INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 4 ISSUE 1 (2018) PAGES 2729 - 2732
RECEIVED : 15.01.2018 PUBLISHED : 13.01.2018
Feb 13, 2018

SEARCH RANK FRAUD AND MALWARE DETECTION IN GOOGLE PLAY

[1]V. Hemalatha, ME Scholar, Department Of Computer Science And Engineering, Idhaya Engineering College for women, Chinna Salem-606201.

[2]S. Jayasundar, Assistant Professor, Department Of Information Technology, Idhaya Engineering College for women, Chinna Salem-606201.

ABSTRACT

Fraudulent behaviors in Google Play, the most popular Android app market, fuel search rank abuse and malware proliferation. To identify malware, previous work has focused on app executable and permission analysis. In this paper, we introduce FairPlay, a novel system that discovers and leverages traces left behind by fraudsters, to detect both malware and apps subjected to search rank fraud. FairPlay correlates review activities and uniquely combines detected review relations with linguistic and behavioral signals gleaned from Google Play app data (87K apps, 2.9M reviews, and 2.4M reviewers, collected over half a year), in order to identify suspicious apps. FairPlay achieves over 95% accuracy in classifying gold standard datasets of malware, fraudulent and legitimate apps. We show that 75% of the identified malware apps engage in search rank fraud. FairPlay discovers hundreds of fraudulent apps that currently evade Google Bouncer's detection technology. FairPlay also helped the discovery of more than 1,000 reviews, reported for 193 apps that reveal a new type of "coercive" review campaign: users are harassed into writing positive reviews, and install and review other apps.

1. INTRODUCTION

The commercial success of Android app markets such as Google Play and the incentive model they offer to popular apps, make them appealing targets for fraudulent and malicious behaviors. Some fraudulent developers deceptively boost the search rank and popularity of their apps (e.g., through fake reviews and bogus installation counts), while malicious developers use app markets as a launch pad for their malware . The motivation for such behaviors is impact: app popularity surges translate into financial benefits and expedited malware proliferation. Fraudulent developers frequently exploit crowdsourcing sites (e.g., Freelancer, Fiverr, BestApp Promotion) to hire teams of willing workers to commit fraud collectively, emulating realistic, spontaneous activities from unrelated people (i.e., "crowdturfing"), see Figure 1 for an example. We call this behavior "search rank fraud".

In addition, the efforts of Android markets to identify and remove malware are not always successful. For instance, Google Play uses the Bouncer system to remove malware. Previous mobile malware detection work has focused on dynamic analysis of app executables as well as static analysis of code and permissions. However, recent Android malware analysis revealed that malware evolves quickly to bypass anti-virus tool.

In this paper, we seek to identify both malware and search rank fraud subjects in Google Play. This combination is not arbitrary: we posit that malicious developers resort to search rank fraud to boost the impact of their malware.Unlike existing solutions, we build this work on the observation that fraudulent and malicious behaviors leave behind telltale signs on app markets. Resource constraints can compel fraudsters to post reviews within short time intervals. Legitimate users affected by malware may report unpleasant experiences in their reviews. Increases in the number of requested permissions from one version to the next, which we will call "permission ramps", may indicate benign to malware (Jekyll-Hyde) transition.

2. RELATED WORK

We focus on the Android app market ecosystem of Google Play. The participants, consisting of users and developers, have Google accounts. Developers create and upload apps, that consist of executables (i.e., " apks ") , a set of required permissions, and a description. The app market publishes this information, along with the app's received reviews, ratings, aggregate rating (over both reviews and ratings), install count range, size, version number, price, time of last update, and a list of "similar" apps. Each review consists of a star rating ranging between 1-5 stars, and some text. Google Play limits the number of reviews displayed for an app to 4, 000. Figure 2 illustrates the participants in Google Play and their relations. The Fraudulent developers attempt to tamper with the search rank of their apps, e.g., by recruiting fraud experts in crowd sourcing sites to write reviews, post ratings, and create bogus installs. While Google keeps secret the criteria used to rank apps, the reviews, ratings and install counts are known to play a fundamental part.

To review or rate an app, a user needs to have a Google account, register a mobile device with that account, and install the app on the device. This process complicates the job of fraudsters, who are thus more likely to reuse accounts across jobs. The reason for search rank fraud attacks is impact. Apps that rank higher in search results tend to receive more installs. This is beneficial both for fraudulent developers, who increase their revenue, and malicious developers, who increase the impact of their malware.

