Malware Detection System on Android Devices

C.Kotteeswari
Assistant Professor
Department of Computer Science and Engineering
Velar College of Engineering and Technology
Tindal Eroode-638012

Anusha S, Dharanya G Kartheeswaran S
Department of Computer Science and Engineering
Velalar College of Engineering and Technology
Thindal, Erode - 638012

Abstract: Google Play Protect makes it easier to protect your device. Before you download apps from the Google Play Store, it checks for safety. It checks your device for applications from other sources that could be harmful. Malware is another name for these harmful applications. It removes known harmful apps from your device and gives you a warning if any potentially harmful apps are found. It alerts you to apps that violate our Unwanted Software Policy by concealing crucial information. It violates our Developer Policy by sending you privacy alerts about apps that may obtain user permission to access your personal information. Present A Novel Ensemble Learning Using PCA in the proposed system. Which can detect fraud based on a user's spending habits and does not require fraud signatures. Based on the spending profile of the mobile app holder and activity monitoring, it tries to find any anomalies in the transaction. 970 positive reviews and 710 negative reviews were downloaded at random by a Java web crawler.

Keywords: Malware, Developer Policy, Novel Ensemble Learning, Anomalies.

I INTRODUCTION

DIGITAL MEDIA

Google operates and developed the digital sharing service known as Google Play, which was formerly known as Android Market. It allows users to browse and download applications created with the Android software development kit (SDK) and published by Google. It serves as the official app store for the Android operating system. Google Play is also a digital media store that sells music, books, magazines, movies, and television shows. Before the launch of a separate online hardware retailer, Google Store, on March 11, 2015, it offered Google hardware devices for purchase. Applications are available for free or for a fee through Google Play. Through the Play Store mobile app or by installing the application on an Android device from the Google Play website, they can be downloaded. Users of devices with particular hardware components, such as a motion sensor (for motion-dependent games) or a front-facing camera (for online video calling), can be targeted by applications that take advantage of a device's hardware capabilities.

ANDROID APP MARKETS

The commercial success of Android app stores like Google Play and the incentives they provide to popular apps make them attractive targets for dishonesty and fraud. While malicious developers use app markets as a launch pad for their malware, several dishonest developers dishonestly increase the search rank and fame of their apps (for example, through fake reviews and bogus installation counts). The motivations behind such actions are: Increases in app popularity bring about economic benefits and accelerate the spread of malware. As a matter of fact, for advancing cell phone Applications, a competitor list of applications is the principally significant method of upslope on the lookout. An app should be ranked higher based on how quickly its growth chart increases and how quickly it can generate many downloads and, ultimately, revenue. There were several different approaches to promoting apps to rise to the top of the leader boards for apps. The official approach is known as white hat promotion, and developers use this strategy to increase their app’s fame and download. In most cases, this strategy employs "internet bots" or "human water armies" to rapidly increase the number of app downloads, ratings, and reviews. Some requirements are shown as having two restrictions to limit fraud. The first restriction is that an app can only be rated once from a user's login; the second restriction uses IP address support to limit the number of users logged in each day. Finally, historical record real-world app data that will be gathered from the App Store for a long time will be used to evaluate the proposed system.

II LITERATURE SURVEY

[1]FAKE REVIEWS IN GOOGLE PLAY AND GOOGLE APP STORE

The security measures that are currently in place to safeguard Internet services are made more difficult by crowd-sourcing systems. Many of these security measures are based on the idea that computer programs create malicious activity automatically. As a result, when attacks are generated by actual users working in a crowd-sourcing system, they would operate badly or be simple to circumvent. It reveals shocking evidence through measurements that malicious crowd-sourcing systems not only exist but are also rapidly expanding in both user base and total revenue. Examine the specifics of the campaigns that are offered and carried out on these websites, and test their overall effectiveness by running your own active, benign campaigns. Finally, investigate and evaluate the personnel source on various crowdsourcing websites. The results show that these systems' campaigns are successful at reaching users and that their continued expansion is a real threat to online communities. Unexpected evidence showed that malicious crowd-sourcing systems not only exist but are rapidly expanding in both their user base and profits.

[2]SCALABLE AND ACCURATE ZERO-DAY ANDROID MALWARE DETECTION

Malicious apps hide within other normal apps, making it difficult to detect them," Malicious apps spike in the Google Play store. Because they rely on known malware samples for signature mining, the currently available mobile antivirus programs are insufficient due to their reactive nature. It outlines a proactive strategy for identifying zero-day Android malware. This scheme is motivated to evaluate potential security risks posed by these untrusted apps without relying on malware samples or their signatures. Specifically, a scalable examination of a particular app that exhibits risky behaviour (such as initiating a root exploit or sending SMS messages in the background) by an automated system known as Risk Ranker. Using these models, it is believed that some prioritize a list of less important apps that require more research. The risk score can be any function that is inversely related to the probability so that lower probability interprets into a higher score.

