

# Review on Examination of Mobile User Behavior Using Data Mining Techniques on App Store Data

Neha S.Pawar, Prof S.S.Dandge

**Abstract**— Mobile apps are software developed for use it on mobile phones and these apps are available in app stores. App stores are highly viable markets where app developers need to provide to a large number of users across multiple countries. There is large country differences present in mobile app user behavior. In proposed application one of the largest surveys will conduct to date of app users across the world, in order to examine the exact nature of those differences. The survey will examine user acceptance of the app store concept, needs, and basis for selecting or discarding an app. It will collect data from different countries. Proposing application will analyze user behavior for app download throughout multiple countries using data mining technique, so the results can be of many benefits for future app development. The results of survey that will release to the app developer and feedback received from the users that are very helpful. Anticipate that the new challenges identified in this study can guide software engineering researchers towards the development of tools and techniques to improve market-driven software engineering for mobile apps.

**Index Terms**—Mobile apps, app store, Data Mining, Cosine similarity

## 1 INTRODUCTION

The smart phones are the new trend in the world and almost everyone carry smart phones. As the smart phones arrived the applications starts flooding to the app store for multiple requirement. The current scenario is that there is millions of app available in app store and the number are still growing. Because of not having knowledge of actual application which is available in the store developers are developing the apps with similar behavior. This makes failure of the app on the store and less downloads leads to no profit of maximum app. Sometimes developers on one country might not be aware about the facts that some applications are still available in the market and is being used by the people of other country. The problem is increasing rapidly and must be addressed.

Developers are basically have very little margin on revenue, so the application must address major crowd. The revenue is basically depends on the number of downloads. The current study examined that more than 80% of the apps available in IOS app store have no download or less download. So it comes to the failure of the app. Despite this fact, thousands of applications are still being made and the development cost is being waste. Proposing application will analyze user behavior for app download throughout multiple countries using data mining technique, so the results can be of many benefits for future app development.

## 2 RELATED WORK

As the app store and Google play store is brought and grown most fast in market, not much research has been in this section. Still some of the efforts have been taken and elaborated their efforts in following section. The research work is basically categorized into those who mine app data, those who

mine activity log of devices and those who conduct survey.

On 16 Spetember 2012 Pagano and Maalej collected data on user reviews and ratings for the top 25 paid and free apps of only one country from all app categories in the Apple iOS App Store [1]. They used various statistical measures to examine how users provide feedback, and when they provide it, as well as study the content of the reviews. Their results showed that the user provided short reviews after new app releases, with a quickly decreasing frequency over time. In addition, user reviews contain various topics, like user experience, bug reports, and feature requests while using those apps. The worth and usefulness of user reviews vary extensively, from helpful advices and innovative ideas to offensive comments [1].

Harman et al. collected data from the Blackberry app store for information like app description, category, user ratings, price and the ranking of the app based on downloads [5]. They found that there is strong correlation between user ratings for the particular app and the ranking of an app, but there is no correlation between prices and number of downloads. Their study only based on priced apps, but extra work may be necessary for to examine free apps. [5].

Chen and Liu collected app information such as name of an app, app developer, app category, current ranking of an app, average rating for particular app, and number of ratings given to that app; they collected all these information from the Apple iOS App Store [6].

Most of the studies focused on collecting requirements for specific apps. For example, Henze et al. developed five game apps for the Android market and made it available for download and examined how the apps were used by the users [7]. Their most accepted applications collected data from 6,907 users. Their data showed that most of the users discarded the apps after a short time and they suggested that developers need to focus on quality of an app and providing incentives to users in order to encourage long-term use of an

- Neha S.Pawar is ME Student at PRMIT&R Badnera, Amravati.
- Prof S.S.Dandge is assistant professor at PRMIT&R Badnera, Amravati

app [7]. Hence et al. examined that most of their users were English-speaking from the United States, hence he was unable to examine the result over country wise [7].

McMillan et al. collected usage data of their iPhone app from 8,676 users over five months [8]. Their Data logging technique seem to be a costly to gather data from a large number of geologically dispersed users. However, the activity logs generated by their result were unable to provide detail knowledge of user behavior, and log analysis unsuccessful to disclose the users' needs and their behavior [8].

