MALWARE DETECTION IN ANDROID APP STORE USING ROBUST ATTRIBUTE GENERATION

Shivani Y M[1], Dr. Ch Ramesh[2]
Computer Network and Information Security, Jawaharlal Nehru Technological University
[1] Department of Information Technology, GNITS, Telangana, Hyderabad yalalashivani@gmail.com
[2] Assistant Professor, Department of Information Technology, GNITS, Hyderabad
chramesh23@gmail.com

ABSTRACT: Now a days usage of android devices is becoming quite more due to day-to day increasing the number of active users. With the first-class platform for create gaming and applications android as gained admiration among all other smartphone. It will also allow users to sell and distribute apps instantly and also offers ample free third-party download and install application from the Google play store and third-party play store. Today more the 350,000 sample of new malware are found per every day and 5.2% increase compared to last year.

Mobile devices can control and track the Internet of Things (IOT) services and make them prone to attacks by various malicious application. Conventional methods fail to detect malicious application because of the growth in numbers, variants, and advancement of the malware. Moreover, android allows to installation and downloading apps from unverified sources. For detecting malware application current solution are not appropriate with the rise of the malware and there are limited by low detection accuracy, advanced implementation and heigh computational cost and power. Therefore, android malware become a real-world challenge. To solve this problem an automatic intelligence technique is required for detecting malware apps with the use of real-word datasets which analyses the source code. To show the uniformities in apps which hold malware content we proposed API calls and use Chi- square for SelectKbest best feature selection and examined combination with six supervised machine-learning algorithms (Random Forest, Decision Tree, K-Neighbours, AdaBoost, SVM Linear, and Naïve Bayes). The detection accuracy of these six algorithms is analyse to identify the most effective classifier for detecting malware.

key words: malware, android devices, machine learning, android features, Feature extraction, features selections, Random Forest, Decision Tree, K-Neighbours, AdaBoost, SVM Linear, Naïve Bayes and chi-square

I. Introduction:

Android has become the most common operation system for tablets, MacBook and smartphone with an estimated market share of 70% to 80%. Android has overtaken many other mobile operating systems to become one of the most popular mobile platforms in the world. Recognition of smartphone and other kind of mobile devices has drop down significantly. Android has gained huge admiration among all other smartphone it is the first-class platform for creating gaming and application which allows to distribute and also offer free third party's apps to install and download from Google play store for selling and distributing android apps.

The Internet of Things (IoT) is an attractive system that connects several logical objects and physical devices with networks to enlarge their communication capabilities. In previous years, the IoT has gained popularity owing to technological advancements in areas like artificial intelligence, cloud computing, machine learning, deep learning, application system, and smart home devices. Now a days mobile users are interested in using online transaction for example using smartphone for paying bills, online shopping, trading, booking tickets etc. such portable device is targeted by hacking and become vulnerable to attack more and more.

Malware refer to malicious code which harms user integrity, availability, and confidentiality. Malicious activity is hidden in the background and appear as a clean application. Some examples of the android malware are stealing user personal information (e.g., bank account number, bank credentials, contact number, passwords), tracking user location, send premium SMS messages that cost more than the standard ones, streaming videos from users' cameras, send SMS or mails which contain malware links, and encrypting personal information such as video, images, SMS and contact.

The shared report by IDC for smartphone increases 25.5% global share market in the 1rd quarter of the year 2021[1], the total market share of Android was 72.83% [2] and more than 3.8 billion active users around the world in 2021.

Android is one of the most popularly used mobile device in the world, because of its framework, compatibility, technological impact, open-source, higher success ratio, inter app

integration, high level multitasking, user friendly development environment, etc. On the other hand, it also has risk in installing and downloading applications from unauthorize web site or third-party. Because of the open -source applications android is getting more prone to attack. It has become target to malware, even though it provides security mechanisms.

To prevent malware attacks, developers and researchers are working in different security solution by applying data science, artificial intelligence, machine learning and deep learning. We can analysis android malware with two techniques static analysis and dynamic analysis approaches. In static analysis we get information about software which are going to analyses without executing it but ware as in dynamic analysis is doing

5	The authors [7] proposed to combine permission and API to use machine learning classifiers to detect malware.	Detected 1200 malware apps and 1200 benign apps
6	The authors [8] proposed extracted feature vector from Android Manifest file and combines PI and	Results shows bet against other traditio permission.