

A Descriptive Analysis of a Large-Scale Collection of App Management Activities *

Huoran Li
Peking University
lihuoran@pku.edu.cn

Wei Ai
University of Michigan
aiwei@umich.edu

Xuanzhe Liu
Peking University
liuxuanzhe@pku.edu.cn

Qiaozhu Mei
University of Michigan
qmei@umich.edu

Feng Feng
Wandoujia Lab
jackfeng@wandoujia.com

ABSTRACT

Smartphone users adopt an increasing number of mobile applications (a.k.a., apps) in the recent years. Investigating how people manage mobile apps in their everyday lives creates a unique opportunity to understand the behaviors and preferences of mobile users. Existing literature provides very limited understanding about app management activities, due to the lack of user behavioral data at scale. This paper analyzes a very large collection of app management log of the users of a leading Android app marketplace in China. The data set covers one month of detailed activities of how users download, update, and uninstall the apps on their smart devices, involving 8,306,181 anonymized users and 394,661 apps. We characterize how these users manage the apps on their devices and identify behavioral patterns that correlate with users' online ratings of the apps.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous

General Terms

Experimentation

Keywords

mobile computing, app management activities, user behavior analysis

*This work is supported by the High-Tech Research and Development Program of China under grant number 2013AA01A605 and the Natural Science Foundation of China under grant numbers 61370020 and 61222203. It is also partially supported by the National Science Foundation under the grant number IIS-1054199.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).
WWW 2015 Companion, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3473-0/15/05.
<http://dx.doi.org/10.1145/2740908.2742771>

1. INTRODUCTION

The rapid growth of smartphone apps has changed people's daily lives. To help the users selecting and exploring apps, most app marketplaces gather user ratings of apps, either in the format of like votes, numerical ratings, and/or free text comments. Much work has been done on analyzing these reviews [2]. However, among hundreds of millions of users who download and use these apps, only a small proportion have left a review [1]. Besides these biases, fake, paid off, and malicious reviews also commonly exist, compromising the credibility of the reported metrics [3].

We study a different signal, hypothesizing that users who like an app use and manage it differently from those who don't. If one can infer user preferences from behavioral patterns in these app management activities, one could provide an accurate and unbiased indicator of the quality and popularity of apps even if they are not rated.

In this paper we present the first empirical analysis on app management activities collected from eight million users. We present a descriptive analysis of how users' app management activities correlate with the popularity and ratings of these apps on different marketplaces. We present sequential patterns of user activities that are correlated with online ratings of apps, which can be potentially utilized to build predictive models for app ratings.

2. A DESCRIPTIVE ANALYSIS

The data is collected through a leading Android app marketplace in China, the Wandoujia.¹ Our data set covers time stamped sequences of 106,930,529 app management activities of 8,306,181 anonymized users on 394,661 apps in one month, as well as the online ratings of these apps. We collected three kinds of user activities: downloads, updates and uninstallations during the period from August 4th to September 2nd, 2013.

Figure 1 plots the percentages of user activities over 24 hours of a day. We aggregated downloading and updating activities in this plot and compared them with uninstallations. All timestamps were converted to the Beijing Time (UTC+8) as most of Wandoujia users are in China. As one may anticipate, app management activities start to increase sharply from 7 a.m. and stay high through out the day. Both downloading/updating and uninstallation activities fall around lunch and dinner.

¹<http://www.wandoujia.com/>

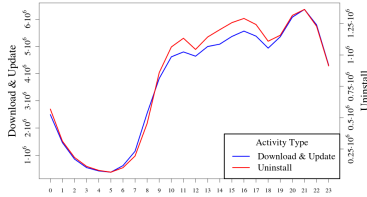


Figure 1: Activities peak in late afternoon and television time.

We hypothesize that users may have a routine schedule of housekeeping the apps on their mobile devices. To further investigate this, we plotted the distribution of the **time intervals** between any two consecutive activities of the same user (Figure 2). Most consecutive activities are conducted within less than an hour, which are likely in the same session. However, when the intervals are larger, there is a peak at every 24 hours. This suggests that a user does have a routine time period of a day for housekeeping - she may not do it every day, but when she does, it is likely to be closer to the same time of a day.