### 3 PROPOSED WORK

#### A. Motivation

The most recent technologies in the world are smart phones and the applications being developed in app store for end users to download. This motivates us to perform research in the sector and obtained a thorough knowledge of app store and the applications. We can develop an application which can mine the data of app store log and can generate the decisions for developers for future app development.

#### B. Problem statement

The major players in the app store market are apple app store and android play store. The stores are flooded with many applications. Almost 80% of applications have less or no download. Still the developers are developing many similar applications irrespective of the user behavior. This leads to the failure of app over app store.

#### C. System Architecture

Mainly there are four modules in the proposed application which are as follows:

##### Module 1: Dataset collection and data gathering

The very first thing is the dataset collection, the app store dataset need to be downloaded from app store and play store. This log will contain the information and reviews about each app downloaded from the store also the little information about the user. This can be available from freely available dataset.

##### Module 2: Survey & Data Cleaning

Have to build the survey which can be conduct and can contain free input fields for the user. The technique known as snowballing will be used in the survey. The data cleaning step will include cleaning of answers obtained from the survey.

##### Module 3: Data Analysis

Have to perform the relationship between different user attribute and based on that and have to calculate resemblance ratio in this. The existing method uses the correlation coefficient approach like RQ1 to RQ4 for resemblance match-

ing. In the Proposing application cosine similarity technique concept can use for resemblance matching.

#### Module 4: Analysis & Testing

The testing of the application and analysis of the results will perform in the last step. This step will also include the comparison of the system with the existing one.

#### D. Proposed System Design

Step1:-  
Data Gathering (Survey)

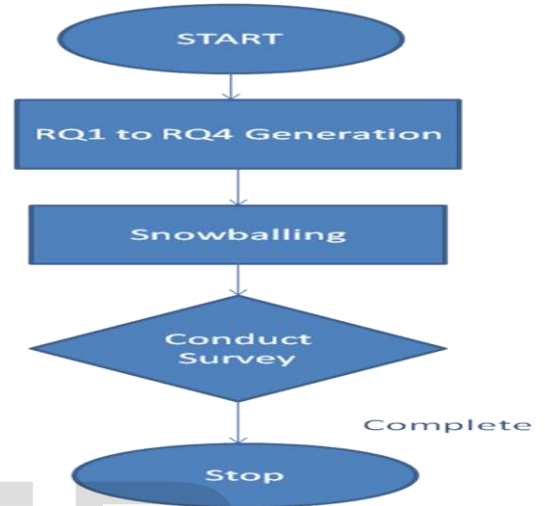


Fig 1.1 Overview of Data gathering

Step 2: Mixing

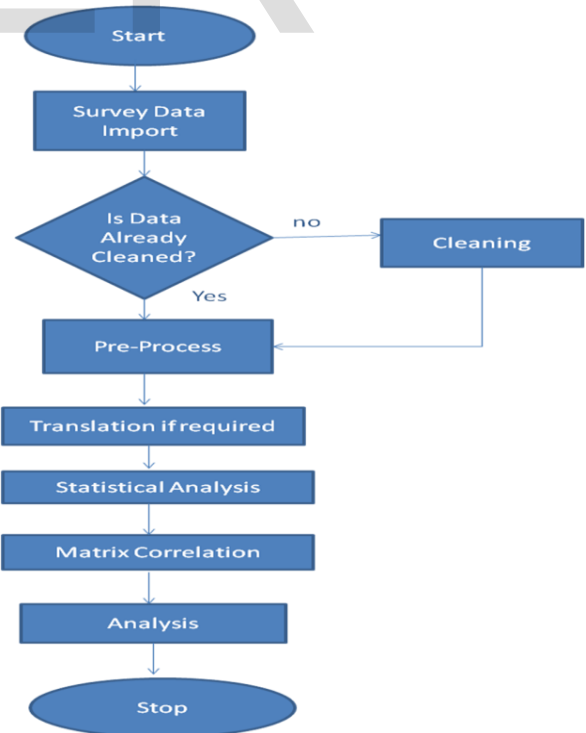


Fig 1.2 Overview of Mixing

Description of proposed Architecture:

Step 1: These steps involve the gathering of dataset. In this step will prepare the questionnaire and apply the snowballing technique to gather the answers from different people from different countries.

