International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 05 | May-2018 www.irjet.net p-ISSN: 2395-0072

DISCOVERY of RANKING FRAUD

Jyoti A. Dhawane[1], Suhas D. Raut[2]

[1,2]*Professor, Department of Computer Science and Engineering, Solapur University*

Abstract - *The smart city, the smart life, India's =young generation was becoming smarter. What is smart, a smart phone, no a smart use of smart phone is a smart life. The different smart applications make the phone, a smart phone. These various applications installed in smart phone make life easier. To download various mobile application smart phone user has to visit respective play store such as Google Play Store, Apples store etc. as per operating system support of smart phone. When user visit play store then he/she is able to see the various application lists. This list is built on the basis of promotion or advertisement. User doesn't have knowledge about the application (i.e. which applications are useful or useless). So user looks at the list and downloads the applications. But sometimes it happens that the downloaded application won't work or not useful and it affects the smartness of phone and at last the users experience. That means it is fraud in mobile application list. To avoid this fraud, we are making application in which we are going to list the applications. To list the application first we are going to find the active period of the application named as leading session. We are also investing the three types of evidences: Ranking based evidence, Rating based evidence and Review based evidence. Using these three evidences finally we are calculating aggregation.*

***Key Words*: fraudulent app, mobile apps, ranking fraud, aggregating function**

1. INTRODUCTION

The number of mobile Apps has grown at a breathtaking rate over the past few years. For example, as of the end of April 2016, there are more than 2.6 million Apps at Apple's App Store and Google Play. The App leaderboard which demonstrates the chart rankings is one of the most important ways of promoting mobile Apps. A higher rank on the leaderboard usually leads to a huge number of downloads and a million dollars in revenue. Therefore, App developers tend to explore various ways such as advertising campaigns to promote their Apps in order to have their Apps ranked as high as possible in such App leaderboards. However, as a recent trend, instead of relying on traditional marketing solutions, shady App developers resort to some fraudulent means to deliberately boost their Apps and eventually manipulate the chart rankings on an App store. This is usually implemented by using so-called "boot farms" or "human water armies" to inflate the App downloads ratings and reviews in a very short time. For example, an article from Venture Beat reported that, when an App was promoted with the help of ranking manipulation, it could be propelled from number 1,800 to the top 25 in Apple's top free leader. Mobile Apps are not generally ranked high on the leaderboard, but rather just in some events ranking that is fraud usually happens in leading sessions. In this manner, the main target is to detect ranking fraud of mobile Apps within leading sessions. First, propose an effective algorithm to distinguish the leading sessions of each App based on its historical ranking records. At that point, an algorithm developed to extract ranking based fraud evidence along with rating and review based evidence. In this way, assist two types of fraud evidence are proposed based on Apps' rating and review history. Moreover, to integrate these three types of an unsupervised evidence-aggregation technique is developed which is utilized for evaluating the credibility of leading sessions from mobile Apps.

2. LITERATURE SURVEY

The paper, detecting product review spammers using rating behaviors [4], aims to detect users generating spam reviews or review spammers. The paper is written by Ee-Peng Lim, Viet-An Nguyen, Nitin Jindal, Bing Liu, Hady W. Lauw. The author identifies several characteristic behaviors of review spammers and models these behaviors seas to detect the spammers. First, spammers may target specific products or product groups in order to maximize their impact. Second, they tend to deviate from the other reviewer in their ratings of products. Paper proposes scoring methods to measure the degree of spam for each reviewer. Then select a subset of highly suspicious reviewers for further scrutiny by our user evaluators with the help of a web-based spammer evaluation software specially developed for user evaluation experiments. Results show that proposed ranking and supervised methods are effective in discovering spammers and outperform other baseline method based on helpfulness votes alone. Finally, the author said that the detected spammers have the more significant impact on ratings compared with the unhelpful reviewers. Evaluative texts on the Web have become a valuable source of opinions on products, services, events, individuals, etc. Existing research has been focused on classification and summarization of opinions using natural language processing and data mining techniques. An important issue that has been neglected so far is opinion spam or trustworthiness of online opinions. In opinion spam and analysis paper, N. Jindal and B. Liu study this issue in the context of product reviews [9], which are opinion rich and are widely used by consumers and product manufacturers. In the past two years, several start-up companies also appeared which aggregate opinions from product reviews. It is thus high time to study spam in reviews. Paper said that opinion spam is quite different from Web spam and email spam, and thus requires different detection techniques. This paper analyses such spam activities and presents some novel techniques to detect them.

3. SYSTEM ARCHITECTURE

Increase in popularity of mobile apps attracts strangers to develop and increases the fraudulent apps which in turn increase the fraud rankings of mobile apps in play store. The various leaderboards of different online mobile app stores contain fraud mobile ranking apps. These fraud ranking misguide the genuine app users and make them download fraud apps rather than genuine apps. Hence, it becomes necessary to discover fraudulent mobile apps. This paper proposes a simple and effective system. Fig. 1 shows the system architecture, framework of the fraud ranking discovery in the mobile app.

