A Longitudinal Study of Removed Apps in iOS App Store

Fuqi Lin[*]

Key Lab of High Confidence Software Technologies
(Peking University), Ministry of Education
Beijing, China
linfuqi@pku.edu.cn

Liu Wang

School of Computer Science,
Beijing University of Posts and Telecommunications
Beijing, China
w_liu@bupt.edu.cn

Haoyu Wang[*][†]

School of Computer Science,
Beijing University of Posts and Telecommunications
Beijing, China
haoyuwang@bupt.edu.cn

Xuanzhe Liu[†]

Key Lab of High Confidence Software Technologies
(Peking University), Ministry of Education
Beijing, China
xzl@pku.edu.cn

ABSTRACT

To improve app quality and nip the potential threats in the bud, modern app markets have released strict guidelines along with app vetting process before app publishing. However, there has been growing evidence showing the ineffectiveness of app vetting, making potentially harmful and policy-violation apps sneak into the market from time to time. Therefore, app removal is a common practice, and market maintainers have to remove undesired apps from the market periodically in a reactive manner. Although a number of reports and news media have mentioned removed apps, our research community still lacks the comprehensive understanding of the landscape of this kind of apps. To fill the void, in this paper, we present a large-scale and longitudinal study of removed apps in iOS app store. We first make great efforts to record daily snapshot of iOS app store continuously in a span of 1.5 years. By comparing each two consecutive snapshots, we have collected the information of over 1 million removed apps with their accurate removed date. This comprehensive dataset enables us to characterize the overall landscape of removed apps. We observe that, although most of the removed apps are low-quality apps (e.g., outdated and abandoned), a number of the removed apps are quite popular. We further investigate the practical reasons leading to the removal of such popular apps, and observe several interesting reasons, including ranking fraud, fake description, and content issues, etc. More importantly, most of these mis-behaviors can be reflected on app meta information including app description, app review, and ASO keywords. It motivates us to design an automated approach to flagging the removed apps. Experiment result suggests that, even without accessing to the bytecode of mobile apps, we can identify the removed apps with good performance (F1=83%). Furthermore, we are able to

[*]Co-primary authors ordered by alphabet
[†]Corresponding via haoyuwang@bupt.edu.cn and xzl@pku.edu.cn

flag the removed apps in advance as long as their inappropriate behaviors appear in their metadata. We believe our approach can work as a *whistle blower* that pinpoints policy-violation behaviors timely, which will be quite effective in improving the app maintenance process.

CCS CONCEPTS

• **Security and privacy** → **Software and application security**.

KEYWORDS

app Store, iOS, app mining, removed app

1 INTRODUCTION

Although the mobile app ecosystem has seen explosive growth in recent years, *app quality remains a major issue across app markets*. A number of low-quality apps, fake/cloned apps, illicit apps, Potentially Harmful Apps (PHAs), and even malware were recurrently found in Google Play, iOS app store, and third-party app markets [16, 37, 40, 42], posing great threats to mobile users.

To improve app quality and nip the potential threats in the bud, most app markets including Google Play and iOS app store have released strict developer policies, along with inspection and vetting processes before app publishing. For example, Google Play has released a set of developer policies [17] that cover 10 main categories, and iOS app store has published a detailed review guidelines [23] covering five main parts including safety, performance, business, design and legal. Ideally, mobile apps that break these policies should not be published on app markets.

However, it is challenging to automatically identify low-quality and potentially harmful apps in the app vetting process. There has been growing evidence [1, 5, 6, 34] showing the ineffectiveness of app vetting in both Google Play and iOS app store, i.e., a large portion of low-quality apps and PHAs sneak to app markets from time to time. Even for iOS app store, one of the most secure app markets due to its locked-down ecosystem, malicious apps [5, 34] and

aggressive adware [6] were recurrently found. Furthermore, some malicious and fraudulent behaviors cannot be explicitly identified during app vetting process. For example, previous work [10, 43] suggested that apps can manipulate metadata and generate fake reviews to boost app ranking and search ranking. Some sophisticated apps (especially the HTML container apps) can behave normally during the app vetting process, while showing their illicit or malicious behaviors once getting in the market. These kinds of fraudulent behaviors, however, cannot be explicitly perceived based on binary-level analysis before they sneak into the market.

Removed Apps. As a result, app markets have been removing low-quality and policy-violation apps continuously in a reactive manner, after the apps were listed in app markets for a certain time. For example, the number of apps in iOS app store peaked at around 2.2 million in 2017, but declined over the next few years (around 1.8 million apps in 2020), as Apple continues to remove low-quality apps and apps that break guidelines [31, 32]. Although a number of reports and news media have mentioned removed apps, *our research community still lacks the comprehensive understanding of the landscape of this kind of apps and the app maintenance (i.e., app removal) behaviors of app markets.* There remain a number of unexplored questions, e.g., *how many apps were removed from the app market? were they removed periodically or occasionally? what are the practical reasons behind those app removals?*

This Work. In this paper, we present a large-scale longitudinal study of removed apps. To be specific, we focus on iOS app store, the official and largest app market for iOS apps. We first make great efforts to collect daily snapshot of the whole iOS app market (i.e., information of all available apps), with over 500 complete snapshots in total (from Jan 2019 to April 2020). By comparing the two consecutive snapshots, we can pinpoint which apps were removed, and their accurate removal date (**see Section 3**). *To the best of our knowledge, this is the largest and most comprehensive dataset of removed apps studied in the research community.* This enables us to perform fine-grained analysis of removed apps. We then characterize the overall landscape of removed apps (**see Section 4**), including the daily app maintenance behaviors across the span of 1.5 years, the life-cycle of the removed apps, their correlation with other app features (e.g., app popularity), and the developers of the removed apps. Our investigation suggests that although most of the removed apps are low-quality apps, a number of them are quite popular. It further motivates us to investigate the practical reasons behind the removed popular apps (**see Section 5**). We observe that a number of features from app meta information (e.g., app description, app comments, and app search optimization keywords) can be used to differentiate these removed popular apps effectively. Therefore, we finally devise a practical app removal prediction model (**see Section 6**) based on machine learning techniques to identifying suspicious apps that should be removed.

To the best of our knowledge, this is the first comprehensive study of app removal practice in iOS app store at *scale*, *longitudinally*, and *across various dimensions*. In addition, compared to preliminary work of removed apps in Google Play [38] that are based on rather coarse-grained snapshots, this work comprehensively measured the overall landscape of removed apps including the daily trend, app popularity, and app life-cycle, etc, and derived

various insightful findings. Among interesting results and observations, the following are prominent:

- *The number of removed apps is surprisingly higher than our expectation.* During the span of 1.5 years we studied, over 1 million apps were removed from iOS app market. Interestingly, app removal in iOS app store shows cyclical patterns.
- *Some developers have the tendency to release policy-violation apps.* Surprisingly, 73.45% of removed app developers have all their released apps being removed. Majority of the removed apps are dominated by a limited number of developers.
- *Removed "popular" apps are prevalent.* Although most of the removed apps are considered to be low-quality apps, 5% of them are popular apps with high rank in app store and have gained a considerable number of user ratings.
- *The removed "popular" apps can be identified proactively without analyzing their app binaries.* The features extracted from the app metadata can be used as strong indicators to differentiate the removed apps from normal ones.
- *The removed apps can be identified prior to its actual removal date due to the initial signals it released.* By analyzing the evolution of its meta information only, we can pinpoint the suspicious policy-violation apps timely as soon as their inappropriate behaviors appear in their metadata.

Our results motivate the need for more research efforts to illuminate the widely unexplored removed apps. We believe that our research efforts can positively contribute to bringing user and developer awareness, attracting the focus of the research community and regulators, and promoting best operational practices across app market operators. To facilitate further study along this direction, we have released our dataset to the research community[1].

2 BACKGROUND

2.1 iOS App Store Guidelines

To improve app quality and eliminate potential security issues, Apple requires all iOS apps to go through a vetting process to determine whether they are reliable and perform as expected. The app vetting process adopts manual efforts with automated tools. App reviewers compare the app with Apple's public App Store guidelines[2], making sure it does not show any undesired behaviors. On average, 50% of apps are reviewed in 24 hours and over 90% are reviewed in 48 hours [22]. If the submission is incomplete, review time may be further delayed or the app may be rejected. In general, the App Store Review Guidelines can be summarized as follows:

Safety. Apps should not include objectionable content that is offensive, insensitive, upsetting, intended to disgust, in exceptionally poor taste, or just plain creepy. Apps should not behave in any way that risks physical harm. Apps should implement appropriate security measures to ensure proper handling of user information.