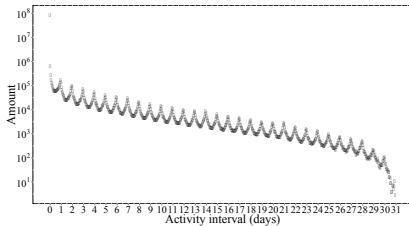


Figure 2: Intervals between activities peak at every 24 hours.

3. INFERRING APP QUALITY

An app is considered to be of high quality if it receives a higher average rating by its users. In practice, however, the online ratings suffer from data sparseness and biases.

We crawled user ratings of all apps from both Wandoujia and Google Play, which is the native Android market place. For the apps that are rated by the users on both marketplaces, we investigated the correlation of the number of ratings an app receives on the two sites, plotted in Figure 3. The number of ratings given to the same app at the two marketplaces are generally positively correlated. However, the noticeable vertical lines in the left and horizontal lines in bottom indicate considerable biases. These data points refer to the apps that have many ratings on one market but very few on the other. For example, the Facebook app has 25,169,686 ratings on Google Play, but has only 1,644 rating on Wandoujia.

We then investigated whether the same app receives similar ratings on the two marketplaces. Figure 4 presents a positive correlation between the average *scores* on Google Play and the *likerates* (e.g., number of positive ratings divided by number of all ratings) on Wandoujia for apps which receive at least five ratings on Wandoujia. It seems that the user ratings are overall coherent on the two marketplaces. However, we can also identify many different or even contradictory ratings for the same apps. The correlation is poor if we include apps with few ratings on Wandoujia.

We hypothesize that signals from app management activities may address the biases and sparseness of online rat-

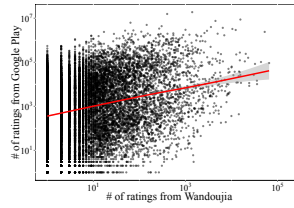


Figure 3: Numbers of user ratings are correlated in different marketplaces, but significant biases exist.

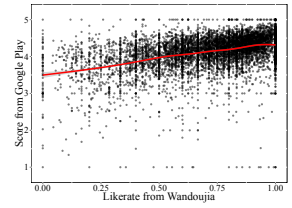


Figure 4: Average ratings at different marketplaces are correlated if there are abundant of ratings.

ings. Our goal is to identify weak signals that are good indicators of user preference. We found that “Uninstallation-Downloading (UD)” activities, i.e., a user uninstalls an app and later on re-installs it, may be a good indicator that the user likes the app (that’s why he installed it back). To verify this, in Figure 5 we plot the frequency of “UD” patterns among the activities of all users of an app and correlate it with the likerate of the app. Since not all apps have an “UD” activity sequence, we only plot those who have. The average number of “UD” sequences per user is in general positively correlated with the likerate of the app. This is promising, indicates that some user activity sequences may be good indicators of user preferences and app quality.

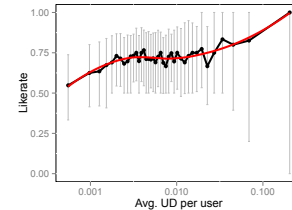


Figure 5: “Uninstallation-Downloading” sequences are positively correlated with Likerate.

4. DISCUSSION

It is encouraging to observe that certain sequential patterns of app management activities are good indicators of user preferences and app quality, which effectively supplement the biases and sparsity of online ratings. Next we will explore how to identify all such indicative patterns from large-scale behavioral data and how to construct a machine learning model that predicts app quality. Our study serves as a preliminary step of understanding individual and collective app using behaviors of smartphone users.

5. REFERENCES

- [1] S. Lim, P. Bentley, N. Kanakam, F. Ishikawa, and S. Honiden. Investigating country differences in mobile app user behavior and challenges for software engineering. *IEEE Transactions on Software Engineering*, 2014.
- [2] B. Liu. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1):1–167, May 2012.
- [3] M. Ott, C. Cardie, and J. Hancock. Estimating the prevalence of deception in online review communities. page 201. ACM Press, 2012.