[3]PROBABILISTIC GENERATIVE MODELS FOR RANKING RISKS OF ANDROID APPS

"PCMag published an article titled "Top Android App a Scam, Pulled from Google Play." This risk communication method is one of Android's primary defences against malicious apps. Before a user installs an app, it informs the user of the permissions it needs and hopes that the user will make the right choice. Because it presents the risk information for each app in a manner that necessitates an excessive amount of practical awareness and some time to extort valuable information, this strategy has been demonstrated to be pointless. Risk-scoring algorithms make use of probabilistic generative models, which can recognize a variety of models, from basic Naive Bayes models to sophisticated hierarchical mixture models. Real-world datasets were used in the experiments, and the results show that Naive Bayes models offer a promising risk-scoring approach and that probabilistic general models perform significantly better than other methods currently in use. Any function that is inversely related to the probability can be used to calculate the risk score, with a lower probability translating into a higher score.

[4]MACHINE LEARNING APPROACHES

Android OS is one of the most widely used mobile operating systems. Learn how to recognize fake apps on the Google Play store. With the ever-increasing number of mobile devices on the market, the number of malicious applications and malware is only getting bigger. There are a lot of commercial signature-based tools on the market to stop malicious applications from getting into and spreading. According to several studies, traditional signature-based detection systems are effective to a certain extent, and malware authors employ a variety of methods to evade these tools. To evade detection, malware authors employ a variety of strategies, including the (i) code obfuscation technique, (ii) encryption, (iii) inclusion of permissions that the application does not require, (iv) request for unwanted hardware, and (v) download or update attack, in which a benign application updates itself or updates another application with a malicious payload, which is difficult to detect. Numerous studies have demonstrated that malicious activities can be accurately detected using machine learning algorithms.

[5]PERMISSION USAGE TO DETECT MALWARE IN ANDROID

How to tell if an app is fake on the Google Play store. Since the advent of Android devices, the number of programs for this operating system has grown by an incredible amount. Applications are already organized in Google's Android Marketplace, which is highly susceptible to misuse. Developers of malware introduce malicious applications not only to this market but also to a variety of other markets. PUMA is a new machine learning-based approach to detecting malicious Android applications by analysing the permissions extracted from the application itself. Installing mobile applications on these devices has been challenging for users over the past ten years. The issues begin when they locate the application they wish to install. Numerous operating systems, including Symbian, used an authentication system based on certificates to protect the device and prevent piracy. However, this system came with some drawbacks for users such as the fact that they were unable to install applications despite having purchased them and the platforms have employed special strategies to guard against this kind of software.

III EXISTING SYSTEM

Fraud is discovered after it has occurred that is, after the owner of the mobile app files a complaint. Also, a lot of online purchases are made these days, so do not know how people use mobile apps online; instead, just capture the IP address for verification. Therefore, to investigate fraud, cybercriminals need to assist. The regularity of ranking fraud activities, and the efficacy of the proposed system in the experiments. It is difficult to manually label each mobile app for ranking fraud because of the large number of apps, so it is important to have a scalable method that automatically detects ranking fraud without using benchmark data. Therefore, detecting ranking fraud in mobile app-leading sessions is detecting ranking fraud in mobile app-leading sessions. Suggest a straightforward but efficient method for determining the most popular sessions of each app based on its previous ranking records.

IV PROPOSED SYSTEM

A Novel Ensemble Learning Using PCA in the proposed system. Which can detect fraud based on a user's spending habits and does not require fraud signatures. Any bank's Fraud Detection System (FDS), which issues detecting Mobile apps to app holders, typically does not know the specifics of items purchased in individual transactions. The FDS is unaware of the kinds of goods purchased in that transaction. Based on the spending profile of the mobile app holder and activity monitoring, it tries to find any anomalies in the transaction.970 positive reviews and 710 negative reviews were downloaded at random by a Java web crawler. The detection of fraudulent use of Mobile apps is found much faster than the existing system because we focus on consumer opinions.

V METHODOLOGY

REGISTRATION AND LOGIN

The user enters his username and password and chooses any one role from the given list. Choose the role as User then the user details will be displayed. Choose the role of Developer then you upload the app. Choose the role of Admin then it contains the app developer details and mobile app details.

UPLOAD APP

Choose the role as Developer then you upload the app by providing the app category including Art & Design, Auto & Vehicles, Book & Reference, Action, Music, and Finance, etc., then give the app name and upload the app image. Describe the app then upload the app.