Step 2: The process of mining will apply in this step and analyze the data using different techniques. For first three set RQ1 to RQ3 we will apply the technique of cosine similarity to match the answers and its uniformness. In the next step will apply the same peterson's match matrix for processing RQ4.

E. Contribution:

Cosine Similarity:-

This cosine similarity metrics is frequently used when trying to determine similarity between two documents. Because there are more words that are common between two documents, it is useless to use the other methods of calculating similarities (that is the Euclidean Distance and the Pearson Correlation Coefficient discussed earlier). As a result, the probability that two documents do not share the majority is very high (as with the Tanimoto Coefficient) and does not generate a satisfactory metric for determining similarities.

In this paper, for the proposed application the technique of Cosine Similarity will apply to match the similarity of the documents for RQ1 to RQ3. These are the sets where answers will be in subjective

## 4 CONCLUSION

The smart phones are the current trend in the world and mostly every person carry smart phones. As the smart phones arrived the applications starts flooding to the app store for multiple requirement. App stores are highly competitive markets with a rapidly increasing number of apps, and developers need to supply to large number of users due to low margins per scale. In this proposing application one of the largest surveys will conduct to date of mobile app users across the world. Demonstrate that app user behavior differs signifi-

cantly across countries, a result that will show in other domains but never before in app-based software engineering, indicating that developers need to carefully consider the countries of their end users.

In this paper, proposing application will also examine user adoption of the app store concept, their app needs, and their basis for selecting or discarding an app. Through analysis of the survey results, Identify new challenges to market-driven software engineering related to covering requirements, element space, quality expectations, app store dependency, price sensitivity, and ecosystem effect, and their indication for software engineering research in terms of research directions and tool development.

Proposing application will release the results of survey to the app developer community and received feedback that the insights are very useful. Few developers have requested for other countries to be studied as they are building apps for those countries. Anticipate that the new challenges identified in this study can guide software engineering researchers towards the development of tools and techniques to improve market-driven software engineering for mobile apps.

## REFERENCES

- [1] D. Pagano and W. Maalej, "User feedback in the appstore: An empirical study," in Proc. 21st IEEE Int. Requirements Eng. Conf., 2013, pp. 125-134.
- [2] M. Bohmer, B. Hecht, J. Schoning, A. Kruger, and G. Bauer, "Falling asleep with Angry Birds, Facebook, Kindle: A large scale study on mobile application usage," in Proc. 13th Int. Conf. Human Comput. Interaction Mobile Devices Services, 2011, pp. 47-56
- [3] O. I. Franko and T. F. Tirrell, "Smartphone app use among medical providers in ACGME training programs," J. Med. Syst., vol. 36, pp. 3135-3139, 2012.
- [4] Y. Van Everdingen, J. Van Hillegersberg, and E. Waarts, "Enterprise resource Planning: ERP adoption by European midsize companies," Commun. ACM, vol. 43, pp. 27-31, 2000.
- [5] M. Harman, Y. Jia, and Y. Zhang, "App store mining and analysis: MSR for app stores," in Proc. 9th IEEE Work. Conf. Mining Softw. Repositories, 2012, pp.108-111.
- [6] M. Chen and X. Liu, "Predicting popularity of online distributed applications: iTunes app store case analysis," in Proc. iConf., 2011, pp. 661-663.
- [7] N. Henze, M. Pielot, B. Poppinga, T. Schinke, and S. Boll, "My app is an experiment: Experience from user studies in mobile app stores," Int. J. Mobile Human Comput. Interaction, vol. 3, pp. 71-91, 2011.
- [8] D. McMillan, A. Morrison, O. Brown, M. Hall, and M. Chalmers, "Further Into the wild: Running worldwide trials of mobile systems," Pervasive Comput. vol. 6030, pp. 210- 227, 2010.