

A Novel Approach of Investigating Deceptive Activities of Developer for Ranking apps

Priyanka.jali^{1*}, Nagavani Biradar²

^{1,2} Department of Computer science,
VTU belgaum, India

www.ijcseonline.org

Received: Feb/22/2016

Revised: Mar/06/2016

Accepted: Mar/14/2016

Published: Mar/31/2016

Abstract—Ranking fraud in the mobile App market refers to fraudulent or deceptive activities which have a purpose of bumping up the Apps in the popularity list. Indeed, it becomes more and more frequent for App developers to use shady means, such as inflating their Apps' sales or posting phony App ratings, to commit ranking fraud. While the importance of preventing ranking fraud has been widely recognized, there is limited understanding and research in this area. To this end, in this paper, we provide a holistic view of ranking fraud and propose a ranking fraud detection system for mobile Apps. Specifically, we first propose to accurately locate the ranking fraud by mining the active periods, namely leading sessions, of mobile Apps. Such leading sessions can be leveraged for detecting the local anomaly instead of global anomaly of App rankings. Furthermore, we investigate three types of evidences, i.e., ranking based evidences, rating based evidences and review based evidences, by modelling Apps' ranking, rating and review behaviours through statistical hypotheses tests. In addition, we propose an optimization based aggregation method to integrate all the evidences for fraud detection. Finally, we evaluate the proposed system with real-world App data collected from the iOS App Store for a long time period. In the experiments, we validate the effectiveness of the proposed system, and show the scalability of the detection algorithm as well as some regularity of ranking fraud activities.

Keywords—Mobile ranking, fraudulent mobile apps.

I. INTRODUCTION

Mobile app is one of the most used applications in mobile industry, where users download the apps from the server and they use it. The real usage and common features of the mobile app are very hard to know without actually using it. Here there are so many ways to deceive the users about the mobile app. Some of the app are not at all used by the people and yet they will have higher ranking than the most commonly used app. This they do by giving false positive review about the app, they also try to give higher ranking for their app so that users should download it and use it. Thus making their app more popular. To make the improvement of portable Apps, numerous developer of the app dispatched every day App updates, which exhibit the graph rankings of most prominent Apps. To be sure, the

App leader board is a standout amongst the most essential courses for advancing mobile Apps. A higher rank on the leader board more often than not prompts countless and million dollars in income. In this way, App designers have a tendency to investigate different routes, for example,

publicizing battles to advance their Apps keeping in mind the end goal to have their Apps positioned as high as could be expected under the circumstances in such App leader boards.

II Literature Survey

Literature survey is the most important step in software development process. Before improving the tools it is compulsory to decide the economy strength, time factor. Once the programmer's create the structure tools as programmer require a lot of external support, this type of support can be done by senior programmers, from websites or from books.

A. *A flexible generative model for preference aggregation*

AUTHORS: M. N. Volkovs and R. S. Zemel

Many areas of study, such as information retrieval, collaborative filtering, and social choice face the preference aggregation problem, in which multiple preferences over objects must be combined into a consensus ranking. Preferences over items can be expressed in a variety of forms, which makes the aggregation problem difficult. In this work we formulate a flexible probabilistic model over

Corresponding Author: Priyanka.j.priyanka.jali29@gmail.com
Dept. of CS, University of VTU belgaum, karanataka, India

pair wise comparisons that can accommodate all these forms. Inference in the model is very fast, making it applicable to problems with hundreds of thousands of preferences. Experiments on benchmark datasets demonstrate superior performance to existing methods

B. Getjar mobile application recommendations with very sparse datasets

AUTHORS: K. Shi and K. Ali

The Netflix competition of 2006 [2] has spurred significant activity in the commendations field, particularly in approaches using latent factor models. However, the near ubiquity of the Netflix and the similar MovieLens datasets may be narrowing the generality of lessons learned in this field. At GetJar, our goal is to make appealing recommendations of mobile applications (apps). For app usage, we observe a distribution that has higher kurtosis (heavier head and longer tail) than that for the aforementioned movie datasets. This happens primarily because of the large disparity in resources available to app developers and the low cost of app publication relative to movies.

M. Bellare, C. Namprempre, and G. Neven Security proofs for identity-based identification and signature schemes. This paper provides either security proofs or attacks for a large number of identity-based identification and signature schemes defined either explicitly or implicitly in existing literature. Here is a framework that helps to explain how the schemes are derived and also permit modular security analyses, which helps to understand, simplify, and unify previous work. We study the generic folklore construction that provides identity-based identification and signature schemes without random oracles.

C. Detecting spam web pages through content analysis

AUTHORS: A. Ntoulas, M. Najork, M. Manasse, and D. Fetterly

In this paper, we continue our investigations of "web spam": the injection of artificially-created pages into the web in order to influence the results from search engines, to drive traffic to certain pages for fun or profit. This paper considers some previously-undescribed techniques for automatically detecting spam pages, examines the effectiveness of these techniques in isolation and when using classification algorithms aggregated. When combined, our heuristics correctly identify 2,037 (86.2%) of the 2,364 spam pages (13.8%) in our judged collection of 17,168 pages, while misidentifying 526 spam and non-spam pages (3.1%).