Performance. Apps should be tested on-device for bugs and stability before submission, and include demo account info. The metadata including app description, screenshots, and previews must be accurate and be kept up-to-date. Apps should satisfy both hardware compatibility and software requirements.

Table 1: An overview of the detailed information collected for the removed apps.

Type	Data Field	Description
Objective	App Name	The name of the app
	App ID	The ID of the app given by the app store
	Developer Name	The name of the developer of the app
	App Category	The category of the app in app store
	Price	The price of the app
	Status	Online, Removed or Relaunched
	Release Date	The release date of the app
	Update Date	The latest update date of the app
	Offline Date	The date when the app is removed
	Relaunch Date	The date when the app is relaunched (if existed)
Subjective	App Ratings	The number of {1,2,3,4,5} ratings of the app
	App Reviews	Detailed app reviews from app users
Popularity	App Ranking	App daily ranking in its category
	ASO Keywords	The daily ASO keywords covered by the app

Business. Apps that attempt to manipulate reviews, inflate chart rankings with paid, incentivized, filtered, or fake feedback, and expensive apps that try to cheat users with irrationally high prices will be punished. Ads displayed in an app must be appropriate and be limited to the main app executable. Besides, the payments of the app shall not be irregular.

Design. Apps' design must meet the minimum standards for approval including no copycats, no spam, containing features, content, and UI, and hosting or containing compliant extensions, etc. Apps that stop working or offer a degraded experience may be removed from the App Store at any time.

Legal. Apps must comply with all legal requirements in any location where the app is available. It is paramount in Apple ecosystem to protect user privacy, including data collection and storage, data use and sharing, and location services, etc. Intellectual property should not be infringed. Gambling, gaming, lotteries and VPN apps must not violate local laws.

2.2 App Removal

Although Apple has made great efforts to improve the app vetting process, low-quality and policy-violation apps were found in the market from time to time. App removal is a common practice in the market. Apps can be removed by either app developer or the app market. For the apps removed by app developer proactively, if the app binary has not been modified, the app developer can re-launch the app at any time, without going through the app inspection and vetting process. In contrast, for the apps removed by app market due to policy-violation reasons, the app developers should address all the issues and then submit for the app inspection and vetting process. However, it is not feasible for us to infer whether the app was removed by app market. Besides the apps that break review guidelines, mobile apps can also be removed by iOS app store due to political reasons (e.g., censorship by governments).

3 STUDY DESIGN

We present the details of our characterization study on removed apps in this section. We first describe our research questions (RQs), and then present the dataset used for our study.

3.1 Research Question

Our study is driven by the following research questions:

RQ1 *How many apps were removed from iOS App Store? What is the distribution of removed apps across time (e.g., daily) and space (e.g., category and app popularity) dimensions?* Although a number of reports and news media mentioned app removal from iOS app store from time to time, our research community still lacks the understanding of the overall landscape. We are unaware of the app maintenance behaviors in app markets and the characteristics of the removed apps.

RQ2 *Why were they removed from iOS App Store?* Although app markets have released strict developer policies, along with inspection and vetting processes before app publishing, a number of low-quality and potentially harmful apps were found in app markets after they were released, which then were removed after receiving massive user complaints. Understanding the reasons of app removal can help us understand the weakness of app vetting processes in iOS app store.

RQ3 *Can we identify the removed apps in advance?* Existing efforts of app markets usually rely on reactive methods to identify potentially harmful apps after they sneak into the market, e.g., based on mobile users' feedback or the anomalies/damages caused by apps. Proactively identifying the removed apps before they are widely spread can help eliminate the potential risks they exposed.

3.2 Dataset Collection

To answer the aforementioned research questions, we first need to harvest a comprehensive dataset of removed apps. However, it is non-trivial to compile such a dataset, as it requires great efforts to monitor apps in the app market continuously to check whether they were taken down, especially when we want to know their accurate removal date. To this end, we cooperate with an anonymous leading mobile app intelligence company to get the daily removed apps from iOS app store. The key idea of identifying removed apps is straightforward, i.e., we crawl the iOS app market to maintain a snapshot of the whole market everyday. By comparing the two consecutive snapshots, it is easy to know which apps were removed, and their accurate removal date.

We have collected the daily removed apps in iOS app store for 1.5 year, *i.e.*, from 1st January, 2019 to 30th April, 2020. To the best of our knowledge, this is the only available and largest dataset on removed iOS apps in our research community. Overall, there are 1,129,615 app removal records, which correspond to 1,033,488 different mobile apps[3]. To enable further analysis, we also collect the detailed information of the removed apps. As shown in Table 1, the information can be classified into the following three categories:

1) *App Objective Information.* App objective information represents the data that created/defined by app developer/app market, e.g., app name, app ID, developer name, and app price, etc. We have maintained a list of 10 objective information for each removed app.

2) *App Subjective Quality.* This kind of data is extracted from the information provided by mobile users. App markets provide

---

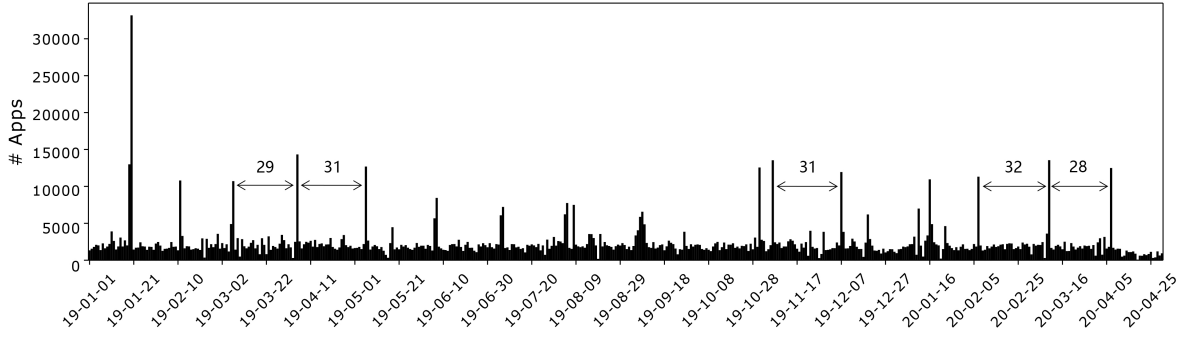[3]Note that one app can be removed from the market more than once.

Figure 1: The number of daily removed apps in iOS app store (from January 1st, 2019 to April 30th, 2020).
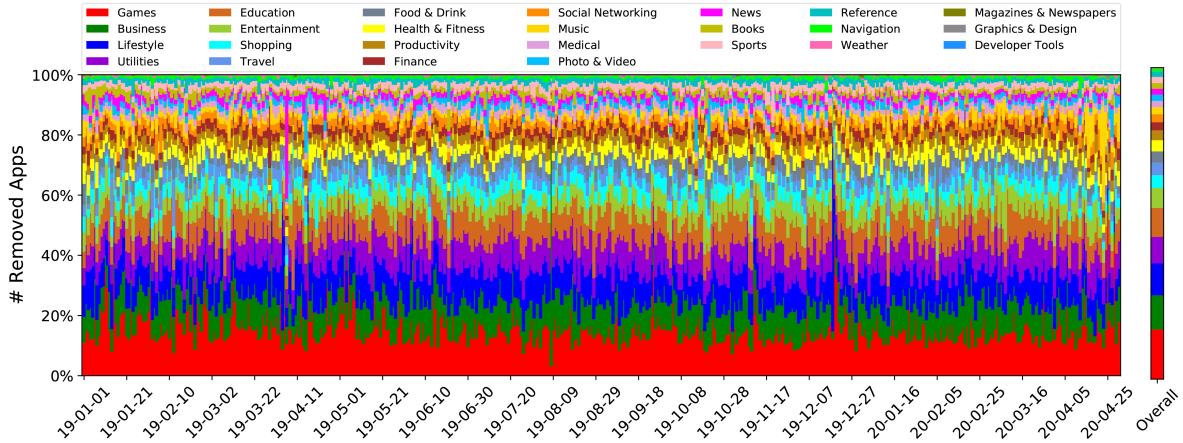

Figure 2: The proportion of daily removed apps across the 25 categories.

multiple ways for users to leave feedback, which eventually are summarized to app ratings and detailed app reviews.

