CYBER VISION A.I

-Team Members

Abhinand I

Minhaj P

Sabin Santhosh

PROBLEM DEFINITION

Nowadays, there is increasing use of android apps in daily life, which poses a lot of cyber incidents. There is a chance of vulnerability appearing in android apps, which may lead to cyber threats.

In order to analyze android applications and detect android malwares, we proposed an android detection method using system calls of apps and network data. The sequence of system calls can help in identifying attack behaviors.

To collect enough information for malware analysis, we gathered apps from the play store and analyzed and traced them using the Monkey tool and strace.

The collected sequence was then given to a transformer-based model known as BERT to develop the android malware detection model. Furthermore, we employed explainability to explain the model's decision.

OBJECTIVES

Main Objective

Develop a comprehensive malware detection and threat analysis system capable of effectively identifying and explaining potential security threats on Android devices.

Specific Objectives:

- Analyze Android applications for vulnerabilities.
- Collect data on system calls and network traffic.
- Develop a transformer-based model for malware detection.
- Implement explainability techniques to interpret model decisions.

SOLUTIONS

Data Collection and Analysis

- Collected Android applications from the Play Store.
- Analyzed apps using tools like Monkey and Strace to understand system calls.

Development of Malware Detection Model

- Used collected system call sequences to build a strong Android malware detection model.
- Leveraged BERT, a transformer-based architecture, to capture malware patterns effectively.

Interpretability and Transparency

- Incorporated explainability techniques into the model.
- Aimed to clarify why certain apps are identified as potential malware, enhancing understanding and transparency.

FEASIBILITY STUDY

- **Technical Feasibility**: Availability of tools like Strace and Monkey for data collection; transformer models like BERT for analysis.
- **Economic Feasibility**: Cost of tools and resources required for data collection and analysis.
- **Operational Feasibility**: Suitability of the proposed methodology within existing operational frameworks.

REQUIREMENT SPECIFICATION

- **Hardware:**

  Android device for application execution and data collection

- **Software:**

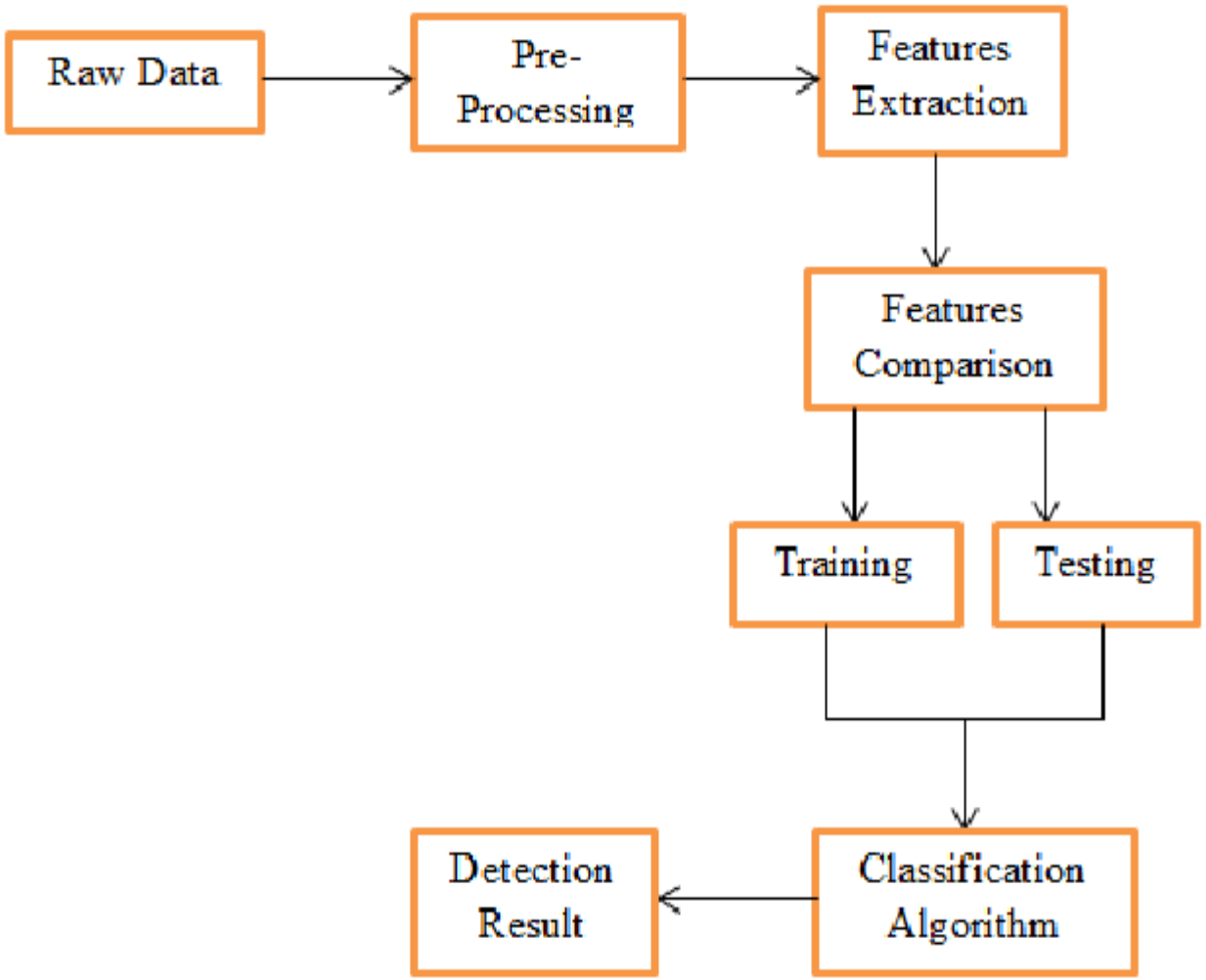
  Android development tools (SDK)

  Tools for system call and network traffic analysis (Monkey tool, strace)

  TensorFlow or PyTorch for implementing the BERT model

  Libraries for explainability techniques

FLOW DIAGRAM



METHODOLOGY

Approach:

Analyze application behavior through system calls and network traffic analysis.

System Calls:

Low-level interactions revealing application intents.

Network Traffic Analysis:

Monitoring data exchanges for communication patterns.

Integration of System Call and Network Traffic Analysis:

Identify behavioral patterns indicative of malicious intent.

Utilization of Transformer-based BERT Model:

Extract features efficiently.

Classify applications based on behavior.

Result:

Enhanced security through accurate malware detection.

PROTOTYPE DESIGN FOR ANDROID MALWARE DETECTION

- **Data Collection:**

  Download apps from Play Store (anonymize data).

- **Application Analysis (Sandbox):**

  Use Monkey tool to simulate user interactions.

  Capture system calls (libstrace).

  Monitor network traffic (libpcap).

- **Feature Engineering:**

  Extract relevant features from system calls (frequencies, sequences).

  Analyze network traffic data (packet size, protocol, destinations).

  Combine features for model training.

- **BERT Model Training:**

  Fine-tune pre-trained BERT model on labeled data (benign/malicious).

  Train to classify new apps based on extracted features.

Outcomes:

Analyze apps, extract features, classify as benign/malicious.

Gain insights into model's reasoning for improved detection.

Future Work:

Explainability - Use some tools understand model decisions(malware indicators).

Thank You...