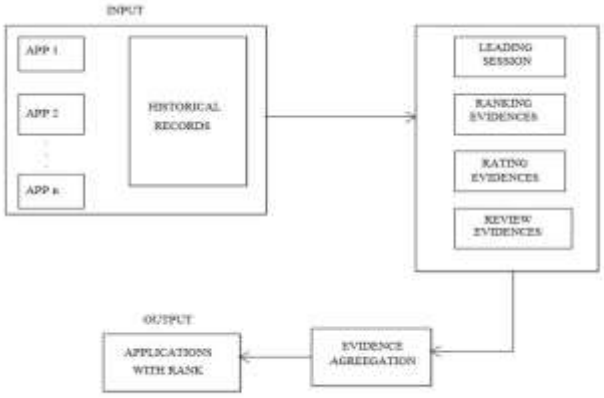

Fig - 1: System architecture, framework of the fraud ranking discovery on mobile app

3.1 Module 1: Leading events

An app contains leading events. A leading event is nothing but the growing period of an app on play store. The increasing in ranking of an app is considered as leading sessions of an app. The leading events will start with growing ranking of an app, will maintain top ranking position for some period and again goes down. These time span of upgrading and downgrading of ranking of a mobile app on leaderboard is repeating process. The growing ranking pattern is leading session. Every leading event contains leading sessions. The genuine apps have different leading session patterns than fraudulent apps. The module first discovers the all leading events and hence leading sessions. The figure 2 shows different ranking phase, such as, rising phase (growing ranking), maintaining phase and recession phase (downgrading ranking) on a mobile app in a leading event on leaderboard.

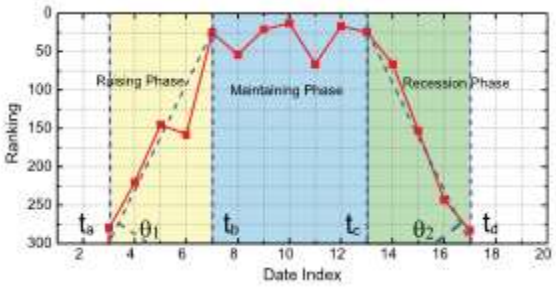

Fig -2: Different ranking phase of leading event

3.2 Module 2: Mining Leading Sessions

Finding out all leading sessions from different leading events and then mining all these leading sessions to discover the fraud ranking is the main objective of this module. Instinctively, mainly the leading sessions of mobile app signify the period of popularity, and so these leading sessions will comprise of ranking manipulation only. Hence, the issue of identifying ranking fraud is to identify deceptive leading sessions. Along with the main task is to extract the leading sessions of a mobile App from its historical ranking records and ongoing ranking records.

Basically, mining leading sessions have two types of steps concerning with mobile fraud apps. Firstly, from the Apps historical ranking records, the discovery of leading events is done and then secondly merging of adjacent leading events is done which appeared for constructing leading sessions. The algorithm is of mining sessions of given mobile App and that algorithm are able to identify the certain leading events and sessions by scanning historical records one by one.

3.3 Module 3: Identifying evidences for ranking fraud detection

The module aims to analysis different evidences for a mobile app in a leaderboard. To achieve these, module performs the fraudulent based analysis for different evidences. The last step is aggregating of these analysis which results in discovering the ranking fraud in mobile apps.

The Ranking Based Evidence, rating based evidences and review based evidences are the three evidences on which module performs the fraudulent based analysis. Let's elaborate the fraudulent analysis on these evidences.

Leading session comprises of various leading events. Hence by analysis of basic behaviour of leading events for finding fraud evidences and also for the app historical ranking records, it is been observed that a specific ranking pattern is always satisfied by app ranking behaviour in a leading event. Ranking of an app on a leaderboard is concluded from ratings and reviews given by app users. Genuine apps have genuine ratings and reviews given by users but fraudulent apps have misleading ratings and reviews. The fraudulent

apps have different and irregular ranking pattern as compared to genuine apps.

Ranking based evidences are useful for detection of fraudulent apps purpose but it is not sufficient. Resolving the problem of "restrict time reduction", identification of fraud evidences is planned due to app historical rating records. As we know that rating is been done after downloading it by the user, and if the rating is high in leaderboard considerably that is attracted by most of the mobile app users. Spontaneously, the ratings during the leading session gives rise to the anomaly pattern which happens during rating fraud. These historical records can be used for developing rating based evidences.

Reviews contains some textual comments as reviews by app user and before downloading or using the app user mostly prefer to refer the reviews given by most of the users. Therefore, although due to some previous works on review spam detection, there still issue on locating the local anomaly of reviews in leading sessions. So based on apps review behaviours, fraud evidences are used to detect the ranking fraud in Mobile app. The reviews are composite of either positive words or negative words. The aggregation and computation of words gives whether given comment if positive or negative for app under consideration. This analysis helps in finding fraudulent apps in a leaderboard.

The proposed new effective algorithm find out fraud ranking in mobile app by aggregating results from above three functions and analysis accordingly.