3) *App Popularity Information.* App popularity information can be reflected as its ranking in the market, and its app search optimization (ASO) capability. ASO [24] is the process of improving the visibility of a mobile app. The key idea is that, by optimizing the keywords (e.g., popular searching words) in their app metadata (e.g., name and description), their apps would appear popular (highly ranked) in the search results. App markets including Google Play and iOS app store asked the developers to avoid user testimonials, excessive details, misleading references to other apps and repetitive, excessive or irrelevant keywords. This kind of information was obtained from the cooperated app intelligence company by daily monitoring the search results of massive keywords in iOS app store. It will be used in Section 5 and Section 6 for characterizing removed apps with ranking fraud.

4 OVERVIEW OF REMOVED APPS

In this section, we seek to investigate the practical practices of iOS app market in its app maintenance behaviors. To be specific, we present the overall landscape of removed apps in iOS app store, including the number of daily removed apps across the span of 1.5 years, the life-cycle of the removed apps, their correlation with other app meta information, and the developers of them.

4.1 Overall trend of removed apps

4.1.1 Overall Statistics. Figure 1 presents the number of daily removed apps in iOS app store. On average, 2,324 apps are removed daily. It can be observed that the number of daily removed apps fluctuates periodically. For example, the peak on 20th January, 2019 indicates a large-scale app removal, with 33,173 apps removed. Interestingly, we can observe a number of peaks during the evolution, while the median of time interval between the adjacent peaks is 28.5 days (examples shown in Figure 1). Thus, although apps can be removed daily from time to time, the large-scale app removal happens regularly every month in iOS app store. *We speculate that the app maintenance behaviors in iOS app store is cyclical (e.g., monthly), which may also be exploited by malicious developers*, i.e., using a number of spamming techniques [10, 43] to reach to unsuspecting users during the "silent time".

4.1.2 Category Distribution. Figure 2 presents the proportion of daily removed apps across each category. It can be observed that apps in Game, Business, Lifestyle, Utilities, and Education categories account for most (over 50%) of the daily removed apps. This is consistent with the overall distribution of removed apps in each category (see the rightmost column in Figure 2). However, we can observe that the proportion of removed apps in some categories are significantly increased on certain days. For example, there is a large-scale app removal in News category on April 6th, 2019, which

accounts for 36.48% of all removed apps that day (compared with roughly 1.8% on other days). Thus, *we speculate that iOS app store would focus on certain kinds of apps in a period of time and remove the inappropriate ones intensively, which will lead to a large-scale app removal in the corresponding categories.* Some anecdotal evidence on news media [15, 33] supports our speculation.

4.2 Popularity of Removed Apps

We next investigate the popularity of removed apps. Generally, we expect that the removed apps are low-quality apps with few users. Note that, unlike Google Play that provides the informative download statistics of the apps, in iOS app store, we can only measure the popularity of apps based on their app ranking and the number of user comments [3].

4.2.1 App Ranking. iOS app store provides the app ranking of top-1500 apps across different categories based on their real time popularity. Thus, during our dataset collection, we make efforts to harvest the daily app ranking information across the 25 categories. Without loss of generality, we define the `popular app` as the app that has ever been ranked in top-1500 in each category in this paper.

Figure 3 presents the proportion of removed popular apps in each category. Overall, 51,704 of the removed apps (5%) has been ranked in top-1500 across categories, while the remaining 95% of removed apps have never got into the app ranking list during their lifecycle. *This observation is basically in line with our expectation that most of the removed apps are non-popular apps.* However, subtle differences arise when looking in detail on the per-category basis. For the removed apps, over 10% of them are popular apps in categories including `Finance`, `Books`, `Sports`, `Reference`, `Weather`, and `Magazines & Newspapers`. One extreme category is `Weather`, and over 40% of the removed apps have ever ranked in top-1500. In contrast, the removed popular apps only take up roughly 1.8% to 2.3% for categories including `Business`, `Education` and `Games`. There are two major reasons to explain this scenario. On one hand, app categories with higher percentage of removed popular apps are generally small categories without many apps. On the other hand, our manually investigation suggests that some policy-violation apps (e.g., illegal gambling apps) usually hide themselves in some specific categories including `Sports`, `Weather` and `Reference` (see Section 5), which makes the proportion of removed popular apps higher in these categories.

Furthermore, we examine the distribution of free and paid removed popular apps in these categories. It is no doubt that free apps take up the majority (75.7%) of removed popular apps. Nevertheless, 12,541 of the removed popular apps are paid apps that have been downloaded many times. For example, some paid gaming apps (e.g., Supreme Glory - Eternal Blood, Sword in the Air, and Sword Dynasty) are quite popular, that even rank top-10 before removal. However, they were found to exaggerate their functional descriptions and post fake positive reviews to entice users to download them for revenue. *It suggests that the low-quality and potentially harmful apps have caused serious impact to unsuspecting users.*

4.2.2 The number of user ratings. The number of user ratings can be served as an implicit indicator of app popularity. In general, the more popular an app is, the more user ratings it will receive [3].
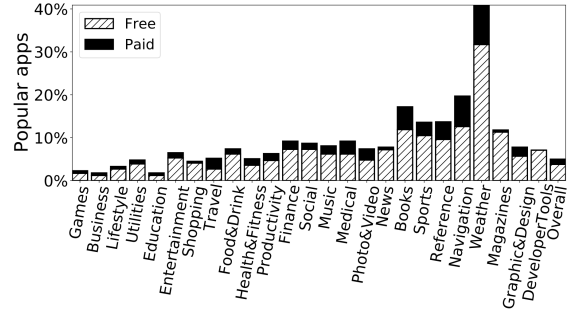

Figure 3: Percentage of removed popular apps in each category. The rightmost bar shows the overall proportion.
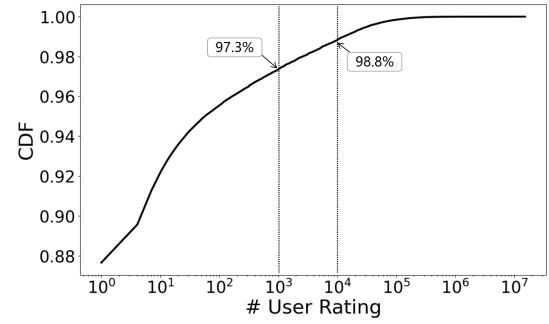

Figure 4: Distribution of the number of user rating.

Figure 4 shows the overall distribution of the number of user ratings received by the removed apps. *As expected, most of the removed apps (85.47%) have received no user ratings, which further indicates that most of them are low-quality and abandoned apps with almost no active users.* Nevertheless, some removed apps have gained a considerable number of user ratings, i.e., roughly 2.7% and 1.2% of them have received more than 1,000 and 10,000 user ratings, respectively. For example, `Fawn weather pass`, one described as weather forecasting but actually a gambling app, has received 119,759 user ratings by the time of its removal. *It suggests that these apps have a large number of active users at least during some period of their life-cycle.* We will further explore the underlying removal reasons in Section 5.

4.3 Developers of Removed Apps

As aforementioned, 420,955 developers have contributed to the over 1 million removed apps in our dataset. We next investigate these developers from two perspectives: 1) *are there any developers who tend to release policy-violation apps?* It can be perceived by analyzing the proportion of removed apps in all the apps they released; and 2) *are the removed apps dominated by a small number of aggressive developers?* It can be perceived by analyzing the developers with the most number of removed apps.

