Grenze International Journal of Engineering and Technology, June Issue

Mobile Malware Attacks, Classification, Propagation, Analysis, Detection, Challenges and Future Directions – A Survey

Dr. B Senthilkumar[1], Dr. M Sujithra[2] and Mani Barathi SP S[3]
[1]Associate Professor, Department of Mechanical Engineering, Kumaraguru College of Technology
Email: senthilkumar.b.mec@kct.ac.in,
[2]Assistant Professor, Department of Computing – Data Science, Coimbatore Institute of Technology
Email: sujithra@cit.edu.in
[3]M.SC Data Science (Integrated Master's Degree Program), Department of Computing – Data Science, Coimbatore Institute of Technology

***Abstract*— Number of Smartphone users are increasing day by day and mobiles have become an integral part of the society. This is because of the rich variety of mobile devices and essential applications provided by its manufacturers. The increasing number of mobile devices invite skilled developers and hackers to develop malware that invades personal and business information in a very efficient manner. Therefore, mobile devices are an ideal target for various security issues and data privacy threats in a mobile ecosystem. Threats posed by malwares include leaking of private information, financial loss to users, system damage. For better protection on computers, researchers and manufacturers making great efforts to produce anti-malware systems with effective detection methods why the most targeted platform for the malware developers is Android and why and how and when malwares propagate into the mobile system with how they are detected, and protection mechanism is discussed in this paper.**

***Index Terms*— Detection algorithm, mechanism, malware, machine learning approach, threats, vulnerabilities.**

I. INTRODUCTION

In recent times, the use of mobile devices for both business and personal purposes has increased significantly. Modern tablets and Smartphone's provide many useful services such as internet browsing, maps, social network clients, internet banking in addition to standard mobile functionality including phone calls, SMS, and Bluetooth. The data used and stored in these services is often highly sensitive and therefore desired by the attackers. Mobile devices may have become the most popular gadgets, but their security is still a developing domain. this is a rising significance and a cumulative need, but it is comparatively weak area for the user's data privacy and protecting. Although mobile companies do think for the user's security and data privacy, the use of applications from the internet creates complex issues in relation to handling threats and vulnerabilities when securing a user's data privacy.

There are thousands of diverse applications accessible from application stores for each mobile device, and these applications have an extensive range of purposes, including web browsing, entertainment (movies, games, and music), social networking, communication (e-mail, internet messaging), banking, and location-based services.

Grenze ID: 01.GIJET.9.2.343

The reasons for the motivations of existence of malware are

- User Information Stealing
- Publicity of User Data
- Spam SMS
- Optimization of search process
- Ransom

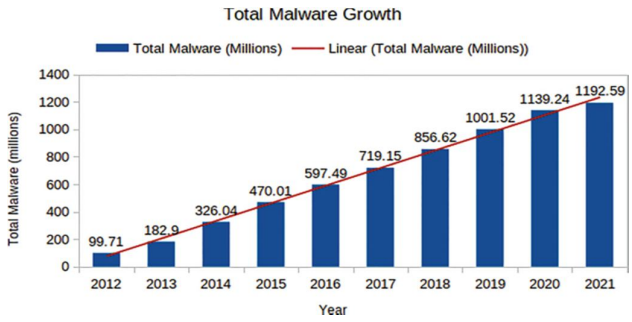

Figure 1: Malware growth in relation with time

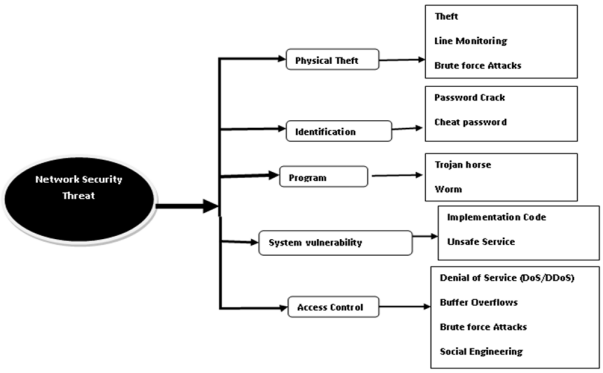

Figure 2: Threats and attacks on mobile OS

Security goal is the processes of achieving are confidentiality, integrity, and availability.

