the inefficient and ineffective practice to back-up and restore files. Future Threats of Ransomware includes rootkit-based ransomware, obfuscation, white-box cryptography, socio-technical attacks [5]

Ransomware has turned into a lucrative business that has increased expanding prominence among assailants. Not at all like customary malware, even after evacuation, is ransomware's impact irreversible and troublesome to relieve without the assistance of its maker. The study advances a novel ransomware scientific classification, from a few points of view. It at that point explains on the components that prompt a fruitful ransomware assaults before examining in detail the investigation into checking ransomware, counting examination, counteractive action, and discovery and forecast arrangements. The study finishes up with a brief discussion on the open issues and potential research headings sooner rather than later. The proposed taxonomy classifies ransomware from three perspectives: severity based, platform based, and target based. [6]

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 2, February 2025

|DOI: 10.15680/IJIRSET.2025.1402074|

Ransomware is the sort of malware that averts or limits client from getting to their framework, either by locking the framework's screen or by securing the clients' documents in the framework except if a payment is paid. More present day ransomware families, exclusively sort as crypto-ransomware, scramble certain record types on tainted frameworks and powers clients to pay the payoff through online payment techniques to get an unscramble key. The analysis demonstrates that there has been a noteworthy enhancement in encryption methods utilized by ransomware. The careful analysis of ransomware conduct can create a viable identification framework that fundamentally lessens the measure of injured individual information misfortune. Some security applications distinguish ransomware dependent on its action, for example, File System Activities, Registry Activities, Gadget control Communications, Network Activity, and Locking component. [7]

Ransomware is a type of vindictive code or malware that taints a computer and spreads quickly to scramble the information or to bolt the machine. The aim of the paper is on giving bits of knowledge on how ransomware have developed from its beginning till March 2016 by breaking down examples of chose ransomware variations from existing ransomware families in Windows and Android conditions. This paper focused around giving the understanding perspective on how ransomware functions and the manner in which it has developed amid this day and age by watching a few examples identified with these chose ransomware families. Also clarifies how the ransomware tests are working and how the discoveries can be utilized to recognize ransomware. [8]

The malwares being planned by attackers are polymorphic and metamorphic which can change their code as they spread. Also, the decent variety and volume of their variations extremely undermine the viability of customary barriers which normally utilize signature based procedures and can't distinguish the already obscure malignant executable. The variations of malware families share common standards of conduct mirroring their cause what's more, reason. The personal conduct standards acquired either statically or powerfully can be misused to recognize and characterize obscure malwares into their realized families utilizing machine learning procedures. This review paper gives a diagram of strategies to breaking down and ordering the malwares. [9]

Malware programs are fit for concealing themselves in frameworks or impairing their movement when they find endeavours to distinguish them. In this manner, it is required to utilize aloof frameworks (i.e., trusted monitoring) that can identify vindictive exercises on focused machines without getting to them. Here presents an end-to-end supervised based system for detecting malware by analysing network traffic. The proposed method extracts 972 behavioural features across different protocols and network layers, and refers to different observation resolutions (transaction, session, flow and conversation windows). A feature selection method is then used to identify the most meaningful features and to reduce the data dimensionality to a tractable size. Finally, various supervised methods are evaluated to indicate whether traffic in the network is malicious, to attribute it to known malware “families” and to discover new threats. [10]

Malware is one of the major security threats that can break computer operation. However, commercial antivirus or anti-spyware that used signature-based matching to detects malware cannot solve that kind of threats. Nowadays malware writers try to avoid detection by using several techniques such as polymorphic, metamorphic and also hiding technique. In order to overcome that issue, here proposed a new framework for malware behaviour identification and classification that apply dynamic approach. This framework consists of two major processes such as behaviour identification and malware classification. These two major processes will integrate together as interrelated process in our proposed framework. The result from this study is a new framework that able to identify and classify malware based on it behaviours. [11]

Ransomware is a type of malicious software (malware) that once executed on a computer system, hinders the user from using the computer or its data, demanding a sum of money (ransom) for the restoration of the computer. Currently, ransomware attacks hinder computer operation in three ways: by blocking accessing to the computer, this form of ransomware is referred to as locker ransomware; by making user data unusable by means of employing encryption algorithms, referred to as crypto ransomware; and a combination of locker/crypto ransomware where a user is blocked from using their computer while their data is being encrypted. This paper provides motivation for the use of machine-learned behaviour for ransomware detection, and in doing so demonstrates technical underpinnings of RansomFlare. [12]