III System Architectur

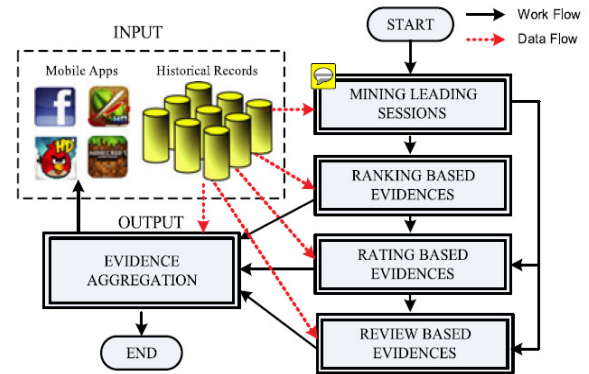


Figure1: Architecture

The above diagram shows the way the application works, here we first do the mining of the leading session, session based information is stored. Then the ranking based evidence is drawn, in ranking based evidence we will try to know weather the raking given to the particular mobile app is positive or negative. Then rating based evidence is drawn. Historical data is used to analyse the ranking of the mobile app. If number of download is less and more reviews people have given then we will assume that the given mobile app is having fraud ranking.

IV. Methodology

Mobile app fraud detection is very important things to be known, the way we do these things is by going though three different steps like ranking based evidence where we will see the ranking information of the mobile app, then in the second step we will see about rating based evidence, one will investigate here what is the ranking of the mobile app, which can be known by studying the rating details of the users. Third step is where review based rating is studying.

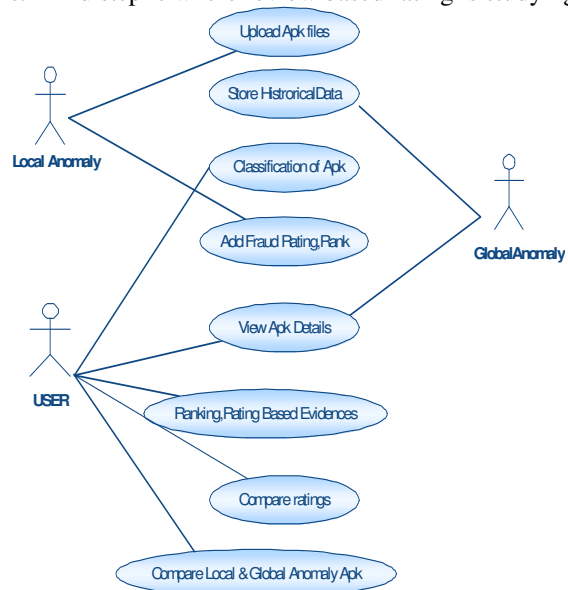


Figure 2: Data flow diagram of mobile fraud detection

V. Algorithm

Step 1: Study the Ranking based evidence and store the information

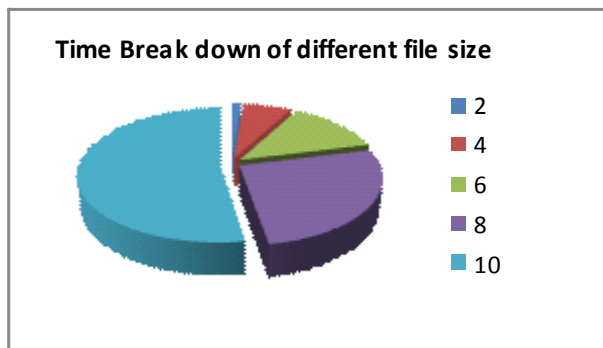
Step 2: Get the details of the Rating based evidence and store it for mining it later.

step 3:review based evidence is studied at the third step,

step 4:Global and local anomaly is studied for the retrieval the true judgment of the mobile app

VI. Results and Discussion

We have studied the mining techniques in determining the mobile fraud acts which happen in making a particular app famous. We studied the various ways of finding the mobile app fraud act. Local and global anomaly is very much important in determining the mobile fraudulent acts. By using the ranking evidence, review evidence session evidence one can get to know the originality of the mobile app which is very important in making the mobile ranking true and distance from forgery..



After getting the true information of the mobile ranking then a graph can be generated which will show the download information, review information of the mobile app. This type of the graph is very important for understanding the download, review information without too much of complexity.

VII. Conclusion and Future Scope

Using the proposed technique we have able to determine the mobile ranking trueness. The proposed technique also gives best performance in terms of less complexity and higher accuracy. The system gives the information in graph format makes it easier and well structured.

In future work we can extend this to get the true information of all other types of software which developer place it online for users to download it. Hence no forgery about the ranking of the software cannot be done, and original information can be placed.

REFERENCES

- [1] L. Azzopardi, M. Girolami, and K. V. Risjbergen, "Investigating the relationship between language model perplexity and its precision-recall measures," in Proc. 26th Int. Conf. Res. Develop. Inform. Retrieval, 2003, pp. 369–370.
- [2] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," J. Mach. Learn. Res., pp. 993–1022, 2003.
- [3] Java The complete Reference.
- [4] My sql Reference Books.