A Review : Mobile App Recommendation Based On Rating Review & Ranking

Dharti A. Bobade[1], Prof.V.S Gangwani[2]

[1]*M.E.Student, Department of Computer Science & Engineering, H.V.P.V College of Engineering, Amravati, India*
[2]*Assistant Professor, Department of Computer Science & Engineering H.V.P.V College of Engineering, Amravati, India*

---***---

Abstract - The Mobile App is a very popular and well known concept due to the rapid advancement in the mobile technology. Due to the large number of mobile Apps, ranking fraud is the key challenge in front of the mobile App market. There are millions of apps are available in market for the application of mobile users. However, all the mobile users first prefer high ranked apps when downloading it. To download application smart phone user has to visit play store such as Google Play Store, Apples store etc. When user visit play store then he 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. 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. In this paper, we provide a brief 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 by using mining leading session algorithm. Furthermore, we investigate three types of evidences, i.e., ranking based evidences, rating based evidences and review based evidences, by studying historical records. we used an optimal aggregation method to integrate all the evidences for fraud detection. Finally, we evaluate the proposed system with real-world App data collected from the Google 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.

Key Words: *Mobile Apps, ranking fraud detection, historical ranking records, evidence aggregation, review, ranking and rating.*

1.INTRODUCTION

AS smartphones emerges new technologies like android and iOS operating system took a boost in market. Mobile application started growing at such a high rate. As a study says millions of apps are there on apple's app store and on Google Play. This started a new business in computer world and became a reason to earn thousands of dollars and downloads. Daily leaderboard is published by these markets contains the most popular apps which will consequently be downloaded and rated most high by users. Some developers may use some marketing strategies like an advertisement campaign for promotion of their app. However this part of technology is also not safe from threats. Mobile app market, we refer it as market, is manipulated by some fraudulent app developers to bump up their app high in the rank list, as an app in leaderboard confirms high downloads and high income. Shady means are used to make such a fraud and implemented using "bot farms" which is also called "Human water armies".

In this area some literature survey is there, for example, spam detection for web ranking, mobile app recommendations, and some online review based spam detection. Our study thus focuses on an integrated approach, for various evidences, to find Mobile App ranking fraud and also recommend the most relevant App that is most genuine. For this we have to go through challenges like first we need to find at what time the fraud is happening it means exact time of fraud is needed. Secondly we know that there is tremendous number of Apps present in market so it is nearly impossible to physically mark ranking fraud for every App, so it's crucial to automatically distinguish fraud without utilizing any essential data. Mobile Apps are not commonly ranked high in the leader board, but instead just in a few events ranking frauds more often than not happens in leading sessions. In this way, fundamental target is to recognize ranking fraud of mobile Apps inside of leading sessions. Initially propose an efficient algorithm to recognize the main sessions of every App depends on its previous ranking records. By then, with the examination of Apps' ranking practices, find the fake Apps consistently have unique ranking examples in every leading session contrasted with ordinary Apps. Along these lines, some fraud confirmations are portrayed from Apps' previous ranking records. By then three limits are produced to concentrate such ranking based fraud confirmations. Thusly, help two kinds of fraud confirmations are proposed taking into account Apps' rating and survey history, which mirror some inconsistency patterns from Apps' previous rating and review records. Also, to coordinate these three kinds of unsupervised proof collection procedure is created which is used for assessing the validity of leading sessions from mobile Apps.