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 2, February 2025

|DOI: 10.15680/IJIRSET.2025.1402074|

In the face threat of the Internet attack, malware classification is one of the promising solutions in the field of intrusion detection and digital forensics. In previous work, researchers performed dynamic analysis or static analysis after reverse engineering. But malware developers even use anti-virtual machine and obfuscation techniques to evade malware classifiers. By means of the deployment of honeypots, malware source code could be collected and analysed. Source code analysis provides a better classification for understanding the purpose of attackers and forensics. In this paper, a novel classification approach is proposed, based on content similarity and directory structure similarity. Such a classification avoids to re-analyse known malware and allocates resources for new malware. Malware classification also let network administrators know the purpose of attackers. [13]

A Classification is a method of predicting similar information from the value of a categorical target or categorical class variable. It is a useful technique for any type of statistical data. These algorithms are used for various purposes like image classification, Predictive modelling, data mining technique etc. The main purpose of supervised learning is to build a simple and unambiguous model of the allocation of class labels in terms of predictor features. The classifiers are then used to classify class labels of the testing instances where the values of the predictor features are known, to the value of the class label which is unknown. In this paper we illustrate various classification techniques used in supervised machine learning. [14]

Ransomwares are notorious threats that have become rampant in the cyber world, the peculiarity about them is that they encrypt user's critical data using strong encryption algorithms. The encrypted data can only be recovered by paying a ransom in bitcoins to the attacker. A traditional signature based detection approach is slow and time consuming and cannot handle the zero-day attacks. Therefore, in this paper, an attempt is made to build a proactive Machine Learning based classifier which is trained to detect the ransomwares based on the static attributes of the PE files. [15]

Ransomware is type of malware and Trojans are types of cyberware that are designed to extort money from a victim. Ransomware can attack on different platform like windows, android etc. It can attack systems in two ways, either can encrypt data or can block information. Ransom attacks can be prevented by paying closer attention to application permission request and by using prevention techniques. Prevention techniques can help detect and remove ransomware without obtaining information about ransomware. The focus of the paper is on ransomware attacks on windows, android and other environments. In windows ransomware attackers can be prevented by monitoring abnormal file system and in android it can be detected by paying close attention to the android manifest file. [16]

Ransomware is now become a bad tool to earn money, theft data, hack the system or to stop the normal functioning of the system. Ransomware is a malware that breaches the security of the system by using malicious codes. It encrypts the information and available data before noticing it. Traditional vaccination system does not cure the infected system without obtaining information on ransomware. Since the data is encrypted hence cannot be recovered without encryption key. Users can avoid the infections of ransomware by updating vaccination system from time to time. However, this method has limited efficacy. This approach cannot trace modified ransomware with new pattern. This paper explores the various ransomware attack. In this paper we converse the analysis of ransomware and the suggested action against ransomware attack. This paper also discusses ransomware removal and preventional methodology [17]

Crypto-ransomware is a challenging threat that ciphers a user's files while hiding the decryption key until a ransom is paid by the victim. This type of malware is a lucrative business for cybercriminals, generating millions of dollars annually. The spread of ransomware is increasing as traditional detection-based protection, such as antivirus and anti-malware, has proven ineffective at preventing attacks. Additionally, this form of malware is incorporating advanced encryption algorithms and expanding the number of file types it targets. This paper discusses ransomware methods of infection, technology behind it and what can be done to help prevent becoming the next victim. The paper investigates the most common types of crypto-ransomware, various payload methods of infection, typical behaviour of crypto ransomware, its tactics, how an attack is ordinarily carried out, what files are most commonly targeted on a victim's computer, and recommendations for prevention and safeguards are listed as well [18]

Ransomware attacks due to code obfuscation techniques and creation of new polymorphic variants every day. Generic Malware Attack vectors are also not robust enough for detection as they do not completely track the specific behavioural patterns shown by Cryptographic Ransomware families. This work based on analysis of an extensive dataset of Ransomware families presents RansomWall, a layered defense system for protection against Cryptographic Ransomware. It follows a Hybrid approach of combined Static and Dynamic analysis to generate a novel compact set

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 2, February 2025

|DOI: 10.15680/IJIRSET.2025.1402074|

of features that characterizes the Ransomware behaviour. Presence of a Strong Trap Layer helps in early detection. It uses Machine Learning for unearthing zero-day intrusions. When initial layers of RansomWall tag a process for suspicious Ransomware behaviour, files modified by the process are backed up for preserving user data until it is classified as Ransomware or Benign. [19]

