A survey of Android malware Detection Technique

Suman R.Tiwari
Silver oak college of engineering, Ahmedabad, India.

Abstract –Signature based approach is used for detection of malware on web based application, which needs more storage memory ,processing power, hence it is not suitable for handheld devices like android smart phone. In this paper, we have surveyed about different types of malware of android devices and different types of technique which can be used to detect android malware.

Index Terms – Signature Based, iOS, API, Permission

1. INTRODUCTION

The number of smart phone users grows from 2.1 billion in 2016 to around 2.5 billion in 2019[6]. Over 36 percent of the world's population is projected to use a smart phone by 2018, up from about 10 percent in 2011[6]. Google's Android and Apple's iOS are the two most popular smart phone operating systems currently used in industry. Android, with 80 percent of all smart phones sales, leads the market[6]. Figure 1.1 shows mobile phone users worldwide.

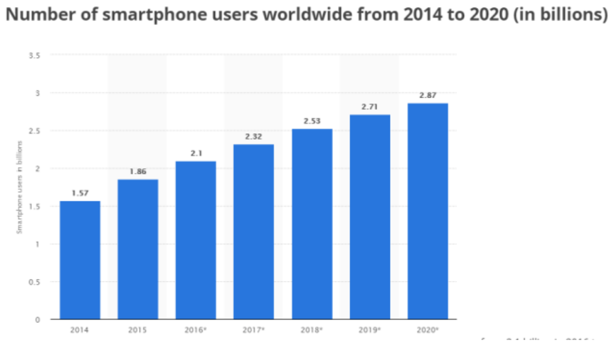

Figure 1 : Trend of Mobile phone users

Based on Trend Micro™ Mobile App Reputation data, 80% of the 1 million apps are malicious in nature[5]. A report by Trend Micro published on 22 June -2016 claimed that 90% of all Android devices are affected by a new family of malware called Godless[5]. According to Norton Mobile Survey, 50% of all Indian smart phone users allowed access to their phone's contacts, and other valuable mobile data , while 40% allowed access to their phone's camera and other hardware features[5]. Google has Released Google Bouncer which is used to scan all the application of Android Play store automatically but Researchers claimed that it is possible to bypass Bouncer security mechanism[5]. Hence any of android phone over internet is high over risks and have threat to their privacy and security.

Malware (short form of Malicious Software), is term used to denote all unwanted software programs. Malware comes in various forms like Viruses, worms, Trojans, spyware, adware, rootkit and ransomware etc. Malware is used as weapon from Hacker community or cyber criminal to compromise victims system, steal personal information, implement DOS and DDOS attack, sending spam emails and for phishing attack.

Following section provides brief description about various malware .

A virus is software which attach themselves with other program or file. When user or host execute program or file than attached virus code get activated. So virus needs human intervention to become active.

First known mobile virus was "Timofonica". "Timofonica" sent SMS messages to GSM mobile phones that read (in Spanish) "Information for you: Telefónica is fooling you." These messages were sent through the Internet SMS gate of the MoviStar mobile operator.

Worms are form of virus which have ability to self propagate. Worms are similar to virus because they can self replicated but unlike virus they does not have attach themselves to file or program. They are extremely dangerous than virus because they can travel through network and can infect all the system within network.

Cabir was worms which was detected in 2004 in android device. When a phone is infected, it displays message 'Caribe' every time the phone is turned on. The worm then attempts to spread to other phones in the area using wireless Bluetooth signals.

Gunpoder is another worm which has infected goggle play store officially in several countries.

Trojans are malicious program which pretend to be useful program to user and perform malicious activity in background.

Ginger master is Trojan which exploits vulnerabilities in Gingerbread version of android OS by privilege escalation. In background it creates services that steal information from infected device like userID, phone number , IMEI number , IMSI number ,location etc.

DroidKungFu is a Trojan which obtains root privileges and installs file com.google.search.apk, which contains backdoor and allows files to removed, open home page to be supplied , and 'open we and download and install' application package. this Trojan collects device information and send it to remote server.

Adware is advertising supporting software that automatically renders advertisement on user screen in order to generate revenue to its master.

Rootkit is malicious software which hides its existence using mask and access the area of devices which is not allowed for un-authorized user. Rootkit access the system for administrator level privileges and perform malicious activity.

Table 1 : Literature Survey Summery

Title/Author	Year & Publication	Description	Analysis
SIGPID[1]	IEEE-2016	Multilevel data pruning to optimize permissons	Does not work on skewed dataset Optimized to 22 permissions
Yang Su[3]	IEEE-2016	Features: Permissions, native permissions, intent priority and sensitive function. Dynamic Analysis check SMS sending facility only.	User manually need to run app in sandbox environment.
CREDROID[2]	ACM-2016	Apk score , Reputation score of URL , Information sent over server and protocol used for communication are analyzed	Can not detect sample it they are not generating traffic. Internet connection is always needed as online tools are used.

2. RELATED WORK

Lichao Sun et. al proposed SigPID[1] in which they have analyzed permissions extracted from Androidmanifest.xml. They have used three levels of data pruning(MLDP - multi level data pruning) methods to filter out permissions ,which helps them to remove extra permissions without affecting on accuracy. First level of pruning is with PRNR(Permission Ranking with Negative Rate) which is used to find out only those permission which contribute more for malware detection. Second level of pruning is SPR(support based permission Ranking) if support(Ranking) for any permission is too low then it is deleted from malware detection features. Third level pruning is PMAR (Permission mining with Association rule) which removes features if they always comes together ex. READ_SMS and WRITE_SMS. After all level of pruning they have obtained only 22 permissions which gives near about same accuracy given by most dangerous 24 permission issued by Google.