4.3.1 Proportion of removed apps per developer. We have collected the detailed app release history for all the developers we considered, through which we can get the proportion of removed apps for each
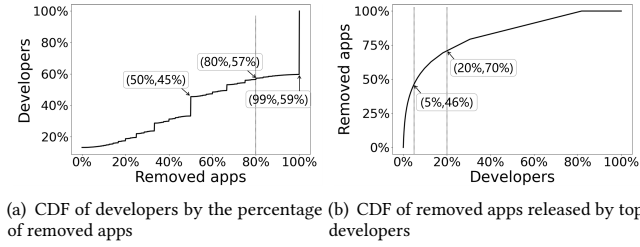
(a) CDF of developers by the percentage of removed apps

(b) CDF of removed apps released by top developers

Figure 5: The developers of removed apps.

developer. Figure 5(a) presents the overall distribution. Note that, over 230K developers (55%) in our dataset have released only one app and that app was removed, thus the proportion of removed apps for each of them is 100%. We eliminate them in Figure 5(a) to make the distribution more clear. It can be seen that more than half of developers have over 55% of their apps being removed from the app store. More seriously, 43% of developers have more than 80% of their apps being removed, 41% of developers even have all their apps removed. When further considering that 55% of developers have released only one app and that app was removed, 73.45% of developers in our dataset have all their apps removed. *It indicates that some developers tend to release policy-violation apps that would be removed by market. Thus, the app market and app users should pay special attention to the apps released by such kind of developers.*

4.3.2 The most aggressive developers. We rank developers by the number of their removed apps, and observe that the maximum number of removed apps per developer is 100, with 438 such developers in total. Figure 5(b) shows the distribution of the number of removed apps per developer sorted in descending order. Interestingly, it follows the typical *power law distribution.* Top 5% of developers have contributed to over 46% of the removed apps, and top 20% of developers have contributed to over 70% of the removed apps. *This observation suggests that majority of removed apps are dominated by a limited number of developers. We argue that the app market should design specific mechanism to label the trustworthiness of each developer, and raise alarm to the apps released by developers with aggressive behaviors.*
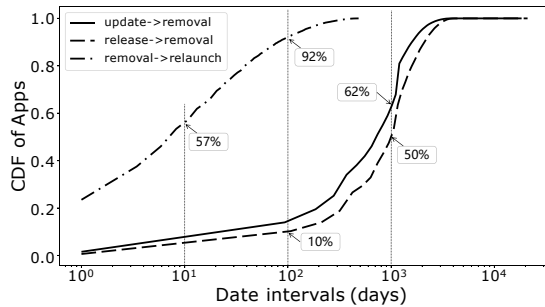

Figure 6: The distribution of time interval between update date and removal date, release date and removal date, removal date and relaunch date.

4.4 Life-cycle of Removed Apps

4.4.1 Life-cycle Analysis. We next analyze the life-cycle of removed apps, i.e., from their launch (release date), to their updates (update date), to their removal (offline date). Note that, the removed apps can be re-launched to the market later, thus we record their re-launch date as well for further analysis.

Update Date VS. Removal Date. Figure 6 presents the time interval between the *offline date* and the *update date* of each removed app. A long interval of time reflects that it is quite possible the app has been outdated or abandoned by the app developer. Roughly 67% of removed apps were not updated within 1 year before being removed, and even over 38% of them remain silent within 1,000 days before being removed. *It suggests that a large portion of the removed apps are outdated ones that may cannot function as intended in new version systems or no longer meet current app store guidelines.*

Release Date VS. Removal Date. We further pay attention to the time interval between *app release date* and *removal date*, which is the overall lifespan of that app. This metric can reveal the exposure time of a removed app, i.e., *how long is it visible to mobile users?* As shown in Figure 6, only 10% of the removed apps have a lifespan within 100 days and half of them last longer than 1,000 days. *This observation indicates that majority of the removed apps can sustain in the app store for long time.* The longer the time is, the more negative effects it will introduce.

Removal Date VS. Re-launch Date. As aforementioned, some of removed apps would be relaunched after a period of time. In our dataset, there are a total of 80,498 (7.78%) removed apps, which were then relaunched (once or multiple times). To measure how long it typically takes for a removed app to be relaunched after being removed, we calculated the distribution of the date interval between each removal and relaunch, as shown in Figure 6. The smaller the time interval, the sooner the app was re-launched. *This metric can be used to measure developer's responsiveness to the controversial apps.* Of all the relaunched apps, 8.7% are re-launched within one day as they were removed and about 57% are re-launched within 10 days after removal. This observation suggests that *most developers tend to fix the issues and relaunch their apps in a rapid manner.*

> **Answer to RQ1:** *During the span of 1.5 years we studied, over 1 million apps were removed from iOS app market, which is surprisingly higher than our expectation. App removal in iOS app store shows some cyclical pattern. Although most of the removed apps are non-popular apps, 5% of them are popular apps with high rank in app store and have gained a considerable number of user ratings. Some developers tend to release controversial apps that were removed, and the majority of removed apps are dominated by a limited number of developers. Most removed apps can sustain in the app store for long time.*

5 DEMYSTIFYING REMOVED POPULAR APPS

Our previous exploration suggests that, although most of the removed apps are low-quality apps (i.e., outdated and abandoned apps with almost no users), a number of the removed apps are quite popular (see Figure 3). Over 5.0% of them were ever ranked in top-1500 of their corresponding categories, and even 0.75% of them

were ever ranked in top-100 when competing with other apps in the same category. This observation motivates us to investigate the practical reasons behind the removal of such popular apps.

5.1 Labelling the Benchmark

Candidate Selection. To demystify the removed popular apps, we first rely on manual efforts to label some apps as the benchmark. To be specific, we select the removed apps whose most recent app rankings are within top-10 across each category before being removed from the market. In our dataset, there are 854 such apps in total. We believe they are definitely popular apps according to their app ranking. To understand their characteristics, we further select normal popular apps that ranked in top-10 as comparison, with 759 such apps in total. Note that, as aforementioned in Section 2, we have gathered all their available information.

Manually Labelling. To figure out the reasons why apps are removed from the app store, the first two authors try to label the reasons according to the review guidelines of iOS app store (see Section 2) independently. Both of the two authors have substantial mobile app analysis experience. To be specific, we rely on the detailed information we collected for labelling. For example, we will manually analyze all user comments before app removal (including user complaints and potential fake reviews), investigate the variation of app popularity including their covered keywords, and check the app update records (e.g., some apps have substantially revised their app description and even developer information to cover some keywords that irrelevant to their contents), etc. Each removed app is labelled by two authors independently. If there exist divergence between two authors, the app will be extracted for further discussion to resolve disagreements. Finally, we figure out the practical reasons of app removal for 722 apps (84.5%) out of the 854 apps. Note that for the remaining over 100 removed popular apps, we cannot find any clue from the information we collect to infer their removal reasons. It is quite possible that their app content has some controversial contents, or they were removed due to some political reasons (e.g., due to the censorship by some governments, iOS app store would also remove some popular apps.). We will further discuss it in Section 7. Nevertheless, our manual efforts can successfully infer the mis-behaviors leading to app removal.

Table 2 presents the practical reasons leading to app removal for the 722 apps. They were classified into five major types. Note that the removal reasons may not be mutually exclusive for a given app. Most apps are removed mainly due to the reasons including `Ranking Fraud` (67.6%), `Fake Description` (56.1%), and `Content Issue` (55.4%). It is worth noting that there exist a large amount of gambling apps in our benchmark. These apps usually contain illegal contents, provide fake description (to evade the app vetting process), and disrupt app ranking using malicious ASO techniques (i.e., covering a large number of sensitive or irrelevant keywords).

Note that, during our manually labelling, we found almost all the mis-behaviors in these removed popular apps can be reflected in *app review*, *app description*, and *ASO keywords*. That is also why we can manually label their removal reasons as aforementioned. For example, for the `Ranking Fraud` behavior, malicious actors always manipulate the app description and fake app reviews in an irregular manner to boost app ranking. Moreover, the ASO keywords it covered usually have quite different patterns compared with other normal apps in its category (see Figure 8(a)). For the `Fake Description` behavior, the semantic information of its description is usually inconsistent with real users' comments (complaints) in app reviews. For the apps with illegal content (`Content Issue`), they usually cover the corresponding illegal/sensitive keywords (e.g., gambling related keywords) during their app search optimization, in order to spread to as many users as possible in a short time. For the remaining removal reasons, we found they are usually not isolated, while most of them would combine behaviors including `Ranking Fraud`, `Fake Description`, or `Content Issue`.

Therefore, we will next detail the behavior patterns of removed popular apps from app review (Section 5.2), ASO keywords (Section 5.3), and app description (Section 5.4), respectively. These characteristics would be used as indicators to facilitate our detection of these apps with mis-behaviors in Section 6.

