Summary of literature review

Sr.No	Year	Author(s)	Focus of the paper	Technique(s) used
1	**2018**	Kul Prasad Subedi, Daya Ram Budhathoki, Dipankar Dasgupta	**It applies data-mining techniques to correlate multi-level code components for finding unique signatures to identify ransomware families and uses a combined approach to unveil the hidden intent of the program**	**Data-mining techniques to correlate multi-level code components for identifying ransomware families and a combined approach to unveil the hidden intent of the program.**
2	2019	Abdullahi Mohammed Maigida, Shafi'i Muhammad Abdulhamid, Morufu Olalere, John K	**Static malware analysis, specifically examining ransomware attacks and detection mechanisms. It delves into the structure and behaviour of ransomware, presents a taxonomy of attack techniques, and analyses research datasets for future study of ransomware anatomy.**	**Cryptographic-based methods, anomaly detection, and statistical monitoring of processes and file directories. These techniques are employed to examine the structure and behaviour of ransomware for detection and prevention.**
3	**2022**	SALEH ALZAHRANI 1, YANGXIAO 1, AND WEI SUN	**Static analysis involves using tools to extract information from ransomware files without executing them, including strings and import functions. It also addresses the challenges posed by obfuscated techniques and anti-debugging mechanisms used by recent ransomware families to hide their behavior during static analysis**	**Static analysis techniques such as using tools like PeStudio, x64dbg, and BinaryNinja to analyse Portable Executable (PE) files of ransomware and extract strings and import functions for static analysis. These techniques help identify the ransomware's capabilities before execution**

4	**2023**	Kenan Begovic ∗, Abdulaziz Al-Ali, Qutaibah Malluhi	**To analyse detection techniques for cryptographic ransomware encryption, with an emphasis on API and system calls, I/O monitoring, and file system activities monitoring, particularly in the context of machine learning.**	**It focuses on API and system calls, I/O monitoring, and file system activities monitoring, with a particular emphasis on the use of machine learning**
5	**2022**	FATIMAH ALDAUIJI, OMAR BATARFI, AND MANAL BAYOUSEF	**Analysing the use of cyber threat hunting techniques to detect and respond to ransomware attacks. It provides a survey of the state of the art in utilizing cyber threat hunting techniques to find ransomware attacks, including an investigation of ransomware research directions and available datasets**	**To find ransomware, including cyber threat intelligence (CTI) methods, data analysis methods, machine learning, deep learning, and behavioural analysis. It also explores the use of practical CTI approaches and different cyber threat hunting (CTH) models, as well as the evolution of ransomware attacks and research directions.**
6	**2023**	Juan A. Herrera-Silva and Myriam Hernández-Álvarez	**To create a dynamic feature dataset for ransomware detection using machine learning algorithms, with the goal of identifying ever-evolving ransomware signatures and predicting new variants of ransomware**	**Detecting ransomware using machine learning algorithms include dynamic analysis using a sandbox, feature extraction, and the application of machine learning algorithms such as Gaussian naive Bayes, random forest, gradient boosted trees, and artificial neural networks**

SRS

1)Functional requirements:

- The system shall be available at all times.
- The system shall be usable without an internet connection.
- The system shall include the following parts: Training dataset, ML training model, Summarized results of training, Ransomware detection script.
- The training dataset shall consist of PE headers of the various ransomware families, threat intelligence via VirusTotal API calls and dynamic analysis features via a sandbox.
- The ML training model shall take the training dataset as input and produce a summary output of the ransomware features
- The ML model shall produce the output using the random forest classifier.
- The ransomware detection script shall require the user to run it and provide the path of the file to be analysed.
- The ransomware detection script shall compare the signature of the file provided by the user with the summarised results produced by the ML model and arrive at a conclusion.
- The ransomware detection script shall conclude whether the given file is a ransomware or a benign file.

2)Non-functional requirements:

- The dataset includes at the least 15-30 ransomware families.
- The VirusTotal API calls are executed in advance using a script to avoid network connection requirements.
- The system as a whole must aim for a storage size under 1GB.
- The Sandbox environment summary is also fed to the dataset in advance by executing the dynamic analysis and then generating the report to be added to the dataset.

3)Hardware requirements:

- CPU: Intel i3 10^{th} gen and above
- GPU: Integrated GPU or Nvidia GT 710 and above
- RAM: 4GB
- Storage space: 2GB
- Operating System: Linux based OS(Kali Linux, Ubuntu, etc)

4)Communication interfaces:

- VirtualBox with Windows 10/11 iso and/or Cuckoo sandbox

5)APIs:

- VirusTotal API

UML diagrams

1)System flow diagram

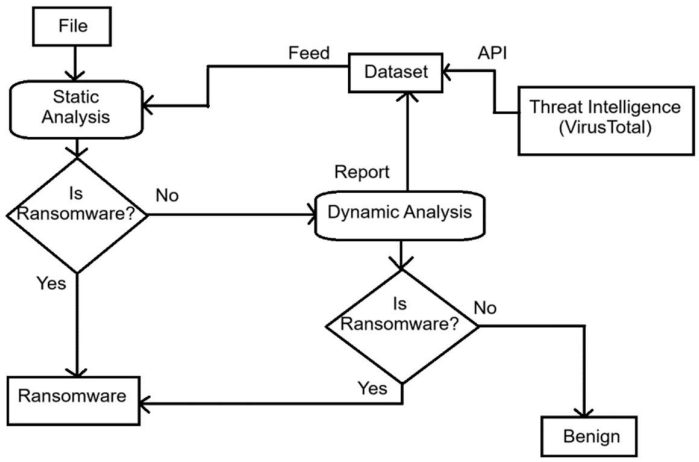
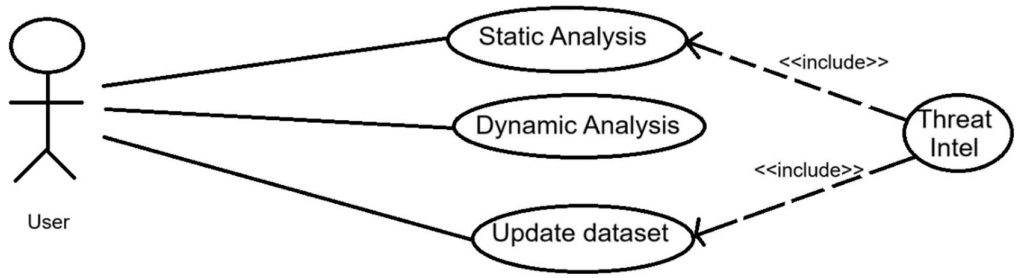


Fig: System flow diagram

---

2)Use case diagram



User: The developer or anyone who has an installed copy of the system

Fig: Use case diagram

---

3)Sequence diagram