Ransomware techniques have evolved over time with the most resilient attacks making data recovery practically impossible. This has driven countermeasures to shift towards recovery against prevention but in this paper, they model ransomware attacks from an infection vector point of view. Then follow the basic infection chain of crypto ransomware and use Bayesian network statistics to infer some of the most common ransomware infection vectors. Also employ the use of attack and sensor nodes to capture uncertainty in the Bayesian network. [20]

III. CONCLUSION

This paper provides a survey on behavioural based classification and identification of ransomware variants using machine learning. Machine learning model is given which describes the overview of machine learning process. Also describes the various machine learning algorithm based on types of machine learning styles. The paper concluded about various ransomware attacks analysis, detection and prevention techniques. Also discussed about the behavioural approach for the classification and identification of ransomware variants. Using this it's able to apply the defence mechanisms to ransomware attacks.

REFERENCES

[1] Hajredin Daku, Pavol Zavarsky, Yasir Malik, "Behavioral-Based Classification and Identification of Ransomware Variants Using Machine Learning", 2324-9013/18/31.00 © IEEE, 2018
[2] M. I. Jordan and T. M. Mitchell,"Machine Learning: Trends, Perspectives, And Prospects", Science 349,255 2015.
[3] Sandhya Ndhage, Charanjeet Kaur Raina, "A Review On Machine Learning Techniques", IJRITCC, ISSN: 2321-8169 Volume: 4 Issue: 3 395 – 399, 2016.
[4] Ross Brewer, Log Rhythm, "Ransomware Attacks: Detection, Prevention and Cure", ELSEVIER, September 2016.
[5] Ziya Alper Gen, Gabriele Lenzini, Peter Y.A. Ryan ,"The Cipher, the Random and the Ransom: A Survey on Current and Future Ransomware", CECC, November 2017.
[6] BanderAli SalehAl-rimy, Mohd AizainiMaarof, Syed Zainudeen MohdShaid, "Ransomware Threat Success Factors, Taxonomy, And Countermeasures: A Survey and Research Directions", ELSEVIER 2018.
[7] Jinal P. Tailor Ashish D. Patel, "A Comprehensive Survey: Ransomware Attacks Prevention, Monitoring and Damage Control", IJRSI VolumeIV, June 2017.
[8] Monika, Pavol Zavarsky, Dale Lindskog,"Experimental Analysis of Ransomware on Windows and Android Platforms: Evolution and Characterization", ELSEVIER 2016.
[9] E. Gandotra, D. Bansal, S. Sofat, "Malware Analysis And Classification: A Survey", Journal of Information Security, vol. 5, no. 02, pp. 56, 2014.
[10] D. Bekerman, B. Shapira, L. Rokach, A. Bar and B. Sheva, "Unknown Malware Detection Using Network Traffic Classification," pp. 134–142, IEEE 2015.
[11] Mohamad Fadli Zolkipli, Aman Jantan, "An Approach for Malware Behavior Identification and Classification", IEEE 2011.
[12] D Nieuwenhuizen, "A Behavioural-based Approach to Ransomware Detection" Information Security 2017.
[13] Chen Chia-Mei, Lai Gu-hsin, "Research on Classification of Malware Source Code", Springer, 2014.
[14] R. Vijaya Kumar Reddy, Dr. U. Ravi Babu, "A Review on Classification Techniques in Machine Learning",ICRTESM March 2018.
[15] Niranjan Agnihotri, "Ransomware Classifier using Extreme Gradient Boosting", IJCSIT 2018.
[16] Sonu B. Surati, Ghanshyam I. Prajapati, "A Review on Ransomware Detection & Prevention", ISSN 2321–2705, IJRSI, Volume IV, Issue IX, September 2017.
[17] Smruti Saxena, Hemant Kumar Soni, "Strategies for Ransomware Removal and Prevention", 978-1-5386-4606-9© IEEE, 2018.
[18] Daniel Gonzalez, Thaier Hayajneh, "Detection and PreventionofCrypto-Ransomware", 978-1-5386-1104-3/17/$31.00 © IEEE, 2017.
[19] Saiyed Kashif Shaukat, Vinay J. Ribeiro, "RansomWall: A Layered Defense System against Cryptographic Ransomware Attacks using Machine Learning", IEEE, 2018.
[20] Aaron Zimba , Zhaoshun Wang , Hongsong Chen, " Reasoning Crypto Ransomware Infection Vectors with Bayesian Networks", 978-1-5090-6727-5/17/$31.00 © IEEE, 2017.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details