5.2 Behavior Patterns of App Review

App stores enable users to give feedback in the form of app reviews, which provides an opportunity to proactively collect user complaints and promptly improve apps' user experience. However, it is difficult to guarantee the quality and credibility of app reviews. During our exploration, we observe that the removed popular apps usually aggressively manipulate app reviews to boost their app ranking. This behavior is explicitly forbidden by iOS app store. Next, we characterize the behavior patterns of these removed apps from three perspectives, including the *the number of duplicated reviews* and *the proportion of 5-star reviews* which are highly related to app ranking fraud, and the *number of abnormal reviewers (users)* who posted the suspicious reviews. The statistics of the app reviews we collected for the labelled 854 removed popular apps and the 759 normal popular apps is shown in Table 3. Note that, we only collect their most recent 1 month reviews. Overall, there are over 16 million reviews posted by over 8 million unique app users in total. Obviously, the removed popular apps have much more user reviews (18,355 reviews on average, and 9,538 unique reviewers on average) than the normal apps (618 reviews on average and 575 unique reviewers on average).

(a) Duplicated reviews   (b) 5 star reviews
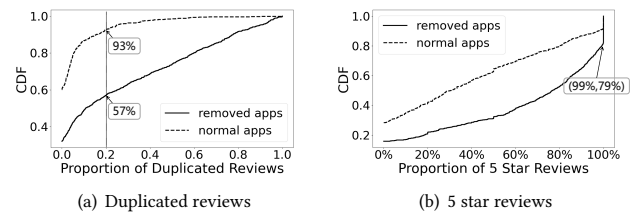
Figure 7: The distribution of duplicated reviews and high score reviews in removed apps and normal apps

5.2.1 Duplicated Reviews. Previous work [20, 21] suggested that the duplication of app reviews can be a simple way to reflect the credibility of reviews. If an app has a higher percentage of duplicated reviews, there is a higher possibility that the app has performed ranking fraud. In this work, we simply label a review is

Table 2: The Manually Labelled Removal Reasons of Removed "Popular" Apps.

Removal Reason	# Apps	Description
Ranking Fraud	488	Apps that manipulate its metadata (including reviews) to boost its app ranking
Fake Description	405	App description is inconsistent with app functions/behaviors
Content Issue	400	Apps that contain illegal content
In-app Purchase Fraud	119	Apps that trick users into purchasing premium services
General Quality Issue	105	Apps that have bugs, poor user experiences, crash issues, etc.
Overall	722	-

Table 3: The number of app reviews collected.

	Removed Apps	Normal Apps
# Reviews	15,675,218	469,041
# Reviewers (Users)	8,145,665	436,655

duplicated as long as there is another review posted within the app with identical content. Figure 7(a) presents the proportion of duplicated reviews in both the removed popular apps and normal popular apps. Obviously, the percentage of duplicated reviews in removed apps is much higher than that in normal apps. Over 43% of removed apps have more than 20% of duplicated reviews, compared to only 7% in normal apps. Over 8.5% of the removed apps have more than 80% of duplicated app reviews (VS. 0.4% in normal apps). This result suggests that *it is highly suspicious that there exist app ranking fraud in the reviews of the removed popular apps, and the proportion of duplicated reviews can be served as an indicator.*

5.2.2 5-star Reviews. In addition to text feedback, users can also review the app by giving a star rating, which is a number between 1 (bad) and 5 (good). We then characterize the proportion of high score reviews, i.e., 5-star reviews, of removed popular apps. Figure 7(b) presents the distribution of 5-star reviews for both removed popular apps and normal popular apps. We can observe a higher proportion of 5 star reviews in removed apps than in normal apps. Surprisingly, there are over 20% of removed apps whose proportion of 5 star reviews is close to 100%. Considering that the removed apps are the ones which in general have some issues, the proportion of 5 star reviews for these apps is expected to be lower. By manually analyzing a large number of such reviews, we find that they are either irrelevant to app content (like advertisement) or duplicated reviews we identified. It further indicates that *the app reviews in the removed apps are highly suspicious to be manipulated.*

5.2.3 Abnormal Users. The prior exploration suggests the high proportion of suspicious reviews. We next seek to explore the presence of abnormal users. Since different users are less likely to give the same review, users who give the same reviews are potentially to be abnormal users. In our analysis, reviews that have more than M words and appear N times in the dataset are considered as *abnormal reviews*. Users who have given *abnormal reviews* are considered as *abnormal users*. To filter potential false positives, we flag abnormal reviews and abnormal users under two different conditions: 1) M equals to 5 and N equals to 10; and 2) M equals to 10 and N equals to 20. Under the first condition, there are 15,350 abnormal users in the normal apps accounting for 3.58% of all users who have commented the normal apps. In contrast, there are 5,099,175 abnormal users

in the removed apps, which accounts for 62.6% of all users who have commented the removed apps. In the second condition, there are 15,187 abnormal users in normal apps (3.54%), compared to 3,958,413 abnormal users in the removed apps (48.6%). The results suggest that *a large portion of users in the removed popular apps are suspicious to be abnormal users, while the proportion of potential abnormal users is quite low in normal apps.*

5.3 Behavior Patterns of ASO Keywords

As aforementioned, ASO is the process of improving the visibility of the mobile app. The goal is to make the app highly ranked in the search results of some keywords. iOS app store disallows developers to use excessive and misleading keywords. Our prior manual investigation suggests that the ASO keywords have been extensively manipulated by app developers.

Motivating Example. We first start by giving a motivating example of the evolution of ASO keywords for a normal popular app and a removed popular app, respectively. Figure 8(a) presents the variation of the number of daily keywords between 17th September, 2019 and 30th September, 2019. `Taste Status` is a gambling app which is removed from the app store on 1st October, 2019, while `iWeekly` is a normal popular app which belongs to `Magazines & Newspapers` category. It can be observed that the number of keywords of Taste Status increased sharply during the period. The number of keywords was less than 10 at the very beginning. However, it increased to more than 7,500 within a single day. By manually analyzing the newly added keywords, we found most of them are irrelevant to its functionalities. In contrast, the number of keywords of `iWeekly` kept stable during the time period. It varies between 4,500 and 5,200. This example indicates that the characteristics of ASO keywords are quite different between normal apps and removed apps.

The Standard Deviation of ASO Keywords. Our initial exploration suggests the the variation of the numbers of keywords of normal apps and removed apps are quite distinct. Thus, we calculate the variation of the number of keywords of each app. Interestingly, there are 300 out of 722 (41.6%) removed apps whose ASO keywords number increased by 1,000 within the week before being removed, while there are only 26 out of 759 (3.4%) normal apps whose keywords number increase by 1,000. This indicates that the number of keywords of removed apps is more likely to increase sharply, while that of normal apps tends to keep stable. Figure 8(b) presents the distribution of standard deviation of the number of ASO keywords. It can be observed that the standard deviation of the number of ASO keywords in removed apps are greater than

Table 4: Keywords Coverage by Description of Removed Apps and Normal Apps

	Removed Apps	Normal Apps
Keyword Coverage (avg)	11.56	54.65
App # (Coverage Rate = 0)	97 (11.36%)	9 (1.05%)

that of normal apps. It further supports that the number of ASO keywords in removed apps in general changes rapidly.

(a) Motivating Example (b) Standard Deviation
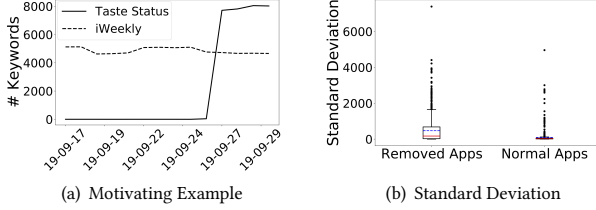
Figure 8: Behavior patterns of ASO keywords.

