CERIAS

The Center for Education and Research in Information Assurance and Security

RANSOMWARE VS MALWARE CLASSIFICATION USING SUBGRAPH MINING OF FUNCTION CALL GRAPHS

Garvit Agarwal[1], Feng Li[1]

1.Department of Computer Information and Graphics Technology, IUPUI.

Abstract

- **Ransomware vs. Malware Differentiation:** This study focuses on the nuanced differentiation between malware and ransomware, addressing the gap in existing research that largely treats malware detection as a homogeneous challenge.
- **API Call Graphs from Cuckoo Sandbox:** Leveraging Cuckoo Sandbox reports, the research extracts dynamic API call data to construct function call graphs that represent the behavior of malware and ransomware samples.
- **Subgraph Mining and Feature Vectorization:** Through subgraph mining, the study identifies critical subsets of API calls, which are then vectorized to capture the characteristics indicative of either malware or ransomware.
- **Deployment of 1D-CNN: A 1D Convolutional Neural Network** (1D-CNN) is employed to classify the vectorized data, demonstrating high precision in distinguishing between the two types of malicious software.

Introduction

- **Need for Differentiation:** Despite the prevalence of malware detection studies, there is a significant gap in research specifically focused on distinguishing ransomware from general malware, which is critical for prioritized incident response.
- **Behavioral Analysis with Cuckoo Sandbox:** The study employs Cuckoo Sandbox to conduct a behavioral analysis of malicious software, extracting API call sequences that form the basis for differentiation.
- **Graphical Representation of Behavior:** A novel methodology is introduced to represent the behavior of malware and ransomware through function call graphs, enabling a structured approach to understanding and classifying software actions.
- **Enhanced Detection Objectives:** The goal is to enhance existing malware detection systems with the ability to distinguish ransomware, facilitating more precise threat mitigation strategies.

Methodology

Dataset Creation & FCG Construction

- Dynamic API call data for malware and ransomware samples are gathered using Cuckoo Sandbox, creating a foundational dataset for analysis. Construct function call graphs from API sequences, categorizing calls as network-related (1), file-related (2), or other (0) to capture the essence of the behavioral patterns.

Subgraph Extraction

- Label each graph node with appropriate tags based on the categorization of the API calls to reflect their relevance to network or file operations. Implement subgraph mining to identify and extract significant subgraphs within the larger function call graphs, focusing on areas indicative of malicious behavior.

Vectorization of Subgraphs

- Transform the extracted subgraphs into numerical feature vectors that encapsulate the structure and frequency of API calls. Adjust the length of feature vectors to fit the model requirements and encode the classification labels for supervised learning.

Deployment of 1D-CNN for Classification

- Develop a 1D-CNN architecture tailored to handle the sequential nature of the feature vectors derived from API call sequences. Train the 1D-CNN model using the prepared dataset, rigorously evaluating its performance in classifying the samples accurately as malware or ransomware.