- Confidentiality is referred as the process of preserving the data from the unauthorized confinements and implies to the proprietary information and personal privacy security.
- Integrity refers to the process of safeguarding the information from the unauthorized attacker's process of destruction or modification and further the guarantee for the authenticity and non-reputation of information.
- Availability is characterized as guaranteed access and utilization of data within the assured time. The term availability ensures that the information and data used are utilized in timely.

Figure 3: Data protected by CIA trait

The important aspects of mobile device security and data privacy issues is discussed. Sensitive security issues affecting on smart phones such as malware attacks, vulnerabilities, and Threats is addressed. Classification of malware, malware propagation and the types of malware analysis techniques is discussed. This survey presents the trusted security countermeasures, various malware detection techniques to help the users to protect their devices. Research questions for future works are introduced in this review. This paper is organized as follows. In

Section 2, the malware classification is discussed. In Section 3, this paper discussed the malware propagation. Section 4 presented the malware analysis and detection approaches taxonomy. In Section 5, discussed the malware detection mechanism. In section 6, we have presented the performance evaluation techniques. Future directions are described in Section 7. In section 8, conclusion is presented.

II. MALWARE CLASSIFICATION

The term malicious software is the abbreviation of the term malware. Put simply, any piece of software that was written with the intent of doing harm to data, devices or to people is called a malware. Malware classification can be done in several ways to distinguish the unique their types from each other. for better understanding of how they can infect computers and devices, the threat level they pose and how to protect against them, Distinguishing and classifying them from each other becomes imperative. The depending on the type of damage malware infuses it can be helpful in categorizing what kind of malicious software you're dealing with. The common types of malwares, but are as follows

- VIRUS: like biological science, viruses glue themselves to clean files and it is contagious and may attack other clean files. They can spread hysterically, and cause damage toa system's core functionality and deleting or corrupting files. They usually appear as an executable file.
- TROJANS: malware which differentiate its own self as the legitimate software or is included in legitimate software that has been tampered with does belong to trojans family. It creates backdoors in your security to let other malware in and may also tend to act discretely.
- SPYWARE: No surprise here: spyware is malware designed to spy on you. It hides in the background and takes notes on what you do online, including your passwords, credit card numbers, surfing habits and more.
- WORMS: Worms infect entire networks of devices, either local or across the internet, by using network interfaces. It uses each consecutive infected machine to infect more.
- RANSOMEWARE: Also called scareware, this kind of malware can lock down your computer and threaten to erase everything unless a ransom is paid to its owner.
- ADWARE: Though not always malicious in nature, particularly aggressive advertising software can undermine your security just to serve you ads which can give a lot of other malwares a way in. Plus, let's face it: pop-ups are annoying.
- BOTNETS: Botnets are networks of infected computers that are made to work together under the control of an attacker.

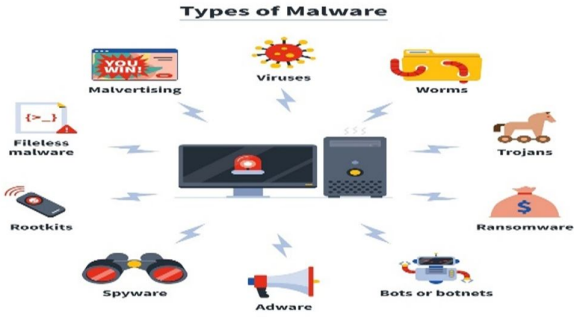

Figure 4: Malware types

III. MALWARE PROPAGATION

There are different types of malware propagation, they are

- REPACKAGING: Malware developers first download any popular app, disassemble that app (i.e., generating the source code written in Java), insert their own code having malicious payload within the original code, reassemble the app and redistribute that app in official or third-party app markets.
- UPDATE ATTACKS: Repackaging technique includes malicious payload within the original app but that is easier to detect by analyzing the source code. To evade detection malware developers instead of including malicious payload within the app, they include only an update component which downloads the malicious

payload at run time after the app is installed on the device. Hence scanning the source code will not be able to detect the malware as initially there is no malicious code within the app.