5.4 Behavior Patterns of App Description

App Description VS. ASO Keywords. Since there are sufficient keywords for a single app, we are interested in the contents of keywords. It is expected that the contents of keywords are relevant to the app itself. Since the description introduces the main function of an app, the keywords should be covered by the descriptions. Therefore, we investigate the number of keywords covered by the description of apps. We first filter the apps whose average number of keywords is too small (i.e., the average number of keywords is smaller than 100). Then we calculate the average keyword coverage by description of removed apps and normal apps. Table 4 presents results. The average keyword coverage of removed apps is 11.56, while that of normal apps is 54.65. Over 11% of removed apps have an average keyword coverage rate equal to 0, compared to only 1.05% for normal apps. *This observation suggests that the removed apps may cover many irrelevant keywords in the process of their app search optimization.*

> **Answer to RQ2:** *The removed "popular" apps are mainly due to reasons including ranking fraud, fake description, content issues, in-app purchase fraud, and general quality issues. Interestingly, by analyzing the meta information only, we can figure out the removal reasons for most of the removed popular apps. Some strong indicators from the app meta information can be used to differentiate removed apps from normal apps.*

6 DETECTING REMOVED APPS

Our prior exploration suggests that the indicators extracted from the app meta information (app description, app review, and ASO keywords) can be used to differentiate the apps that should be removed from normal apps. It motivates us to design an automated approach to flag suspicious apps that should be removed. Therefore, in this section, we seek to explore whether we can identify the controversial apps based on meta information only without accessing their app binaries (see Section 6.1). More importantly, as

Table 5: Features used for App Removal Prediction.

Feature Name	Description
Review Count	The number of daily reviews.
Review Standard Deviation	The standard deviation of daily review number.
Rating Percentage	The proportion of reviews of different ratings.
Duplicate Reviews Percentage	The proportion of duplicated reviews.
Abnormal Users	The number of users posting duplicated reviews.
ASO Keyword Count	The number of daily keywords.
ASO Keyword Standard Deviation	The standard deviation of daily keyword number.
# Keyword-Description Coverage	The number of keywords covered by the description
% Keyword-Description Coverage	The percentage of keywords covered by the description.

Table 6: Results of Prediction of App Removal.

	AUC	Precision	Recall	F1	Accuracy	FPR
LR	0.9028	0.8624	0.7778	0.8058	0.8180	0.1370
SVM	0.9021	0.8909	0.7600	0.8056	0.8236	0.1054
KNN	0.8703	0.8853	0.6725	0.7482	0.7874	0.0844
RF	0.9000	0.8661	0.7918	0.8168	0.8328	0.1278
GBDT	0.9116	0.8589	0.8214	0.8326	0.8397	0.1358

it may take rather long time for the app store to remove potential harmful apps, which may cause negative impacts on users, thus we are wondering if we can further flag the policy-violation apps in advance (see Section 6.2).

6.1 Predicting the Removed Apps

We first perform a binary classification on the labelled apps in Section 5, to predict the apps that should be removed by combining all the indicators we identified from app meta information.

Feature Selection. Table 5 shows the features used in the prediction. According to the behavior patterns we summarized in Section 5, we apply three sets of features. The first set of features represents characteristics of app reviews. These include the variation of reviews (number of daily reviews, standard deviation of the number of daily reviews), and the ratings of reviews (the proportion of reviews with different ratings). We also extract features from the suspicious reviews, including the proportion of duplicated reviews, and the number of users who have posted duplicated reviews. The second set of features is extracted from ASO keywords, including the number of daily keywords and standard deviation of the number of daily keywords. The third set of features is extracted from app description, including the number and the proportion of ASO keywords covered by app description.

Machine learning models. We select five representative models for the classification tasks, which are Logistic Regression, Support Vector Machine (a.k.a, SVM), K-Nearest Neighbors (a.k.a, KNN), Random Forest, and Gradient Boosting Decision Tree (a.k.a, GBDT). The implementation is based on scikit-learn[4] and LightGBM[5].

Evaluation Metrics. We use the widely used metrics including area under curve (a.k.a, AUC) score, precision, recall, F1 score, accuracy, and false positive rate to evaluate the performance of our model. A larger AUC score, precision, recall, F1 score, and accuracy indicates a better model. A smaller false positive rate indicates a better model. We perform a 10-fold cross validation and report the average performance on each of the metric, respectively.

---

[4]https://scikit-learn.org/stable/
[5]https://lightgbm.readthedocs.io/en/latest/

Results. Table 6 presents overall result. GBDT achieves the best performance in AUC score, recall, F1 score, and accuracy, while SVM achieves the best performance in precision. KNN achieves the best performance in false positive rate. The best performances for AUC score, precision, recall, F1 score, accuracy and false positive rate are 0.9116, 0.8909, 0.8214, 0.8326, 0.8397, and 0.0844, respectively. The results suggest that it is feasible to predict whether an app would be removed from the market based on its meta information. We next investigate the feature importance in the classification. *The number of abnormal users* is the most important feature while predicting the removal of an app. As indicated in Section 5.2.3, the proportion of abnormal users is much higher in removed apps. Besides, *Keyword Coverage Percentage*, *Keyword Standard Deviation*, and *Review Standard Deviation* are also important while predicting removal apps. This is in line with our findings in Section 5.

6.2 Predicting the Removed Apps in Advance

To investigate whether we can flag the policy-violation apps in advance, we conduct a series of attempts, i.e., predicting whether an app will be removed from one day in advance to six days in advance, and predicting whether an app will be removed once its mis-behaviors appear in app metadata.

6.2.1 Predicting the removed apps in advance. We first start by predicting whether an app will be removed from one day in advance to six days in advance. Therefore, we utilize the features extracted from a shorter time period. For example, when we predict whether an app will be removed a day in advance, we will exclude the information in the last day and extract the features from the remaining information. *Our expectation is that the policy-violation apps may release some initial signals that can be caught by our model.*
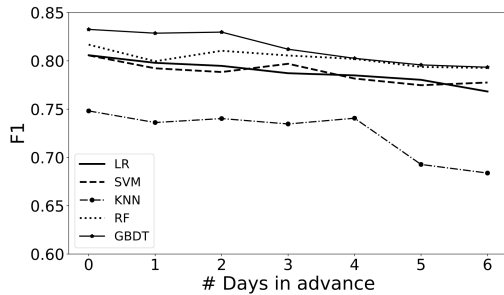

Figure 9: Predicting removed apps in advance.

Experiment Results. Figure 9 presents the overall result (F1 score) of app removal prediction from 0 day to 6 days. Obviously, the overall performance drops slightly along with the increment of time in advance for all models. Nevertheless, we observe the prediction result is quite good. For example, even we only rely on the information obtained prior to 6 days, the best model (GBDT) can achieve an F1 score of 79.3%. *Therefore, this result suggests that it is feasible to advance the process to identify and remove the suspicious apps from the market.*

6.2.2 Predicting the removed apps once the mis-behaviors appear. Since the policy-violating behaviors may appear at any time, we next evaluate whether we can detect such apps as early as possible when their inappropriate behaviors appear in their metadata. To figure out when the mis-behaviors first appear in app metadata, we rely on manual efforts to label the ground truth dataset. To be specific, two authors try to label the date when mis-behaviors first appear according to reviews and keywords. Each removed app is labelled by two authors independently. If there exists divergence between two authors, the app will be extracted for further discussion to resolve disagreements. Finally, we figure out the exact date when the mis-behaviors appear for 337 removed apps. Based on this dataset, we further apply our model to these removed apps to evaluate the "time window" of our approach to detecting them.

Experiment Results. On the first day when the mis-behaviors appear, our models can achieve an F1 score of 95.5%. When the mis-behaviors could be observed for two days, our model can achieve an F1 score of over 97.0%. *This result suggests that we can pinpoint the policy-violation behaviors timely once their inappropriate behaviors appear in their metadata.*

> **Answer to RQ3:** *Even without accessing to app binaries, we can identify the removed apps with good performance (F1=83%) based on the indicators extracted from app meta information. Moreover, even with less information, we are able to accurately flag the removed apps in advance.*

7 DISCUSSION

In this section, we discussed the findings across platforms (i.e., Android and iOS), the insights and implications for practice, and the limitations that can potentially affect the results of our study.

7.1 iOS App Store VS. Google Play

Our study is focused on iOS app store, one of the universally acknowledged most secure app stores, due to its locked-down ecosystem. Our study reveals a number of issues in the ecosystem, which we believe to exist in Android app markets like Google Play as well. Similar spamming techniques abused by the removed apps can be directly adopted to Google Play without any technical barriers. Moreover, our study in this paper does not rely on app bytecode, thus our investigation methods can be transferred to analyzing Google Play apps easily. The only concern is to harvest the comprehensive dataset of apps in Google Play. Our investigation is built atop of the unique app removal dataset that was collected extensively by monitoring the iOS app market daily across the span of 1.5 years. To the best of our knowledge, there is no such kind of datasets available for Android app markets yet.

