JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Proactive Anti Agent against Ransomware Threats Using RanGAN and Hash Conceal

[1]Yuvashri M, [2]Vijayakumar S
[1]PG scholar, [2]Associate Professor
[1]Department of Electronics and Communication Engineering
[1]Paavai Engineering College, Namakkal, India

Abstract: Ransomware attacks have become a pervasive and evolving threat in the digital landscape, demanding innovative defensive strategies to protect organizations and individuals. Ransomware creators continually refine their tactics, rendering traditional signature based detection methods ineffective. Additionally, ransomware is often delivered through a variety of vectors including phishing emails, malicious attachments and compromised websites. Detecting and thwarting these diverse delivery methods presents a significant challenge. This proposed method introduces a proactive defense strategy against ransomware threats, leveraging "RanGAN (*Ransomware Generative Adversarial Network*)" for early detection and "Hash Conceal" for data protection. RanGAN employs advanced machine learning to detect ransomware behavior patterns in real time, while Hash Conceal secures critical data from malicious encryption. So we can provide a robust defense, ensuring rapid threat identification and minimizing data loss. By using a real time ransomware sample, the experiment result shows that our proposed method protects the files from the attack in an effective manner.

Index Terms- Ransomware, Cybersecurity, Defense mechanism, Machine Learning

I. INTRODUCTION

Ransomware is a malware designed to hack files on the computer or other devices of a user or organization. By demanding a payment to user for rescue the file and cracking the files using decryption key, cyber bushwhackers place association in a position where paying the rescue is the easiest and cheapest way to recapture access to their files. Ransomware attacks are via malicious links or attachments in emails having 45% and 21% of attacks are from a remote attack on servers, according to the Sophos survey [4]. The remaining attacks are through misconfigured systems and USB devices. Attackers continually come out with novel techniques to get around systems for detection. Although the effectiveness of detection methods have three abecedarian boundaries, as understood by studies suggesting detecting mechanism or researching the action of ransomware:

1. The efficiency of static and dynamic analysis techniques decrease with the ongoing development of evasion techniques through malware developers [11].
2. File loss may occur from behavior based detection methods until detection is accomplished [12].
3. The detection mechanism may become ineffective, if the monitoring process is interrupted [13].

Previous ML algorithms and other methods for ransomware detection, classification and prevention encountered several challenges and problems. One of the foremost issues has been the imbalance in the dataset used for training and testing these models. Imbalance datasets made it challenging for algorithms to learn effective patterns and often resulted in biased models. This project is to develop and implement a proactive defensive strategy against ransomware threats using a combination of RanGAN and Hash Conceal. The primary goal is to create a robust and innovative approach to prevent, detect and mitigate ransomware attacks effectively, including unknown and evolving variants, while also enhancing data protection and privacy.

II. LITERATURE REVIEW

Ransomware detection based related works are mostly on either static or dynamic analysis or a few hybrid analyses, which consist both static and dynamic. Here we discuss some of those related works.

Smith et al. [7] proposed a comprehensive analysis of ransomware detection frameworks and common Machine Learning (ML) algorithms used to identify evolving ransomware characteristics. This paper explores various algorithms such as decision trees, random forests, Support Vector Machines (*SVM*), deep learning models and clustering algorithms. However limitations include the rapidly evolving nature of ransomware and the need for continuous adaptation of detection frameworks to evolving ransomware tactics.

Zhang et al. [8] introduced the TGAN-IDS framework, which uses dual generative adversarial networks to detect unknown encryption ransomware attacks. The framework uses a multistage approach, starting with a Deep Convolutional Generative Adversarial Network based on a pre-training model for binary classification. A reconstruction loss function is introduced to prevent a decrease in detection rate of normal samples. It experienced the challenge of maintaining detection rates for normal samples during adversarial training.

Poudyal et al. [11] explained an AI powered hybrid approach. The methodology involves multi-level profiling of crypto ransomware, capturing distinctive features at the Dynamic Link Library, function call and assembly levels. It provided behavioral chaining which creates unique behavioral chains for ransomware detection. The experimental results demonstrate the effectiveness of this method achieving the highest accuracy of 99.72% with ML algorithms.

Chai Ming Hsu et al. [14] proposed a novel approach that prioritizes file detection over executable program detection, to establish a backup system to protect user files. The study analyze 22 encrypted file formats, extract specific features and employs the SVM as a classifier to differentiate between encrypted and unencrypted files. Dynamic nature of ransomware poses a challenge to achieve detection rate.