3. LITERATURE SURVEY

Zach Miners comes with the number of mobile apps infected with malware in Google's Play store nearly quadrupled between 2011 and 2013, a security group has reported. In 2011, there were approximately 11,000 apps in Google's mobile marketplace that contained malicious software capable of stealing people's data and committing fraud, according to the results of a study published Wednesday by RiskIQ, an online security services company. By 2013, more than 42,000 apps in Google's store contained spyware and information-stealing Trojan programs, researchers said. Apps designed to personalize people's Android-based phones were most susceptible, as well as entertainment and gaming apps. Some of the most malicious apps in the Google Play store downloaded since 2011 were Wallpaper Dragon Ball, a wallpaper app, and the games Finger Hockey and Subway Surfers Free Tips.

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 4 ISSUE 1 (2018) PAGES 2729 - 2732
RECEIVED : 15.01.2018 PUBLISHED : 13.01.2018
Feb 13, 2018

4. PROPOSED SYSTEM

In this, we introduce FairPlay, a novel system that discovers and leverages traces left behind by fraudsters, to detect both malware and apps subjected to search rank fraud. FairPlay correlates review activities and uniquely combines detected review relations with syntactical and behavioral signals gleaned from Google Play app data, in order to identify doubtful apps. FairPlay achieves over 95% accuracy in classifying gold standard datasets of malware, fraudulent and rightful apps. We show that 75% of the identified malware apps engage in search rank fraud. FairPlay discovers hundreds of fraudulent apps that currently evade Google Bouncer's detection technology. FairPlay also helped the discovery of more than 1,000 reviews, reported for 193 apps that reveal a new type of forceful review operation. We uncover these malicious acts by picking out such trails. For instance, the high cost of setting up valid Google Play accounts forces fraudsters to reuse their accounts across review writing jobs, making them likely to review more apps in common than regular users. Resource constraints can compel fraudsters to post reviews within short time intervals.

5. RESULT ANALYSIS

We have implemented FairPlay to extract data from parsed pages and compute the features, and the R tool to classify reviews and apps. We have set the threshold density value u to 3, to detect even the smaller pseudo cliques. We have used the Weka data mining suite to perform the experiments, with default settings. We experimented with multiple supervised learning algorithms. Due to space constraints, we report results for the best performers: Multi Layer Perceptron (MLP), Decision Trees (DT) and Random Forest (RF) , using 10-fold cross validation . For the back propagation algorithm of the MLP classifier, we set the learning rate to 0.3 and the momentum rate to 0.2. We used MySQL to store collected data and features. We use the term "positive" to denote a fraudulent review, fraudulent or malware app; FPR means false positive rate. Similarly, "negative" denotes a genuine review or benign app; FNR means false negative rate. We use the Receiver Operating Characteristic (ROC) curve to visually display the trade-off between the FPR and the FNR. TPR is the true positive rate. The Equal Error Rate (EER) is the rate at which both positive and negative errors are equal. A lower EER denotes a more accurate solution.

CONCLUSION

I have introduced FairPlay, a system to detect both fraudulent and malware Google Play apps. Our experiments on a newly contributed longitudinal app dataset, have shown that a high percentage of malware is involved in search rank fraud; both are accurately identified by FairPlay. In addition, we showed FairPlay's ability to discover hundreds of apps that evade Google Play's detection technology, including a new type of coercive fraud attack.

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 4 ISSUE 1 (2018) PAGES 2729 - 2732
RECEIVED : 15.01.2018 PUBLISHED : 13.01.2018
Feb 13, 2018

REFERENCES

[1] Report Zach Miners. Report: Malware-infected Android apps spike in the Google Play store. PC World, 2014.
[2] Stephanie Mlot. Top Android App a Scam, Pulled From Google Play. PCMag, 2014.
[3] Daniel Roberts. How to spot fake apps on the Google Play store. Fortune , 2015.
[4] Andy Greenberg. Malware Apps Spoof Android Market To Infect Phones. Forbes Security, 2014.
[5] Freelancer. http://www.freelancer.com.
[6] Fiverr. https://www.fiverr.com/.
[7]Yajin Zhou, Qiang Zhang, Shihong Zou, and Xuxian Jiang. Riskranker: Scalable and Accurate Zero-day Android Malware Detection. In Proceedings of ACM MobiSys, 2012.
[8] Bhaskar Pratim Sarma, Ninghui Li, Chris Gates, Rahul Potharaju, Cristina Nita-Rotaru, and Ian Molloy. Android Permissions: a Perspective Combining Risks and Benefits. In Proceedings of ACM SACMAT, 2012.
[9] Hao Peng, Chris Gates, Bhaskar Sarma, Ninghui Li, Yuan Qi, Rahul Potharaju, Cristina Nita-Rotaru, and Ian Molloy. Using Probabilistic Generative Models for Ranking Risks of Android Apps. In Proceedings of ACM CCS, 2012.
[10] S.Y. Yerima, S. Sezer, and I. Muttik. Android Malware Detection Using Parallel Machine Learning Classifiers. In Proceedings of NGMAST, Sept 2014.