7.2 Implications

Our observations are of key importance to the stakeholders in the mobile app ecosystem. *First, it shows the ineffectiveness of market regulation of iOS app store.* A number of apps with mis-behaviors were found in the market and then were removed, which usually takes a rather long time. This, however, can be improved through proactively monitoring the whole market using techniques like the one proposed in this paper. Our observation shows that some developers have the tendency to release policy-violation apps, while

the market should pay special attention to apps released by these developers. Furthermore, we observe that the app ranking and app searching recommendation mechanisms in iOS app store are seriously abused by spam apps, while indicates the emergency for detecting and regulating this kind of behaviors. *Second, app developers should be aware of the iOS app store guidelines to prevent their apps from being removed from the market.* Our observation suggests that, over 30% of removed apps were re-launched to the market after a specific time. Although most of them seek to fix issues in a rapid manner, it would greatly impact the reputation and the popularity of the app. *Third, even the "popular" apps are not trustworthy, as they can abuse a number of ASO techniques to disrupt the app market.* Thus, mobile users should pay special attention to the suspicious high ranking apps.

7.3 Limitations

Our study, however, carries several limitations. First, in this paper, we did not touch the bytecode of iOS apps, while some removal reasons can only be analyzed via the mobile apps. Nevertheless, we show that most of the removed popular apps can be identified proactively with features extracted from app metadata only. In the future, we can combine static or dynamic analysis of app binaries to get a deeper understanding of the removed apps. Second, this paper relies on some manual efforts to label the removal reasons. Although we strictly follow the iOS app store guidelines to understand their removal reasons, it is quite possible that our labelling reasons are inconsistent with their actual reasons. Third, as aforementioned, app removal can be perform by both app developers and the app market. However, it is almost impossible for us to to infer whether the app was removed by iOS app market.

8 RELATED WORK

8.1 Removed App Analysis

This is the first comprehensive study of iOS app removal practice at scale, longitudinally, and across various dimensions. The most related work is a preliminary study of removed apps in Google Play [38]. However, their work is coarse-grained as they only created two snapshots of Google Play in 2015 and 2017. They cannot infer the accurate app removal time, and they cannot measure the overall landscape of removed apps as we did including the daily trend, app popularity, and app life-cycle, etc. More importantly, we have collected the most comprehensive dataset including app reviews and ASO keywords, which are not covered by their work. Based on such information, we can analyze the behaviors patterns of the removed apps and design an automated approach to detect them, which were not touched by existing studies.

8.2 Mobile App Analysis

Mobile app analysis has been widely explored from various aspects, such as malware detection [18, 41, 45], app privacy and security [2, 4, 11, 13, 14, 19, 27, 28, 35, 36], app usage [8, 26, 44], and so on. These efforts focus on detecting and understanding the characteristics of malware, revealing the privacy and security problems of mobile apps, and characteristic of the usage pattern of different users under different situations. However, most of the existing studies focused on Android app analysis, while only a few of them focused iOS apps [2, 7, 9, 11, 14, 25, 27]. For example, iCryptoTracer [27] was proposed to detect misuse of cryptography functions in iOS apps based on static and dynamic analyses. Chen et al. [11] proposed to detect potentially-harmful libraries in iOS apps based on the observations that many iOS libraries have their Android versions that can potentially be used to understand their behaviors and relations. PiOS [14] was proposed to detect privacy leakage for iOS apps by analyzing data flows in Mach-0 binaries. Existing methods which aim to identify malware and potential harmful apps are mostly based on code analysis, which can be served as a complementary of this work to understand the deeper reasons of app removal.

8.3 App Ecosystem Analysis

There are plenty of efforts focusing on analyzing mobile app ecosystem through the perspective of app marketplaces [12, 29, 30, 37, 39, 40]. For example, Petsas et al. [30] analyzed the mobile app ecosystem through four popular third-party Android app marketplaces. They found out that app downloads follow a power-law distribution. Wang et al. [37, 40] analyzed the evolution of mobile app ecosystems through Google Play and Chinese Android app markets. They found that although the overall mobile app ecosystem shows promising progress, there still exists a number of unsolved issues for all mobile app marketplaces. The condition in Chinese app marketplaces is even worse. Chen et al. [12] proposed AR-Miner, which aims to extract informative informative reviews from app marketplaces. These efforts mainly focus on analyzing the overall patterns of apps and developers in different app marketplaces. However, these studies are limited to investigate the characteristics of removed apps like this paper, as it usually needs a continuous effort to monitor the market.

9 CONCLUSION

In this paper, we have conducted a large-scale and longitudinal study to understand the removed apps in iOS app store. Specifically, our analysis covers over 1 million removed apps obtained by monitoring iOS app market daily across the span of 1.5 years. Overall, our analysis suggests that the number of removed apps is surprisingly high, and the removal behaviors show cyclical patterns in iOS app store. We further pay special attention to the removed popular apps and design an automated approach to predict whether an app would be removed. Experiment results suggest that we can accurately predict the removal of apps, even we are able to flag them several days in advance. We believe our research efforts can positively contribute to the mobile app ecosystem and promote best operational practices across app markets.

REFERENCES

[1] 9to5mac. 2017. Hundreds of apps infected by fake Xcode tools, Apple removing known malicious software from App Store. https://9to5mac.com/2015/09/20/xcode-ghost-app-store-malware-malicious-app/

[2] Yuvraj Agarwal and Malcolm Hall. 2013. ProtectMyPrivacy: Detecting and Mitigating Privacy Leaks on IOS Devices Using Crowdsourcing. In *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services* (Taipei, Taiwan) *(MobiSys '13)*. Association for Computing Machinery, New York, NY, USA, 97–110.

[3] M. Ali, M. E. Joorabchi, and A. Mesbah. 2017. Same App, Different App Stores: A Comparative Study. In *2017 IEEE/ACM 4th International Conference on Mobile Software Engineering and Systems (MOBILESoft)*. 79–90.

[4] Hazim Almuhimedi, Florian Schaub, Norman Sadeh, Idris Adjerid, Alessandro Acquisti, Joshua Gluck, Lorrie Faith Cranor, and Yuvraj Agarwal. 2015. Your location has been shared 5,398 times! A field study on mobile app privacy nudging. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 787–796.

[5] BankInfoSecurity. 2020. Apple Battles App Store Malware Outbreak. https://www.bankinfosecurity.com/apple-battles-app-store-malware-outbreak-a-8538

[6] BRIAN BARRETT. 2020. How 18 Malware Apps Snuck Into Apple's App Store. https://www.wired.com/story/apple-app-store-malware-click-fraud/

[7] Zinaida Benenson, Freya Gassmann, and Lena Reinfelder. 2013. Android and iOS users' differences concerning security and privacy. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. 817–822.

[8] Matthias Böhmer, Brent Hecht, Johannes Schöning, Antonio Krüger, and Gernot Bauer. 2011. Falling asleep with Angry Birds, Facebook and Kindle: a large scale study on mobile application usage. In *Proceedings of the 13th international conference on Human computer interaction with mobile devices and services*. 47–56.

[9] Barry Brown, Moira McGregor, and Donald McMillan. 2014. 100 days of iPhone use: understanding the details of mobile device use. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*. 223–232.

[10] Hao Chen, Daojing He, Sencun Zhu, and Jingshun Yang. 2017. Toward Detecting Collusive Ranking Manipulation Attackers in Mobile App Markets. In *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security* (Abu Dhabi, United Arab Emirates) *(ASIA CCS '17)*. Association for Computing Machinery, New York, NY, USA, 58–70.

[11] Kai Chen, Xueqiang Wang, Yi Chen, Peng Wang, Yeonjoon Lee, XiaoFeng Wang, Bin Ma, Aohui Wang, Yingjun Zhang, and Wei Zou. 2016. Following devil's footprints: Cross-platform analysis of potentially harmful libraries on android and ios. In *2016 IEEE Symposium on Security and Privacy (SP)*. IEEE, 357–376.

[12] Ning Chen, Jialiu Lin, Steven CH Hoi, Xiaokui Xiao, and Boshen Zhang. 2014. AR-miner: mining informative reviews for developers from mobile app marketplace. In *Proceedings of the 36th international conference on software engineering*. 767–778.

