

Determination of Ranking Fraud for Mobile Applications

Sadigale Rohit Birudev¹

Department of computer engineering,
DGOI, FOE, Daund.
Pune, Maharashtra- India
sadigalerohit@gmail.com

Amrit Priyadarshi²

Department of computer engineering,
DGOI, FOE, Daund.
Pune, Maharashtra- India
amritpriyadarshi@gmail.com

Abstract: Mobile application is important for all the smart phone users to play or perform different tasks. There large numbers of mobile application developers are available; they can develop the different mobile applications. For making larger users for their mobile applications some developers refers fraudulent activities. Due to these fraudulent activities the mobile applications jump up in the application popularity list. Such fraudulent activities are used by more and more application developers. The fraudulent activities are like mobile application rating, review and its ranking. For this issue large number of users makes a mistake and downloads the mobile applications which have higher review, rating and ranking. So in this paper, we determine the ranking fraud happens in mobile applications and develop ranking fraud detection system. For identifying the ranking fraud, first we consider leading sessions of mobile applications. Then we examine three types of evidences, these are 1) Ranking based evidence, 2) Rating based evidence and 3) Review based evidence. After this we can aggregate all these evidences for fraud detection. Finally, we develop a system that determines fraud happened in mobile applications.

Keywords: Mobile Applications, ranking fraud detection, ranking evidence, rating evidence, review evidences.

I. Introduction:

There are many mobile applications that have grown at very higher rate over past few years. For example: - at the end of April 2015 there were more than 1.9 million applications at Apple's application store and Google play store. For provoking the development of applications, many application stores provided daily application leader boards, which gave a chart of ranking of most popular applications. For promoting the mobile applications the application leader board is one of the most important ways. The application which has the higher rank on the leader board usually gives the huge number of downloads and million dollars in revenue. Hence, to have their applications ranked as high as possible in such application leader board the application developers does many advertisements to promote their applications in order to get the higher downloads and revenue. But, instead of being dependent such traditional marketing solutions, the fake application developers used some fraudulent techniques to boost their applications and also manipulate the charts rankings on an application store. This is done by using "bot farms" or "human water armies" to inflate the application downloads, ratings and reviews in a very small time. For example, an article from VentureBeat [1] reported that, when an application was launched with the help of ranking manipulation, it could be jumped from number 1,800 to the top 25 in Apple's top free leader board and larger than 50,000-100,000 new users could be acquired within a couple of days. In fact, such ranking fraud raises great concerns to the mobile application industry.

II. Related Work:

The related works of this study consist of spam detection, which contains the web ranking spam refers to rank the selected web pages which belongs to the web site that we have to rank. As there are large numbers of web sites are available, so each developer of website wants higher rank of their website. That fraudulent website ranking spam detection is encountered. Another task is related to the online review spam. When the users buy any product or applications they can give their review, so that review may be fraudulent by the developer. Such Fraud review spams are detected, and the latest fraud is related to the mobile applications.

III. Methodology:

To highlight such problem, in this paper, we have to develop a ranking fraud detection system for mobile applications. To cover such a problem we have to decide the ranking fraud. As the ranking fraud not always happen in the whole life duration of a mobile application, so we need to detect the time when fraud happens. In the real world, the huge numbers of mobile applications are developed by the developer; and that why it is difficult to manually detect ranking fraud for each application. Hence it is important to find out the different way that automatically detects ranking fraud without considering any of the benchmark information. The leather board which is chart that display ranking of mobile applications, and due to the dynamic nature of the chart rankings, it is too difficult to identify and

calculate the evidences that are linked with ranking fraud, which motivates us to examine such ranking fraud by considering some ranking fraud patterns of mobile applications as evidences. That's why, we mainly focus on examining evidences from mobile application historical information like ranking, rating and review records for ranking frauds detection. We also proposed approach that can detect ranking frauds happened in applications historical leading sessions. Fig. 1 shows the framework of ranking fraud detection system for mobile applications.

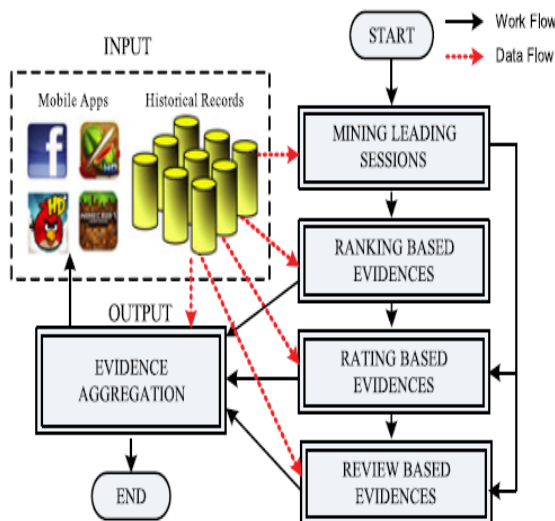


Fig. 1: Framework of ranking fraud detection system for mobile applications.

1) Mining of leading Sessions in mobile Applications:

The leading sessions of a mobile application represent popularity time of mobile application, that the number of users purchases the mobile application, so the ranking manipulation will only take place in these popularity times which is called as leading sessions. Due to this, the problem of detecting ranking fraud is to detect fraudulent leading sessions. Some time mobile application has very low popularity, but that popularity can increase rapidly that session is act as fraudulent leading session and we have to identify such leading session. There are two main phases for mining leading sessions. First, we need to examine leading events from the mobile applications historical ranking information. Second, we need to add adjacent leading events for making the leading sessions.