RATING BASED ON EVIDENCE

Ranking-based evidence is useful for ranking fraud detection. An app that has been published is not sufficient, it can be rated by the user who downloaded it. A higher rating app may attract more users to download the app. Compared its historical ratings used for constructing rating-based evidence.

REVIEW-BASEDEVIDENCE:

In addition to ratings, most App stores permit users to write textual app reviews. These reviews may reflect the individual perceptions and usage experiences of current mobile app users.

FRAUD APPS DETECTING

Downloading or purchasing read its historical reviews. The app can contain more positive reviews to attract users to download. Imposters often post fake reviews to inflate app downloads. Problems of detecting the local anomaly of reviews in the leading sessions. Extracting types of evidence combined for ranking fraud detection. Users frequently first 5, read its historical reviews to aid in decision-making.

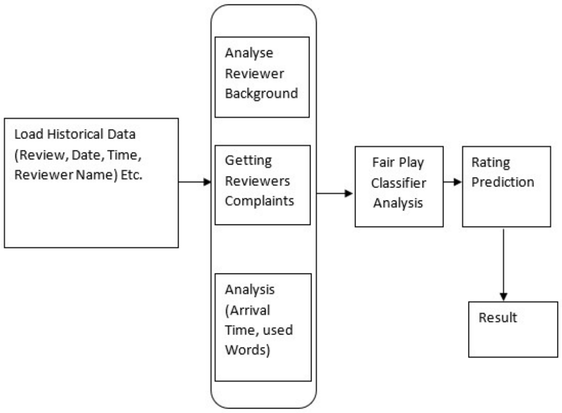

Load Historical Data (Review, Date, Time, Reviewer Name) Etc.
Analyse Reviewer Background
Getting Reviewers Complaints
Analysis (Arrival Time, used Words)
Fair Play Classifier Analysis
Rating Prediction
Result

Figure 1: System Architecture

FAIRPLAY: PROPOSED SOLUTION

Fair Play organizes the analysis of longitudinal app data into the following 4 modules,

1. The Co-Review Graph(Coreg) module identifies apps reviewed in a contiguous time window by groups of users with significantly overlapping review histories.
2. The Review Feedback (RF) module exploits feedback left by genuine reviewers.
3. Inter Review Relation (IRR) module leverages relations between reviews, ratings and install counts.
4. The Jekyll-Hyde (JH) module monitors app permissions, with a focus on dangerous ones, to identify apps that convert from benign to malware.
5. The Review Feedback (RF) module exploits feedback left by genuine reviewers.
6. The Co-Review Graph(Coreg) module identifies apps reviewed in a contiguous time window by groups of users with significantly overlapping review histories.

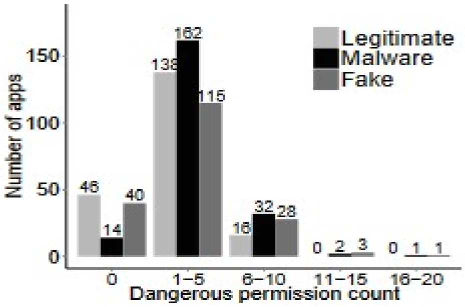
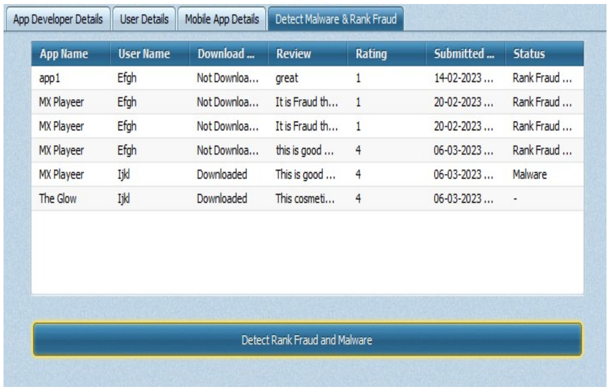


Figure 2: Malware, Fraudulent and benign Apps

Each module produces several features that are used to train an app classifier. Fair play also uses general features such as the app's average rating, and the total number of reviews.FairPlay combines the results of this approach with behavioural and linguistic clues, extracted from longitudinal app data, to detect both search rank fraud and malware apps. Emphasize that search rank fraud goes beyond opinion spam, as it implies fabricating not only reviews but also user app installs events and ratings.

The Coreg module identifies suspicious, time-relatedbehaviours. The RF module uses linguistic tools to detect suspicious behaviours reported by genuine reviews.

The IRR module uses behavioural information to detect suspicious apps. The JH module identifies permission ramps to pinpoint possible Jekyll-Hyde app transitions.