III. RANSOMWARE ANALYSIS

Malware analysis is a crucial method for understanding the components and behavior of malware including ransomware. It helps detect and prevent future attacks by analyzing binary file contents and processes during execution. Static analysis analyzes binary fickle contents while dynamic analysis studies the behavior and action of a process during execution. Signature based malware detection is a static approach that uses unique patterns within the malicious file to detect it, such as unique sequences of bytes, function calls or ransomware notes. In behavior based detection, the system monitors the behavior of files and processes in real time. It is prone to generating false positives, which can lead to operational disruptions and alert fatigue. Anomaly detection algorithms are used to identify outliers in data that deviate from expected pattern.

The researchers conducted a comprehensive analysis of various ransomware families to identify their common features, with a focus on crypto ransomware executed in the Windows Operating System (OS). They identified key insights from the attacker's perspective:

1. Reliable Encryption Environment: Attackers aim to encrypt files reliably without destroying the system to demand a ransom. They avoid encrypting files that may affect the OS's functionality, allowing them to execute the entire encryption process and inform the victim of the damage [13].
2. Fast Impact: Attackers want to encrypt files quickly and efficiently selecting files based on their extensions and paths to reduce discovery time. They also select the appropriate encryption method based on the file size [16].
3. Evolving Attack Techniques: Attackers use various initial access and defense evasion techniques, constantly evolving tactics for successful attacks [17].

EXECUTION FLOW OF RANSOMWARE: The researchers identified four common phases of ransomware attack: initial access, defense evasion, targeting and lateral movement. In the targeting phase, ransomware excludes system related files from encryption for a reliable attack and classifies files based on extensions and paths for fast encryption.

IV. PROPOSED METHOD

In this section, we proposed a novel method to encounter ransomware targeting by applying a RanGAN and Hash Conceal technology. RanGAN is a Deep Learning (DL) model, with dynamic analysis and behavioral anomaly detection to provide a proactive defense against ransomware threats. Dynamic analysis techniques monitor file and process behavior in real time, assessing deviation from normal patterns as indicative of ransomware activities.

The RanFooler Web tool is a cloud based ransomware analysis tool that scans registered devices from a remote server, providing users with a report on system or device security threats. Bidirectional Long Short Term Memory (*BiLSTM*) model and Generative Pre-trained Transformer 2(*GPT2*) model detect ransomware code pieces by examining assembly instructions from static analysis results of Portable Executable file. In this project, we can use byte files and Asm files as the load dataset, which used for model training and processing. For feature extraction using a shallow deep learning based method (*word2vec*) to represent ransomware based on its opcodes. A stacking method is used to classifier initial process, starting with an embedding layer that creates an embedding vector for each word index. The binary cross entropy is a loss function that used for classifying samples in two orders. Document Level Analysis Model (*DLAM*) and Sentence Level Analysis Model (*SLAM*) are developing for the analysis. The system's MAC and IP addresses used for the end device configuration.