RESULTS

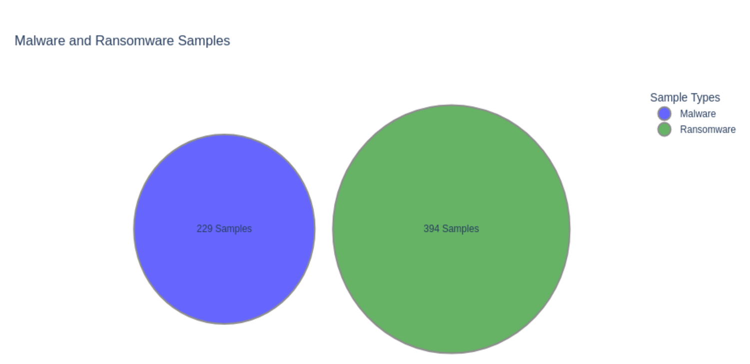

Fig 1. Samples in the Created Dataset

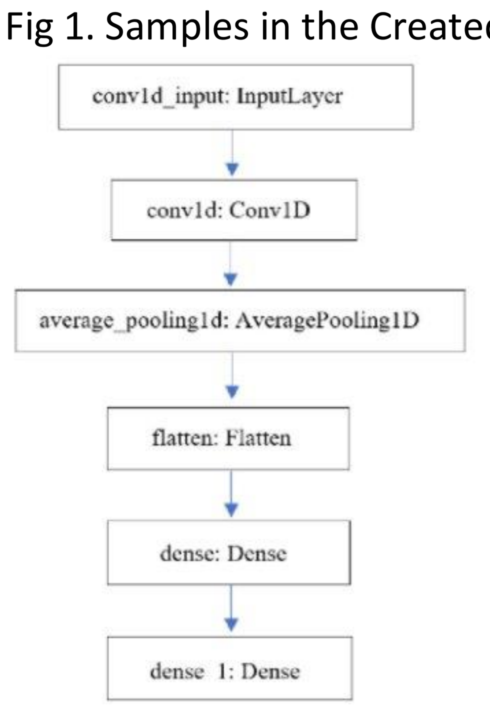

Fig 2. The 1D-CNN Model utilized

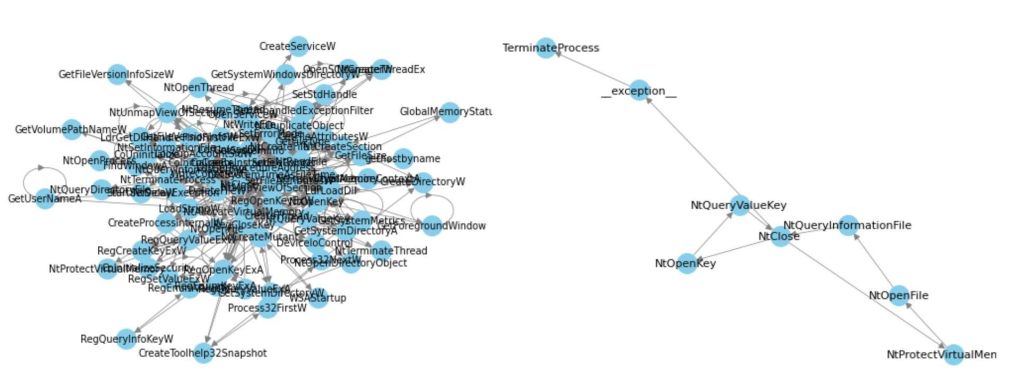

Fig 3. An example of a FCG and its extracted subgraph

```
Tested max_distance: 1, Accuracy: 0.7267779111862183
Tested max_distance: 2, Accuracy: 0.8116835951805115
Tested max_distance: 3, Accuracy: 0.8824383020401001
Tested max_distance: 4, Accuracy: 0.9179971218109131
Tested max_distance: 5, Accuracy: 0.9005805253982544
Tested max_distance: 6, Accuracy: 0.9546444416046143
Tested max_distance: 7, Accuracy: 0.9397677779197693
Tested max_distance: 8, Accuracy: 0.9223512411117554
Tested max_distance: 9, Accuracy: 0.9288824200630188
Tested max_distance: 10, Accuracy: 0.9550072550773621
Tested max_distance: 11, Accuracy: 0.9444847702980042
Tested max_distance: 12, Accuracy: 0.9433962106704712
Tested max_distance: 13, Accuracy: 0.9270682334899902
Tested max_distance: 14, Accuracy: 0.914731502532959
Tested max_distance: 15, Accuracy: 0.9063860774040222
Tested max_distance: 16, Accuracy: 0.9288824200630188
Tested max_distance: 17, Accuracy: 0.9383164048194885
Tested max_distance: 18, Accuracy: 0.9187228083610535
Tested max_distance: 19, Accuracy: 0.927431046962738
Best max_distance: 10 with accuracy: 0.9550072550773621
```

Fig 4. Finding the optimal distance between graph nodes for subgraph mining.

	precision	recall	f1-score	support
malware	0.99	0.99	0.99	831
ransomware	0.99	1.00	1.00	1925
accuracy			0.99	2756
macro avg	0.99	0.99	0.99	2756
weighted avg	0.99	0.99	0.99	2756

Fig 5. Final classification results

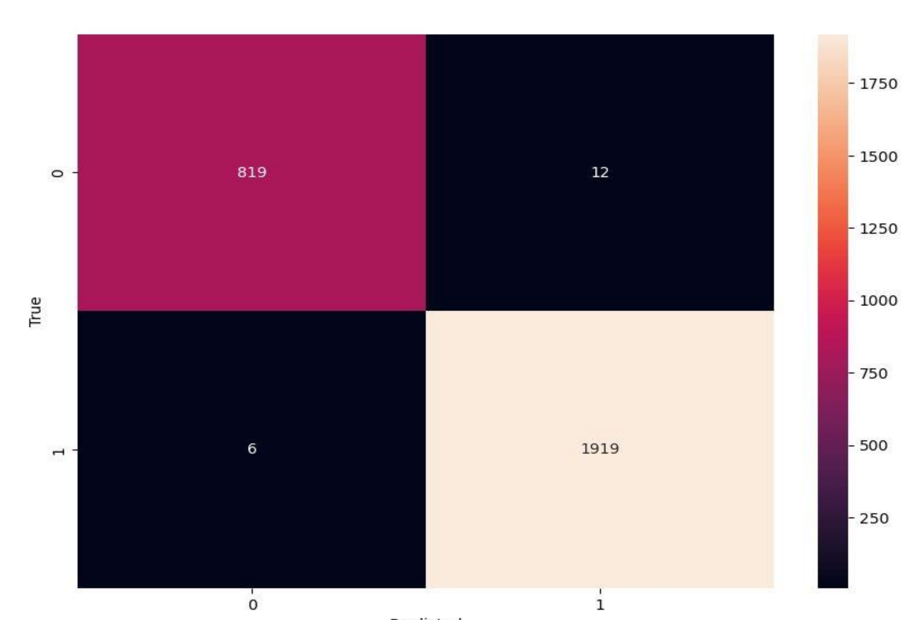

Fig 6. Final confusion matrix for subgraph classification

Future Works

- •**Extended Dataset Inclusion**: We aim to expand the dataset to include a wider variety of ransomware and malware samples, ensuring the robustness and generalizability of the model.
- •**Feature Expansion and Optimization**: Expanding feature sets to include static analysis attributes such as binary file characteristics and applying feature selection techniques to refine the model input for optimal performance.
- •**Network Behavior Profiling**: Integrating network traffic analysis to provide a more comprehensive view of malware communication patterns, which could help to further distinguish between different types of threats.
- •**Improved Machine Learning Models**: Investigating the applicability of recurrent neural networks (RNNs) and other sequence-based deep learning models that are well-suited for the temporal nature of API call sequences and network traffic data.
- •**Adaptive Learning Mechanisms**: Incorporating adaptive learning mechanisms that can update the model in response to new malware and ransomware strains, maintaining high classification accuracy over time.
- •**Adversarial Attack Scenarios**: Testing the model against adversarial attack scenarios to evaluate the robustness of the model against evasion techniques used by sophisticated malware and ransomware.
- •**Behavioral Correlation Analysis**: Conducting correlation analysis between API calls and network behavior to identify patterns that are highly indicative of ransomware, thus fine-tuning the classification process.
- •**Multi-platform Compatibility**: Ensuring the approach is effective across different operating systems and environments, addressing the diversity of platforms that malware and ransomware may target.

PURDUE UNIVERSITY

CERIAS