Malik et al. proposed CREDROID[2] which used 3rd party tools for detection of malware. They have focused research on Apk score , reputation score of URL , Information sent over server and protocol used for communication . Main limitation of this method is that they can not analyze app which is malicious and not generating network traffic. Internet connection is always needed otherwise application cannot be analyzed.

Miang-Yang Su[3] has proposed method which is based on hybrid analysis. They have used permissions , native functions , intent filter and sensitive Functions. They extracted all features and classified with various classifier like K-NN, SVM, Naive bays, Logistic Regression from which K-NN gives them highest accuracy . For dynamic analysis a Sandbox environment is used where user can run app to analyze its behavior .Main Limitation of this method is , it is non optimized method and only SMS sending facility is checked during dynamic analysis , they does not analyzed receiving number or content of message.

Table 1 shows the summery of literature survey.

3. MALWARE DETECTION TECHNIQUE

The Signature based method is traditional and popular method that is used to detect malware. This method stores the signature of existing malware and detects signature of incoming malware with stored signature. It needs to update existing signature database with new signature of malware for higher accuracy and require high processing power, storage capacity which is limited with android phone. So there is need of a system which should not be based on signature based method and can detect malicious application without using signature.

3.1. Static Analysis Technique

Static analysis is most popular method for malware detection and better for preliminary level analysis. Almost 80% of research paper based on static analysis method or this method are combined with another method for better accuracy. Static Analysis Technique Analyze android application before it is installed on device. In this method APK file is decompiled and analysis is done based on code . This method based on extraction of permissions, intent filter , broadcast receiver , method name, class name , package name from existing APK file and analyze these strings for malware detection.

Main advantage of this method is there is no need to install app within system for analysis and major drawback is time required for reverse engineering of application and can used for initial screening of app.

3.2 Dynamic Analysis Technique

Dynamic Analysis technique analyze android application after it is installed in android device. This technique is also known as behavior analysis technique. For dynamic analysis application are installed and executed in sandbox or in isolated environment and its runtime behavior is observed. Network traffic analysis is example of dynamic analysis where network packet are captured and analyzed.

Main advantage of dynamic analysis method is that it checks behavior of app during execution and detects if malicious activities found from app. This method can be used to provide advance level of security and can detect polymorphic and encrypted malware if analyze properly. Major drawback of this method is that app needs to be installed inside system before analysis.

3.3 Hybrid Analysis Technique

To combine the advantage of static and dynamic analysis hybrid analysis method is used. This technique initially perform static analysis followed by dynamic analysis. Main drawback is that system become complex.

Sr.No	**Static Analysis**	**Dynamic Analysis**
1	Static analysis can anlyze application befor it is installed into user device.	Dynamic analysis analyzes the app after it is installed in the user's device.
2	They do not need to run continuously in users device in background.	Need to run continuously on the user's device in background.
3	As they do not need to run into a user's device in background they need less processing capacity.	Need more processing capacity
4	They consume less power	Consume more power
5	The Mobile device has limited memory and power capacity, hence static is more favorable at initial level of screening.	Hence dynamic analysis is not favorable with device having less processing and power capacity**.**
6	Need to perform reverse engineering on application, which takes time and response becomes slower.	No need to perform reverse engineering on application, hence faster than static analysis.
7	Can detect malware before it breaks security.	Detect malware after application shows its behavior or after security is break**.**

Table 2 Static and dynamic approach summery

4. CONCLUSION

Signature based approach requires more memory and processing capacity, hence for android malware detection technique static or dynamic technique can be used . Due to advantage of static analysis like entry level screening , analyze the app before it installed into device and no need to run always in background static technique is more favorable than dynamic method.

REFERENCES

[1] L. Sun, Z. Li, Q. Yan, W. Srisa-an and Y. Pan, "SigPID: significant permission identification for android malware detection," 2016 11th International Conference on Malicious and Unwanted Software (MALWARE), Fajardo, 2016, pp. 1-8.

[2] Malik, J., & Kaushal, R. (2016, July). CREDROID: Android malware detection by network traffic analysis. In Proceedings of the 1st ACM Workshop on Privacy-Aware Mobile Computing (pp. 28-36). ACM

[3] Su, M. Y., & Fung, K. T. (2016, July). Detection of android malware by static analysis on permissions and sensitive functions. In Ubiquitous and Future Networks (ICUFN), 2016 Eighth International Conference on (pp. 873-875). IEEE.

[4] Li, X., Liu, J., Huo, Y., Zhang, R., & Yao, Y. (2016, August). An Android malware detection method based on AndroidManifest file. In Cloud Computing and Intelligence Systems (CCIS), 2016 4th International Conference on (pp. 239-243). IEEE.

[5] Malicious and High-Risk Android Apps Hit 1 Million [online]. : https://www.trendmicro.com/vinfo/au/security/news/mobile-safety/malicious-and-high-risk-android-apps-hit-1-million-where-do-we-go-from-here.

[6] Number of smartphone users worldwide from 2014 to 2020 (in billions). : https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

[7] Google Bouncer: http://blog.trendmicro.com/trendlabs-security-intelligence/a-look-at-google-bouncer/