[13] Feng Dong, Haoyu Wang, Li Li, Yao Guo, Tegawendé F Bissyandé, Tianming Liu, Guoai Xu, and Jacques Klein. 2018. Frauddroid: Automated ad fraud detection for android apps. In *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. 257–268.

[14] Manuel Egele, Christopher Kruegel, Engin Kirda, and Giovanni Vigna. 2011. PiOS: Detecting Privacy Leaks in iOS Applications. In *Proceedings of the Network and Distributed System Security Symposium, NDSS 2011, San Diego, California, USA, 6th February - 9th February 2011*. The Internet Society.

[15] Rachel England. 2020. Apple set to remove thousands of games from Chinese App Store in license crackdown. https://www.engadget.com/apple-will-remove-thousands-of-unlicensed-games-from-its-app-store-in-china-130534474.html

[16] L. García and R. J. Rodríguez. 2016. A Peek under the Hood of iOS Malware. In *2016 11th International Conference on Availability, Reliability and Security (ARES)*. 590–598.

[17] Google. 2020. Providing a safe and trusted experience for everyone. https://play.google.com/intl/en-GB/about/developer-content-policy/.

[18] Michael Grace, Yajin Zhou, Qiang Zhang, Shihong Zou, and Xuxian Jiang. 2012. Riskranker: scalable and accurate zero-day android malware detection. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*. 281–294.

[19] Yangyu Hu, Haoyu Wang, Ren He, Li Li, Gareth Tyson, Ignacio Castro, Yao Guo, Lei Wu, and Guoai Xu. 2020. Mobile app squatting. In *Proceedings of The Web Conference 2020*. 1727–1738.

[20] Yangyu Hu, Haoyu Wang, Li Li, Yao Guo, Guoai Xu, and Ren He. 2019. Want to earn a few extra bucks? a first look at money-making apps. In *2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER)*. IEEE, 332–343.

[21] Yangyu Hu, Haoyu Wang, Yajin Zhou, Yao Guo, Li Li, Bingxuan Luo, and Fangren Xu. 2019. Dating with scambots: Understanding the ecosystem of fraudulent dating applications. *IEEE Transactions on Dependable and Secure Computing* (2019).

[22] Apple Inc. 2020. App Review. https://developer.apple.com/app-store/review

[23] Apple Inc. 2020. App Store Review Guidelines. https://developer.apple.com/app-store/review/guidelines/

[24] Apple Inc. 2020. Optimizing for App Store Search. https://developer.apple.com/app-store/search/

[25] Hammad Khalid. 2013. On identifying user complaints of iOS apps. In *2013 35th international conference on software engineering (ICSE)*. IEEE, 1474–1476.

[26] Huoran Li, Xuan Lu, Xuanzhe Liu, Tao Xie, Kaigui Bian, Felix Xiaozhu Lin, Qiaozhu Mei, and Feng Feng. 2015. Characterizing smartphone usage patterns from millions of android users. In *Proceedings of the 2015 Internet Measurement Conference*. 459–472.

[27] Yong Li, Yuanyuan Zhang, Juanru Li, and Dawu Gu. 2014. iCryptoTracer: Dynamic Analysis on Misuse of Cryptography Functions in iOS Applications. In *Network and System Security - 8th International Conference, NSS 2014, Xi'an, China, October 15-17, 2014, Proceedings (Lecture Notes in Computer Science, Vol. 8792)*, Man Ho Au, Barbara Carminati, and C.-C. Jay Kuo (Eds.). 349–362.

[28] Tianming Liu, Haoyu Wang, Li Li, Xiapu Luo, Feng Dong, Yao Guo, Liu Wang, Tegawendé Bissyandé, and Jacques Klein. 2020. MadDroid: Characterizing and Detecting Devious Ad Contents for Android Apps. In *Proceedings of The Web Conference 2020*. 1715–1726.

[29] William Martin, Federica Sarro, Yue Jia, Yuanyuan Zhang, and Mark Harman. 2016. A survey of app store analysis for software engineering. *IEEE transactions on software engineering* 43, 9 (2016), 817–847.

[30] Thanasis Petsas, Antonis Papadogiannakis, Michalis Polychronakis, Evangelos P Markatos, and Thomas Karagiannis. 2013. Rise of the planet of the apps: A systematic study of the mobile app ecosystem. In *Proceedings of the 2013 conference on Internet measurement conference*. 277–290.

[31] TechCrunch. 2017. Apple goes after clones and spam on the App Store. https://techcrunch.com/2017/06/21/apple-goes-after-clones-and-spam-on-the-app-store/

[32] TechCrunch. 2018. App Store shrank for first time in 2017 thanks to crackdowns on spam, clones and more. https://techcrunch.com/2018/04/04/app-store-shrank-for-first-time-in-2017-thanks-to-crackdowns-on-spam-clones-and-more/

[33] James Vincent. 2019. Apple removes all vaping-related apps from the App Store. https://www.theverge.com/2019/11/15/20966362/apple-vaping-e-cigarette-apps-remove-app-store

[34] Paul Wagenseil. 2020. More than 1,200 iPhone apps infected with malware — what you need to know. https://www.tomsguide.com/news/iphone-apps-infected-malware

[35] Haoyu Wang, Yao Guo, Ziang Ma, and Xiangqun Chen. 2015. Wukong: A scalable and accurate two-phase approach to android app clone detection. In *Proceedings of the 2015 International Symposium on Software Testing and Analysis*. 71–82.

[36] Haoyu Wang, Jason Hong, and Yao Guo. 2015. Using text mining to infer the purpose of permission use in mobile apps. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 1107–1118.

[37] Haoyu Wang, Hao Li, and Yao Guo. 2019. Understanding the evolution of mobile app ecosystems: A longitudinal measurement study of google play. In *The World Wide Web Conference*. Association for Computing Machinery, New York, NY, USA, 1988–1999.

[38] Haoyu Wang, Hao Li, Li Li, Yao Guo, and Guoai Xu. 2018. Why are android apps removed from google play? a large-scale empirical study. In *2018 IEEE/ACM 15th International Conference on Mining Software Repositories (MSR)*. IEEE, 231–242.

[39] Haoyu Wang, Zhe Liu, Yao Guo, Xiangqun Chen, Miao Zhang, Guoai Xu, and Jason Hong. 2017. An explorative study of the mobile app ecosystem from app developers' perspective. In *Proceedings of the 26th International Conference on World Wide Web*. 163–172.

[40] Haoyu Wang, Zhe Liu, Jingyue Liang, Narseo Vallina-Rodriguez, Yao Guo, Li Li, Juan Tapiador, Jingcun Cao, and Guoai Xu. 2018. Beyond google play: A large-scale comparative study of chinese android app markets. In *Proceedings of the Internet Measurement Conference 2018*. Association for Computing Machinery, New York, NY, USA, 293–307.

[41] Haoyu Wang, Junjun Si, Hao Li, and Yao Guo. 2019. Rmvdroid: towards a reliable android malware dataset with app metadata. In *2019 IEEE/ACM 16th International Conference on Mining Software Repositories (MSR)*. IEEE, 404–408.

[42] Quan Wang, Beibei Li, and Param Vir Singh. 2018. Copycats vs. Original Mobile Apps: A Machine Learning Copycat-Detection Method and Empirical Analysis. *Info. Sys. Research* 29, 2 (June 2018), 273–291.

[43] Zhen Xie, Sencun Zhu, Qing Li, and Wenjing Wang. 2016. You Can Promote, but You Can't Hide: Large-Scale Abused App Detection in Mobile App Stores. In *Proceedings of the 32nd Annual Conference on Computer Security Applications* (Los Angeles, California, USA) *(ACSAC '16)*. Association for Computing Machinery, New York, NY, USA, 374–385.

[44] Sha Zhao, Julian Ramos, Jianrong Tao, Ziwen Jiang, Shijian Li, Zhaohui Wu, Gang Pan, and Anind K Dey. 2016. Discovering different kinds of smartphone users through their application usage behaviors. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 498–509.

[45] Yajin Zhou and Xuxian Jiang. 2012. Dissecting android malware: Characterization and evolution. In *2012 IEEE symposium on security and privacy*. IEEE, 95–109.