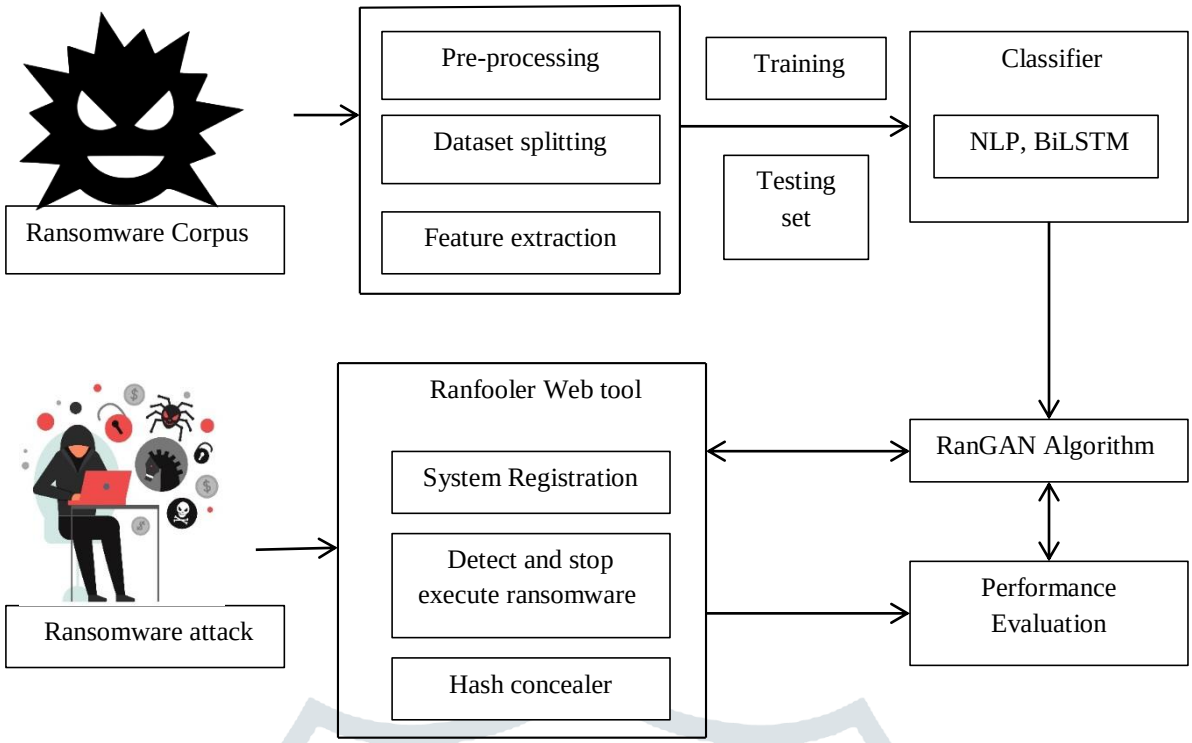

Figure1. Execution Flow of Proposed Method

Algorithm 1: RanGAN algorithm

```
def  create_generator():
def create_dicriminator():
def binary_cross_entropy_loss(predictions, targets):
def train_gan(generator, discriminator,data_loader, num_epochs, learning_rate):
     Generator_optimizer = Adam(generator.parameters(), lr=learning_rate)
discriminator_optimizer = Adam(discriminator.parameters(), lr =learning_rate)
  for epoch in range(num_epochs):
      for real_data in data loader:
          Real_labels = torch.ones(real_data.size(0),1)
    fake_lables = torch.zero(real_data.size(0),
      real_outputs = discriminator(real_data)
            real_loss = binary_cross_entropy_loss(real_outputs, real_labels)
            fake_data = generator(generate_noise(real_data.size(0)))
            fake_outputs = discriminator(fake_data.detach())
    fake_loss = binary_cross_entropy_loss(fake_outputs,fake_labels)
     discriminator_loss = real_loss + fake_loss
     discriminator_optimizer.zero_grad()
     discrcrimmminator_loss.backward()
     discriminator_optimizer.step()
        updated_fake_outputs = discriminator(fake_data)
            generator_loss = binary_cross_entropy_loss(updated_fake_outputs, real_label
generator_optimizer.zero_grad()
generator_loss.backward()
geenerator_optimizer.step()
def generate_noise(batch_size):
    return torch.randn(batch_size, latent_dim)
```

Hash Conceal technology is integrated to enhance data protection and privacy. Files and data are hashed, providing an extra layer of security to prevent unauthorized access and ensure data integrity. A mapping table and hash table facilitate file recovery by pairing original file with hidden file. RanFooler offers two types of online scans for ransomware code detection: Antivirus scan and Prevent scan.

Algorithm 2: Hash Conceal

```
hidden_files = { } #Dictionary to store hidden files and their paths
hash_table = { } # Hash table to store hash values and hidden file paths
function hide_file(original_path):
      hash_value = hash_function(original_path)
      hidden_path = generate_hidden_path(hash_value)
      hidden_files[original_path](hidden_path)
      Hash_table[ hasj_value] = hidden_path
      move_file(original_path, hidden_path)
function retrieve_file(original_path):
     if original_paa in hidden_files:
     hidden_path = hidden_files[original_path]
     move_file(hidden_path, original_path)
  del hidden_files[original_path]
      del hash_table[hash_function(original_path)]
function access_hidden_fillllllllle(link_file_path)
      hash_value = extract_hash_from_;link(link_file_path)
      hidden_path = hash_table[hash_value]
        associate_file_with_program(hidden_path)
function hash_function(file_path);
      return hash(file_path)
function generate_hidden_path(hash_value):
      return "/hidden/"+hash_value+ "/"
function move_file(source_path, destination_path):
function extract_hash_from_link(link_file_path);
```