- DRIVE BY DOWNLOADS: This technique employs traditional drive-by-downloads to Android devices as well in which users are enticed to download interesting or attractive apps. For example, Tracker malware has in-app advertisement library. After clicking on that advertisement link user is redirected to a website which displays the message to download an app which can save battery of the device. However, that downloaded app is a malware which subscribes to premium rate services without user's knowledge.

Figure 5: Types of malware propagation

IV. MALWARE ANALYSIS TECHNIQUES

Malware can be analyzed with the help of detection techniques. Malware analysis is the method or a process of understanding the code, behavior, and functionality of malware so that critical act of attack can be measured. static analysis, dynamic analysis and permission-based analysis are the three broad categories of Detection techniques- parameters-static code analysis, taint tracing and control flow dependencies are the ways in which static analysis can be done with this is depicted in the figure given below. Dynamic analysis considers parameters including-network traffic, native code, and user interaction. Permission-based analysis can be done with the help of permissions specified in manifest file. As told in literature, various techniques exist for detection of mobile malware.

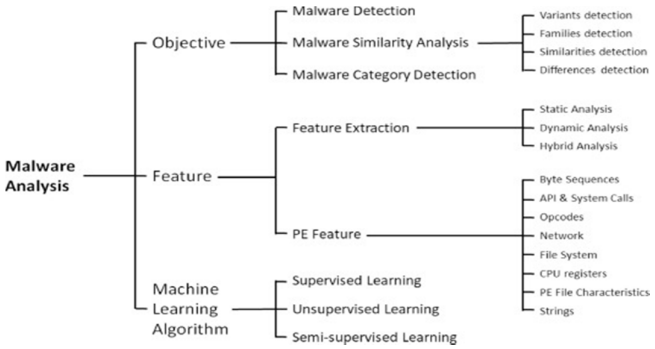

Figure 7: Malware analysis techniques

- Static Analysis: Static analysis can perform the investigation in the downloaded app by inspecting its own software properties and its source code. It is an inexpensive way to find malicious activities in code segments without executing application and notifying its behavior. Many techniques can be used for static analysis: decomplication, decryption, pattern matching and static system call analysis etc. The obfuscation and encryption techniques embedded in software makes static analysis hard. Static analysis is further divided into two categories-
  - misuse detection
  - anomaly detection traditionally used by anti-viruses.
- Misuse detection: signature-based approach for detection of malware based on security policies and rulesets by matching of signatures is where the Misuse detection is used. Data flow dependency and control flow dependencies in source code that would help to understand the behavior of apps, this is possible in case of static analysis.
- Anomaly detection: machine learning algorithms for learning of known malwares and predicting unknown malware is used in case of Anomaly detection. identifying action of malware rather than pattern is the most

common application of this approach. The procedure undergone here are first used to construct suspicious behavior of applications and then observed signatures are matched against database of normal behavior applications. by training network with classifier such as support vector machine (SVM) It can distinguish between malicious and normal behavior.

- Dynamic Analysis: Dynamic analysis perform the implementation of application in secluded environment to keep in track of its execution behavior. Various heuristics are considered for monitoring dynamic behavior which includes-monitoring network activity, file changes and system call traces. Android applications can run in an Android SDK, a mobile device emulator running on desktop computer for emulation of software and hardware features except generating phone calls. For testing purposes emulator supports Android Virtual Device (AVD) configurations. When applications start running on the emulator, it can use all services like to invoke other applications, accessing network state, play audio and video, store and retrieve data.
- Permission Analysis: Permissions play key role while analyzing android applications. They are listed in Manifest.xml file while each application is installed. Install time permissions limits application behavior with control over privacy and reduces bugs and vulnerabilities. Users have right to allow or deny the installation of applications, but he cannot go for the selection of individual permissions. These permissions are required in android applications because the use of resources in android phones is based on this permission set. Some researchers can detect malicious behavior of android applications based on permissions specified in Manifest.xml.