4. PERFORMANCE ANALYSIS

The algorithm proves effectiveness in discovering fraud ranking in mobile apps published in leader board of different stores of mobile apps. The aggregating function that aggregates results of ranking based evidences, rating based evidences and review evidences executed effectively and gives the better result as compared to previous works in same field. The previous works mainly did in detecting product review spammers using rating behaviours (DPRSRB) [4] and opinion spam or trustworthiness of online opinions (OSA) [9]. The paper, detecting product review spammers using rating behaviours aims to detect users generating spam reviews or review spammers. We identify several characteristic behaviours of review spammers and model these behaviours seas to detect the spammers. Results show that our proposed ranking and supervised methods are effective in discovering spammers in opinion spam and analysis paper, N. Jindal and B. Liu study this issue in the context of product reviews, which are opinion rich and are widely used by consumers and product manufacturers. The figure 3 shows the performance chart of proposed effective algorithm against previous work done in product review spammers and opinion spam in the context of product reviews.

The DPRSRB paper performs analysis on mainly on reviews using rating behaviours while the paper OSA performs analysis mainly on opinions such that reviews given by users. The paper does not consider rating of products. The proposed paper, DRFMA covers all evidences such that, ranking based evidences, rating based evidences and review based evidences.

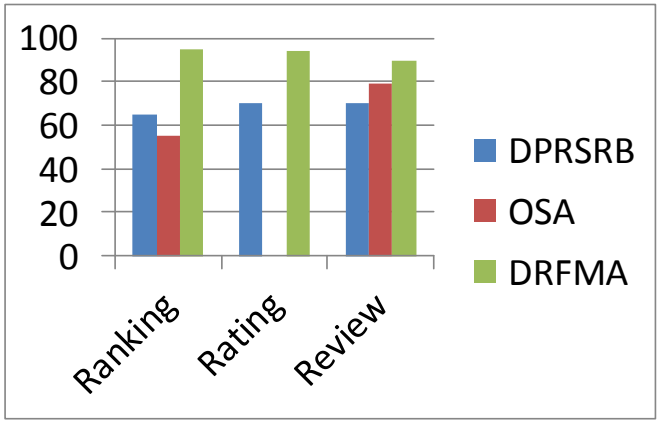

Fig - 3: Performance analysis comparison

5. CONCLUSION

A ranking fraud detection system developed for mobile Apps to protect genuine users from getting fraud apps which effects on an experience of users. Specifically, we first showed that ranking fraud happened in leading sessions and provided a method for mining leading sessions for each App from its historical ranking records. Then an application identified ranking based evidences, rating based evidences and review based evidences for detecting ranking fraud. Moreover, we proposed an optimization based aggregation method to integrate all the evidences for evaluating the credibility of leading sessions from mobile Apps. A unique perspective of this approach is that all the evidences can be modelled by statistical hypothesis tests, thus it is easy to be extended with other evidences from domain knowledge to detect ranking fraud.

ACKNOWLEDGEMENT

I have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals and organisations. I would like to extend my sincere thanks to all of them. I am highly indebted to Mr. Suhas D. Raut for their guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in completing the project. I would like to express my gratitude towards my parents & member of Nagesh Karjagi Orchid College of Engineering and Technology, Solapur for their kind co-operation and encouragement which help me in completion of this project.

REFERENCES

[1] B. Zhou, J. Pei, and Z. Tang. A spamicity approach to web spam detection. In Proceedings of the 2008 SIAM International Conference on Data Mining, SDM'08, pages 277–288, 2008.

[2] A. Ntoulas, M. Najork, M. Manasse, and D. Fetterly. Detecting spam web pages through content analysis. In Proceedings of the 15th international conference on World Wide Web, WWW '06, pages 83–92, 2006.

[3] N. Spirin and J. Han. Survey on web spam detection: principles and algorithms. SIGKDD Explor. Newsl., 13(2):50–64, May 2012.

[4] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw. Detecting product review spammers using rating behaviors. In Proceedings of the 19th ACM international conference on Information and knowledge management, CIKM '10, pages 939–948, 2010.

[5] Z.Wu, J.Wu, J. Cao, and D. Tao. Hysad: a semisupervised hybrid shilling attack detector for trustworthy product recommendation. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '12, pages 985– 993, 2012

[6] Ranking fraud Mining personal context aware preferences for mobile users. H. Zhu, E. Chen, K. Yu, H. Cao, H. Xiong, and J. Tian. In Data Mining (ICDM), 2012 IEEE 12th International Conference on, pages1212–1217, 2012.

[7] Exploiting enriched contextual information for mobile app classification, H. Zhu, H. Cao, E. Chen, H. Xiong, and J. Tian. In Proceedings of the 21st ACM international conference on Information and knowledge management, CIKM '12, pages 1617–1621, 2012.

[8] spammers using behavioral Footprints A. Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu, M. Castellanos,and R. Ghosh. In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '13, 2013.

[9] Detecting product review spammers using rating behaviors. E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw In Proceedings of the 19th ACM international conference on Information and knowledge management, CIKM '10, pages 939–948, 2010