V. PERFORMANCE ANALYSIS

In this section, we can analyze the performance of this method by using metrics such as confusion matrix, accuracy, precision, recall and F_1 score. The Confusion Matrix is a tabular representation of the model's performance. It provided the breakdown of the number of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) predictions. Accuracy measures the overall correctness of the system's predictions,

$$accuracy = ((TP + TN))/((TP + TN + FP + FN)) \quad \text{(Eq.1)}$$

Precision measures the proportion of correctly predicted positive instances out of the total instance predicted as positive.

$$precision = TP/(TP + FP) \quad \text{(Eq.2)}$$

Recall known as sensitivity or true rate which measure the proportion of correctly predicted positive instances out of the total actual positive instance

$$recall = TP/(TP + FN) \quad \text{(Eq.3)}$$

The F_1 score combines precision and recall into a single value, providing a balanced measure of the system's accuracy by considering both false positives and false negatives.

$$F\_1 = (precision * recall)/(precision + recall) \quad \text{(Eq.4)}$$

By using these parameter metrics we can provide the higher detection rate and prevention rate in this methodology.

VI. RESULT AND DISCUSSION

Agenda, AvosLocker, Black Basta, BlackCat/ALPHV, CHOP, Conti, Cuba, Dagon, DarkSide / BlackMatter, FarGo/ TargetCompany, GandCrab, HolyGhost, Hive, Koxic, Lilith, LockBit, Magniber, Maze, Medusa, Moisha, NightSky, Nokoyawa, ONYX, Phobos, Pysa, Ragnarok, Ryuk, SolidBit, STOP/Djvu, Sugar, SunCrypt, Surtr, WannaCry, and Yashma/ Chas a few of the ransomware variants that the suggested method was able to detect valuable files form. Considering most ransomware families do not encrypt concealing files, we were still able to access our files regularly. The recovery function was able to restore files effectively, proving how well our strategy works to provide affordable, all-around ransomware defense.

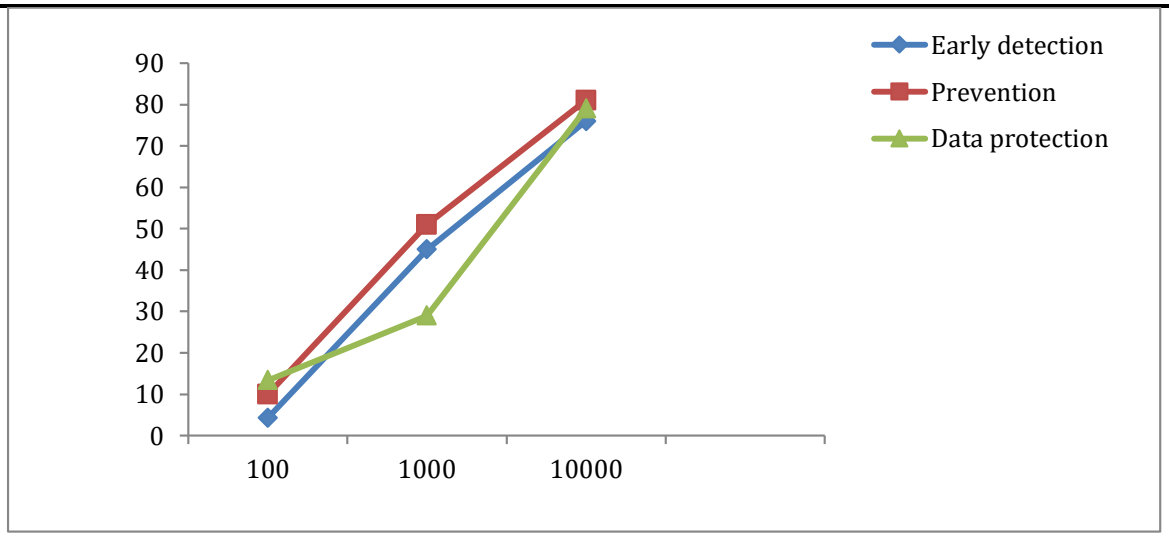

Figure2. Data trained Versus Performance

VII. CONCLUSION

In order to defend against ransomware attacks, this study suggests a novel tactic that detects the ransomware early and prevents from the threat and concealed important data by using information obtained from the viewpoint of the attacker. By successfully obscuring target files, the technique makes it more difficult for ransomware to find and encrypt important data. The technique is improved and the attack surface decreased with the addition of an encrypted database. We have taken usability into account while preserving security, as this is frequently a deciding factor in the effective adoption of defensive methods. The suggested approach proved successful in protecting important files in an economical way through tests using real- world ransomware strains, indicating its usefulness as a backup line of defense in situations where the primary detection techniques might not work. Future work anticipated to target additional enhancements in areas improving usability or practical applications.