2) Evidences for Ranking Fraud Detection:

In this section, we study how to extract and combine fraud evidences for ranking fraud detection.

2.1 Evidence Based on Ranking:

The leading session is made up of combining several leading events. For evaluating fraud evidences, we should first identify the basic characteristics of leading events. Then we have to examine the historical ranking information of mobile application, and after examining the historical ranking information we observe the applications ranking behaviors in a leading event that always satisfy a specific ranking phases, which consists of 1) rising phase, 2) maintaining phase and 3) recession phase. Here, in each leading event, the mobile applications ranking first increases to a peak position in the leader board (i.e., rising phase), then it keeps such peak position for a period (i.e., maintaining phase), and after that it is finally decreases till the end of the event (i.e., recession phase).

2.2 Evidences Based on Rating:

The ranking based evidences are helpful for ranking fraud detection. However, sometimes, it is not enough to use only the evidences based on ranking. For e.g., some applications created by the popular developers, such as Gameloft, may have some leading events due to the developers' credibility and the "word-of-mouth" advertising effect [10]. Further, some of the valid marketing offers, like "limited-time discount", may also result in significant evidences based on ranking. To handle such a situation, we have to study how to calculate the fraud evidences from mobile applications' historical rating information. Specifically, after an application has been published, the users can download it and give rating to it. Even, user rating is one of the most important factors of mobile application advertisement. An application which has higher rating that attract more users to download it and which can also be ranked higher in the leader board. Thus, rating manipulation is also an important thing of ranking fraud.

2.3 Evidences Based on Review:

Besides the ranking and ratings, most of the application stores also allow users to write some textual comments or opinion as application reviews. Such reviews can reflect the personal thinking and usage experiences or opinion of users for particular mobile applications. Rather, review exchange is one of the most important factors of application ranking fraud. Specifically, before downloading or purchasing a new mobile application, users may first read its historical reviews or information for making their decision. If mobile application contains more positive or good reviews then it may fetch more and more users to download it. Therefore, the developer often post fake reviews in the leading sessions of a specific application in order to increase the application downloads, and thus jump up the applications ranking position in the leader board.

2.4 Aggregation of Evidences:

After considering such three types of fraud evidence, the next step is to combine all these evidences for determine ranking fraud detection. Indeed, there are many ranking and evidence aggregation methods, such as permutation based models [5], [6], score based models [3], [8] and Dempster-Shafer rules [2], [7]. However, some of these methods focus on learning a global ranking for all candidates. This is not proper way for detecting ranking fraud for new applications. Other methods are based on supervised learning techniques, which depend on the labeled training data and these are hard to be exploited. Instead, we propose an unsupervised method which is based on fraud similarity to combine these evidences.

Conclusion:

In this paper, we studied about the ranking fraud detection system for mobile applications. For that purpose we consider the leading sessions that are useful to identifying ranking fraud happened in leading sessions and gave a method for mining leading sessions. Then, we identified all evidences like, ranking based evidences, rating based evidences and also review based evidences for detecting ranking fraud. After this, we proposed an aggregation method to combine all these evidences for determining the fraudulent mobile applications.

ACKNOWLEDGMENT

I express great many thanks to Prof. Amrit Priyadarshi for their valuable effort of supervising and leading me, to accomplish this fine work, also thankful for Prof. Sachin S. Bere. H.O.D. of my college. To college and department staff, they were a great source of support and encouragement. To my friends and family, for their warm, kind encourages and loves. To every person gave us something too light my pathway, I thanks for believing in me.

References

- [1] (2012). [Online]. Available: <http://venturebeat.com/2012/07/03/apples-crackdown-on-app-ranking-manipulation/>
- [2] Y. Ge, H. Xiong, C. Liu, and Z.-H. Zhou, "A taxi driving fraud detection system," in Proc. IEEE 11th Int. Conf. Data Mining, 2011, pp. 181–190.
- [3] D. F. Gleich and L.-h. Lim, "Rank aggregation via nuclear norm minimization," in Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2011, pp. 60–68.
- [4] A. Klementiev, D. Roth, and K. Small, "An unsupervised learning algorithm for rank aggregation," in Proc. 18th Eur. Conf. Mach. Learn., 2007, pp. 616–623.
- [5] A. Klementiev, D. Roth, and K. Small, "Unsupervised rank aggregation with distance-based models," in Proc. 25th Int. Conf. Mach. Learn., 2008, pp. 472–479.
- [6] A. Klementiev, D. Roth, K. Small, and I. Titov, "Unsupervised rank aggregation with domain-specific expertise," in Proc. 21st Int. Joint Conf. Artif. Intell., 2009, pp. 1101–1106.
- [7] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ, USA: Princeton Univ. Press, 1976.
- [8] M. N. Volkovs and R. S. Zemel, "A flexible generative model for preference aggregation," in Proc. 21st Int. Conf. World Wide Web, 2012, pp. 479–488.
- [9] H. Zhu, H. Cao, E. Chen, H. Xiong, and J. Tian, "Exploiting enriched contextual information for mobile app classification," in Proc. 21st ACM Int. Conf. Inform. Knowl. Manage., 2012, pp. 1617–1621.
- [10] H. Zhu, H. Xiong, Y. Ge, and E. Chen, "Ranking fraud detection for mobile apps: A holistic view," in Proc. 22nd ACM Int. Conf. Inform. Knowl. Manage., 2013, pp. 619–628.
- [11] H. Zhu, E. Chen, K. Yu, H. Cao, H. Xiong, and J. Tian, "Mining personal context-aware preferences for mobile users," in Proc. IEEE 12th Int. Conf. Data Mining, 2012, pp. 1212–1217.