V. MALWARE DETECTION MECHANISM

Malware detection system is a method to examine if a program has malicious or non-malicious. Detection system includes two process-

- Analysis
- Detection.

It takes two inputs one is the signature or behavioral parameters of a given code and second is the program under inspection, it can employ its own detection mechanism to decide if the program is malware or not. They can be divided into

- signature-based detection
- specification-based detection
- behavior-based detection
- Application Permission Analysis
- Cloud Based Malware Detection
- Data Mining Based Malware Detection.
- Signature-Based malware detection

A pattern-marching approach commercial antivirus is an example of signature-based malware detection where the scanner scans for a sequence of byte within a program code to identify and report a malicious code. This approach to malware detection adopts a syntactic level of code instructions to detect malware by analyzing the code during program compilation. This technique usually covers complete program code and within a short period of time. However, this method has limitation by ignoring the semantics of instructions, which allows malware obfuscation during the program's run-time.

A. Specification-Based Malware Detection

Specification based detection makes use of certain rule set of what is considered as normal to decide the maliciousness of the program violating the predefined set of the rules. The specification-based system there exists a training phase which attempts to learn all valid behavior of a program or system which needs to inspect. The main limitation of specification-based system is that it if very difficult to accurately specify the behavior the system or program.

B. Behavioral-Based Detection

The behavior-based malware detection system is composed of several applications, which together provide the resources and mechanisms needed to detect malware on the Android platform. Each program has its own specific functionality and purpose in the system and the combination of all of them creates the Behavior-Based malware detection system. For collecting data from Android applications, the Android data mining scripts and

applications mentioned in are the responsible, and the script running on the server will be the responsible for parsing and storing all collected data.

C. Permission –based Detection

Applications run in a sandbox environment however they need permissions to access certain data. At the time of installation, Android platform asks the user to grant or deny permission for the application based on the activities the application can perform. This is to overcome a limitation in Android platform where the developers can intentionally hide permission label to a component. If no label is specified there is no restriction as it had default allow policy.

D. Cloud Based Malware Detection

The Google Play apps are examined for the malware. Bouncer are the service used automatically to examine the apps on the Google Play Store for malware. As soon as an application is uploaded, the Bouncer checks it and then compares it to other known malware, Trojans, and spyware. Every application is run in a simulated environment to see if it will behave maliciously on an actual device.

E. Data Mining Based Malware Detection

In data mining methods for detecting malicious executables, a malicious executable as a program that performs function, such as compromising a system's security, damaging a system, or obtaining sensitive information without the user's permission. Their data mining methods detect patterns in large amounts of data, such as byte code, and use these patterns to detect future instances in similar data. Their framework used classifiers to detect new malicious executables.

VI. PERFORMANCE EVALUATION METRICS

Evaluation metric plays a critical role in achieving the optimal classifier during the classification training. Thus, a selection of suitable evaluation metric is an important key for discriminating and obtaining the optimal classifier. This paper systematically reviewed the related evaluation metrics that are specifically designed as a discriminator for optimizing generative classifier. Generally, many generative classifiers employ accuracy as a measure to discriminate the optimal solution during the classification training. However, the accuracy has several weaknesses which are less distinctiveness, less discriminability, less informativeness and bias to majority class data.