App Name	User Name	Download ...	Review	Rating	Submitted ...	Status
app1	Efgh	Not Downloa...	great	1	14-02-2023 ...	Rank Fraud ...
MX Playeer	Efgh	Not Downloa...	It is Fraud th...	1	20-02-2023 ...	Rank Fraud ...
MX Playeer	Efgh	Not Downloa...	It is Fraud th...	1	20-02-2023 ...	Rank Fraud ...
MX Playeer	Efgh	Not Downloa...	this is good ...	4	06-03-2023 ...	Rank Fraud ...
MX Playeer	Ijkl	Downloaded	This is good ...	4	06-03-2023 ...	Malware
The Glow	Ijkl	Downloaded	This cosmeti...	4	06-03-2023 ...	-

Figure 3: Detect Rank Fraud and Malware

COMPARING WITH THE EXISTING SYSTEMS

Figure 4: Comparisons with existing algorithms

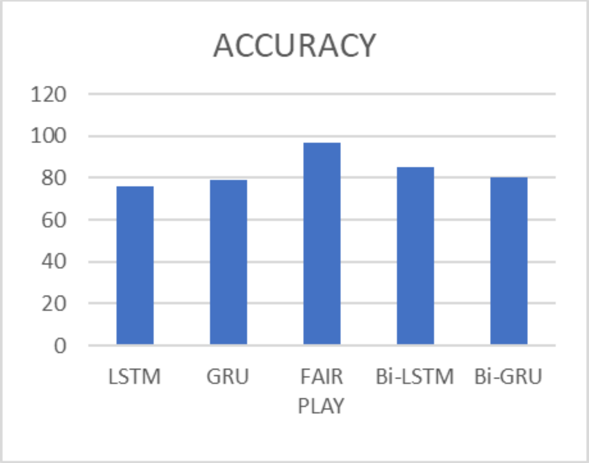


Figure 5: Table for Accuracy and Algorithm

ALGORITHM	**ACCURACY**
LSTM	76.2
GRU	79
FAIR PLAY	96.79
Bi-LSTM	85.44
Bi-GRU	80.5

When compared with the existing algorithm the accuracy of FAIRPLAY is higher because of their protection according to the rules. LSTM provides 76.2% accuracy, GRU provides 79% accuracy, Bi-LSTM provides 85.44% accuracy, and Bi-GRU provides 80.5% accuracy.

VI CONCLUSION

Fairplay is a system for identifying malware and fraudulent Google Play apps. A high percentage of malware is involved in search rank fraud, as demonstrated by experiments on a new longitudinal app dataset; FairPlay correctly identifies both. Additionally, it demonstrated Fair Play’s capacity to uncover hundreds of apps that defy Google Play's detection technology, including a novel form of coercive fraud.

FUTURE ENHANCEMENT

There is a growing risk of malware targeting mobile devices due to the recent emergence of mobile platforms that can execute increasingly complex software and the increasing prevalence of mobile platforms in sensitive banking applications. Due to the limited resources and privileges granted to the user, the problem of detecting such malware presents unique challenges, but also unique opportunities in the required metadata attached to each application. Malware on Android devices is detected by a machine learning-based system. To take advantage of the more powerful computing power of a server or cluster of servers, the system trains a One-Class Support Vector Machine offline (off-device).

REFERENCES

[1] S. Chen, L. Fan, G. Meng, T. Su, M. Xue, Y. Xue, Y. Liu, and L. Xu, "An empirical assessment of security risks of global Android banking apps," in ICSE, 2020.

[2] C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of ELECTRICAL ENGINEERING, Vol.63 (6), pp.365-372, Dec.2012.

[3] G.Neelakrishnan, M.Kannan, S.Selvaraju, K.Vijayraj, M.Balaji and D.Kalidass, "Transformer Less Boost DC-DC Converter with Photovoltaic Array", IOSR Journal of Engineering, October 2013; 3(10): 30-36.

[4] R. Feng, J. Q. Lim, S. Chen, S.-W. Lin, and Y. Liu, "Seqmobile: An efficient sequence-based malware detection system using rnn on mobile devices," in ICECCS, 2020

[5] C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011.

[6] Q. Guo, S. Chen, X. Xie, L. Ma, Q. Hu, H. Liu, Y. Liu, J. Zhao, and X. Li, "An empirical study towards characterizing deep learning development and deployment across different frameworks and platforms," in ASE, 2020.

[7] C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor & Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011.

[8] R.Baskar, R.Jayaprakash, M.Balaji, M.Kannan, A.Divya and G.Neelakrishnan, "Design of Nanoscale 3-T DRAM using FinFET", IOSR Journal of Electrical and Electronics Engineering, November-December 2013; 8(1):1-5.