REFERENCE

1. J. Choi, J.Lee, G.Lee, J.Yu, and A.Park, " A defense mechanism against attacks on files by hiding files, "J.Korea soc.Ind.Inf.Syst., vol.27, no.2, pp.1–10, 2022, doi:10.9723/jksiis.2022.27.2.001
2. J.Yuste and S.Pastrana, "Avaddon ransomware: An in-depth analysis and decryption of infected systems, '' Comput. Secur., Vol.109, Oct.2021, Art.no.102388,doi:10.1016/j.cose.2021102388.
3. S.Homayoun, A. Dehghantanha, M. Ahmadzadeh, S. Hashemi, and R. Khayami, "Know abnormal, find evil: Frequent pattern mining for ransomware threat hunting and intelligence, " IEEE Trans, Emerg. Topics Comput., vol. 8, no,pp.341–351,apr.2020.
4. Sophos 2020 threat report. Accessed: Jul. 20, 2019. [online]. Available: https://www.sophos.com/en-us/medialibrary/Gated-Assets/whitepapers/sophos-the-state-of-ransomware-2020-wp.pdf
5. K.Lee, S.Lee, and K.Yim, "Machine Learning based file entropy analysis for ransomware detection in backup system," IEEE Access, vol.7,pp.110205–110215,2019.
6. B.Zhou, A. Gupta, R.Jahanshahi, M. Egele, and A. Joshi, "Hardware performance counters can detect malware: Myth or fact?" in Proc. Asia Conf. Comput Commun. Secur.,May 2018, pp.457–468.
7. Daryle Smith, Sajad Khorsandroo , Kaushik Roy " Machine learning Algorithms and Frameworks in Ransomware Detection" IEEE Digital Object Identifier 10.1109/ACCESS.2022.3218779.
8. Xueqin Zhang , Jiyan Wang, Shinan Zhu " Dual Generative Adversarial Networks Based Unknown Encryption Ransomware Attack Detection" Digital Object Identifier 10.1109/ACCESS.2021.3128024.
9. I. GoodFELLOW, J. Pought-Abadie, and M. Mirza, "Generative adversarial nets," in Proc. Adc. Neural Inf. Process. Syst.,2014, pp.2672_2680.
10. H. Lee, S. Han, and J. Lee, "Generative adversarial trainer: Defense to adversarial perturbations with GAN," 2017, axXiv:1705.03387.
11. S. Poudyal and D. Dasgupta, " Analysis of crypto-ransomware using MLbased multi-level profiling," IEEE Access, vol.9, pp.122532–122547,2021, doi:10.11099/ACCESS.2021.3109260.
12. S. Sheen, K.A. Asmitha, and S. Venkatesan, "R-Sentry:Deception based ransomware detection using file access patterns," Comput.Electr.Eng., vol.103,Oct.2022, Art.no.108346, doi:10.1016/j.compeleceng.2022.108346.
13. Y. Lemmou, J.Lanet, and E.M.Souidi," A behavioral in-depth analysis of ransomware infection," IET Inf.Secur., vol.15, no.1,pp. 38–58, Jan.2021, doi:10.1049/ise2.12004.
14. Chia-Ming Hsu; Chia-Cheg Yang; Han-Hsuan Cheng; 'enhancing File Entropy Analysis to Improve Machine Learning Detection Rate of Ransomware', IEEE Access, Digital Object Identifier 10.1109/ACCESS.20213114148.

15. WANPING LIU, Member IEEE " Modeling Ransomware Spreading by a Dynamic Node-Level Method" IEEE Digital Object Identifier 10.1109/ACCESS.2019.2941021.
16. S. Alzahrani, Y. Xiao, and W. Sun, "An analysis of conti ransomware leaked source codes, IEEE Access, vol.10, pp.100178–100193, 2022,doi: 10.1109/ACCESS.2022.3207757.
17. G. Hull, H. John, and B.Arief, Ransomware deployment methods and analysis: Views from a predictive model and human responses, " Crime Sci., vol.8, no. 1, pp.1–22, Feb 2019, doi:101186/s40163-019-0097-9.