Parshvanath Charitable Trust's

A. P. SHAH INSTITUTE OF TECHNOLOGY

(Approved by AICTE New Delhi & Govt. of Maharashtra, Affiliated to University of Mumbai)
(Religious Jain Minority)

Android Malware Detection Using Machine learning

Group no. 11

Rishab Agrawal 17204006
Vishal Shah 17204003
Sonam Chavan 16104067

Project Guide & Co-guide
Prof. Ganesh gourshete
Prof. Nahid Shaikh

Contents

- Abstract
- Introduction
- Objectives
- Literature Review
- Problem Definition
- Existing System Architecture/Working
- Proposed System Architecture/Working
- Scope Of Project
- References

Abstract

Android OS experiences a blazing popularity since the last few years. This predominant platform has established itself not only in the mobile world but also in the Internet of Things (IoT) devices. This popularity, however, comes at the expense of security, as it has become a tempting target of malicious apps. Hence, there is an increasing need for sophisticated, automatic, and portable malware detection solutions.

Introduction

- Malware or Malicious Software is defined as software designed to distort and interrupt the mobile or computer applications, collect important information and hence perform malicious operations.
- These malicious operations include gaining access over private information, covertly steal this valuable information over the system, display undesirable advertisement, and spy on the activities of the users.

Objectives

- To analyze the malware from the apk files and from the applications.
- To effectively detect the malware and it will give an warning to the user.
- To watch and scan the apps that which file is performing which type of activity if any suspicious activity found it will detect that.
- To Experiments on real-world Apps with more samples validate the algorithm performance.

Literature Review

Paper title:-Android mobile security by detecting and classification of malware based on permissions using machine learning algorithms

Authors:-Varma, P. Ravi Kiran, Kotari Prudvi Raj, and KV Subba Raju

Publication:-2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC). IEEE, 2017.

Findings:-They have collected different apps from google play store and the collect the standard malware dataset then they extracted manifest.xml file then they identified the permissions given to each app and convert it to arff file

Disadvantages:-It is time taking as it is extracting manifest file and also it doesn’t have GUI

Paper title:-Android Malware Detection Using Machine Learning on Image Patterns

Authors:-Darus, Fauzi Mohd, Salleh Noor Azurati Ahmad, and Aswami Fadillah Mohd Ariffin.

Publication:- 2018 Cyber Resilience Conference (CRC). IEEE, 2018

Findings:-GIST descripter is used to extract feature from the images and classification is done using 3 different machine learning descripter (KNN) k nearest neigbour,Random Forest(RF) and decision tree

Disadvantages:-Some malware sample could not be generated into images.because the APK files are either corrupted or they did not classes.dex file

Paper title:-An Android Behavior-Based Malware Detection Method using Machine Learning

Authors:-Chang, Wei-Ling, Hung-Min Sun, and Wei Wu.

Publication:-2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). IEEE, 2016.

Findings:-In the existing system they have 4 main components preprocessing,data monitor ,decision model and evaluation result

Disadvantages:-They have mainly classified data usingRandom Forest no other algorithm is used

Problem Definition

- Sometimes it has been impossible for the user to identify the malware in his own system.
- And at a time when he gets to know about that malware it has actually destroyed his system.
- Malware enter the system while user give permissions to application.
- User don’t even get to know that when the malware is entered in a system
- Therefore, it also can reduce the damage or loss of data from our android devices.

Solution Provided

- We will divide Android.apk file into two categories,malware and benign.
- We will unpack the apk file to extract classes.dex file which contain the dalvik opcode.These opcode can run on android platform.
- Then we will convert it to 8 bit grayscale of image to visualize the file structure.
- Then will use GIST Descriptor to extract features from these images and will classify them using 3 different machine learning descriptor.KNN,RF,DT.

Existing System

- There are two types of detection techniqes used by malware analyst in existing system

- Static

Static detection is based on specific strings from the dissassembled code without executing the binary file.

- Dynamic

It analyses the malware behaviour such as network activity,system calls and file operations by executing the malware.

- In existing system they are using 300 benign APK file.
- They only able to generate 183 and 300 benign gray scale images
- The other 117 malware sample could not be generated into image because the apk file was either corrupted or they did not classes.dex file

Proposed System

- In this system we are detecting the maximum accuracy by using the machine learning algorithms on image features generated from the APK samples.
- The image will be generated from the maximum APK's samples consisting of Malware and benign samples, and the features are extracted using the descriptor .
- For validating our system will collect the maximum samples of android apps and those have to extracted for each and every application .

Scope of Project

- This will be helpful to determine which type of malware is intruding to your system.
- This system will be helpful to provide security to the android application as android user increasing day by day.
- It will detect the malware which is installed in an by using machine learning.

Conclusion

This system will be able to detect the malware more accurately as we are using different classification algorithm for malware with the large data set availability.

References

- Varma, P. Ravi Kiran, Kotari Prudvi Raj, and KV Subba Raju. "Android mobile security by detecting and classification of malware based on permissions using machine learning algorithms." 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC). IEEE, 2017.

- Darus, Fauzi Mohd, Salleh Noor Azurati Ahmad, and Aswami Fadillah Mohd Ariffin. "Android Malware Detection Using Machine Learning on Image Patterns." 2018 Cyber Resilience Conference (CRC). IEEE, 2018.

- Chang, Wei-Ling, Hung-Min Sun, and Wei Wu. "An Android Behavior-Based Malware Detection Method using Machine Learning." 2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). IEEE, 2016.

THANK YOU!!