Metrics	Formula	Evaluation Focus
Accuracy (acc)	$\frac{tp+tn}{tp+fp+tn+fn}$	In general, the accuracy metric measures the ratio of correct predictions over the total number of instances evaluated.
Error Rate (err)	$\frac{fp+fn}{tp+fp+tn+fn}$	Misclassification error measures the ratio of incorrect predictions over the total number of instances evaluated.
Sensitivity (sn)	$\frac{tp}{tp+fn}$	This metric is used to measure the fraction of positive patterns that are correctly classified
Specificity (sp)	$\frac{tn}{tn+fp}$	This metric is used to measure the fraction of negative patterns that are correctly classified.
Precision (p)	$\frac{tp}{tp+fp}$	Precision is used to measure the positive patterns that are correctly predicted from the total predicted patterns in a positive class.
Recall (r)	$\frac{tp}{tp+tn}$	Recall is used to measure the fraction of positive patterns that are correctly classified
F-Measure (FM)	$\frac{2*p*r}{p+r}$	This metric represents the harmonic mean between recall and precision values

Figure 8: Performance Evaluation Metrics

VII. FUTURE DIRECTIONS

In case of security and privacy, the Smartphone users are not able to figure out the number of attacks on their devices and how much money malicious apps may steal from their accounts. In this survey, first, we have discussed different types of the mobile device vulnerabilities and threats, Secondly, we have classified malware and malicious applications focusing on how the attack is executed and what is the target of the attackers. Finally, discussed the possible malware detection defense mechanisms for mobile device security and then, suggested some future directions to improve the detection of malicious or abnormally behaving applications before its propagation. Using new machine learning techniques for providing real-time behavior analysis and identifying

fake apps. Deep learning algorithms can be utilized for the features extraction with more accuracy during the malware testing. The Mobile OS companies, especially popular ones, should consider more security mechanisms for preventing against unpredictable attacks.

VIII. Conclusion

Smart phones are becoming popular in terms of power, sensor, and communication. With the rapid proliferation of the Smartphone gadgets and developing apps with a lot of features, as several sensors and connections, the number of malware and attacks is raising. Modern, smart phones provide lots of services such as messaging, browsing internet, emailing, playing games in addition to traditional voice services. Increase in the number of smart phones on the market, the need for malware analysis is an urgent issue. Malware is a critical threat to user's computer system in terms of stealing

- confidential information
- corrupting or disabling security system.

This survey paper explains some occurred technologies used by security researchers to gear these threats. It says malware types, static, dynamic and hybrid malware analysis techniques, malware detection mechanisms. Among the various existing approaches Machine Learning methods have shown the results with high Accuracy in the detection of malicious activities. With this categorization, we want to provide an easy understanding for users and researchers to improve their knowledge about the security and privacy of smart phones.

References

[1] Ammar Ahmed E. Elhadi, MohdAizainiMaarof and Ahmed Hamza Osman, Malware Detection Based on Hybrid Signature Behaviour Application Programming Interface Call Graph, American Journal of Applied Sciences 9 (3): 283-288, 2020, ISSN 1546-9239, 2021, Science Publications

[2] Sujithra, M., Padmavathi, G., (2012): A Survey on Mobile Device Threats, Vulnerabilities, and their Defensive Mechanism. International Journal of Computer Applications (0975-8887) Volume 56-- No.14

[3] Kirti Mathur, Saroj Hiranwal, A Survey on Techniques in Detection and Analyzing Malware Executables, International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277 128X, Volume 3, Issue 4, April 2020.

[4] Sujithra, M.; Padmavathi, G.: Enhanced permission-based malware detection in mobile devices using optimized random forest classifier with PSO-GA. Res. J. Appl. Sci. Eng. Technol. 12, 732–741 (2016)

[5] Matthew G. Schultz, Eleazar Eskin, Erez Zadok, and Salvatore J. Stolfo, Data Mining Methods for Detection of Study New Malicious Executables, in Proceedings of the Symposium on Security and Privacy, 2021, pp. 3849.

[6] Sujithra M, Padmavathi G, Narayanan S (2015) Mobile device data security: a cryptographic approach by outsourcing mobile data to cloud. Procedia Computer Science. 47:480–485

[7] Jonathan joseph bloun, adaptive rule-based malware detection employing learning classifier systems, Thesis- Master of science in computer science, Missouri University of science and technology, 2022.