2. Literature Survey

In this section discuss existing work done by the researchers for text mining process. In paper [1], author made ranking fraud location framework for mobile Apps. Specifically, they at first showed that ranking fraud happened in leading sessions and gave a strategy to mining leading sessions for every App from its previous ranking records. By then, they recognized ranking based confirmations, rating based proofs and review based proofs for distinguishing ranking fraud. They likewise proposed a optimization based collection strategy to consolidate each one of the proofs for surveying the validity of leading sessions from mobile Apps. In paper [2], author have focused on various parts of substance build spam regarding the Web and showed different heuristic schedules for recognizing content based spam. Here, they continue with examinations of ”web spam”: the infusion of misleadingly made pages into the web with a particular deciding objective to affect the results from web crawlers, to direct individuals to particular pages for the purpose of excitement or advantage. In paper [3], author has reported a review on Web spam location, which altogether exhibits the rules and algorithm in the literature. Undoubtedly, the work of Web ranking spam identification is principally in light of the examination of ranking measures of web searchers, for instance, PageRank and question term frequency. This is not the same as ranking fraud location for mobile Apps. They sort each present algorithm into three classifications in light of the type of information they use: content-based techniques, link based strategies and techniques on the basis of non-conventional information, for instance, customer behavior, snaps, and HTTP sessions. In paper [4], authors have seen a couple of representative behaviors of review spammers and model these practices to identify the spammers. This paper expects to perceive clients creating spam overviews or review spammers. They perceive a couple trademark practices of review spammers and model these practices with a specific end goal to recognize the spammers. Authors attempt to exhibit the going with practices. In any case, spammers may target specific things or item stores up in order to grow their impact. Second, they have a tendency to go out of order from exchange experts in their assessments of items. In paper [5], authors have analyzed the issue of discovering hybrid shilling attacks on rating information. The methodology depends on can be utilized for reliable item suggestion and the semi-supervised learning. This paper displays a Hybrid Shilling Attack Detector or HySAD for short, to handle this issue. Specifically, HySAD familiarizes MC-Relief with select successful recognition metrics and Semi managed Naive Bays (SNB) to correctly isolate Random-Filler model aggressors and Average- Filler model attackers from standard clients. In paper [6], authors have analyzed the issue of singleton survey spam detection. Specifically, they handled this issue by recognizing the co-anomaly pattern in different review based time arrangement. Also some of above strategies can be used for anomaly detection from previous rating and overview records, they are not prepared to focus fraud evidences for a given time period (i.e., leading session). In paper [7], author created a mobile App recommender framework, Appjoy, which depends on user's App use records to assemble an inclination matrix despite utilizing explicit client ratings. In paper [8], author analyzed a few suggestion models and proposed a content-based collaborative separating model, called Eigenapp, for prescribing Apps in their Web website Getjar. Also, a few researchers analyzed the issue of misusing advanced logical data for mobile App suggestion.

3. Proposed Work

In this paper, we propose to develop a ranking fraud detection system for mobile Apps. Along this line, we identify several important challenges. First, ranking fraud does not always happen in the whole life cycle of an App, so we need to detect the time when fraud happens. Such challenge can be regarded as detecting the local anomaly instead of global anomaly of mobile Apps. Second, due to the huge number of mobile Apps, it is difficult to manually label ranking fraud for each App, so it is important to have a scalable way to automatically detect ranking fraud without using any benchmark information. Finally, due to the dynamic nature of chart rankings, it is not easy to identify and confirm the evidences linked to ranking fraud, which motivates us to discover some implicit fraud patterns of mobile Apps as evidences. Detecting ranking fraud of mobile Apps is actually to detect ranking fraud within leading sessions of mobile Apps. Specifically, we first propose a simple yet effective algorithm to identify the leading sessions of each App based on its historical ranking records. Then, with the analysis of Apps' ranking behaviors, we find that the fraudulent Apps often have different ranking patterns in each leading session compared with normal Apps. Thus, we characterize some fraud evidences from Apps' historical ranking records, and develop three functions to extract such ranking based fraud evidences.

A. Rating based evidences :- Rating to app is given by the user who downloaded it, specifically after the app is published in the market. Hence rating is one of the main evidence in ranking fraud of apps. In this module it performs preprocessing of ratings that is it removes ratings that are less than or equal to two and calculates rating score by summing all the ratings class collected and decision is taken on the basis of rating which scores high amongst all.

B. Review based evidences :- Reviews are familiar to all which provides the way for app user to write some textual comments regarding the personal experience of usage of that particular app. Therefore, manipulation of reviews is one way used by shady app developers to promote their app. Hence reviews are used to detect the ranking fraud in Mobile App industry. This module performs pre-processing of reviews and then performs sentiment analysis on pre-processed reviews. It will find out whether the comment is positive, negative or neutral. If word is positive then it will add plus one to score if word is negative it will minus one from score. Sometimes it is unable to find sentiment of some reviews, that time it makes the use of Naïve Bayes classifier. In this way it will find final score by analyzing sentiment of each review and determine whether app is fraud or not on the basis of review evidences.

C. Ranking based evidences **:-** As per the observation the mobile apps does not always ranked high in the leaderboards, in fact in some leading events only. Further, App having adjacent leading events are merged to form leading sessions. Hence, the problem of identifying ranking fraud is to find out vulnerable leading sessions. There are two phases for mining leading sessions. Firstly, we need to discover the leading events.

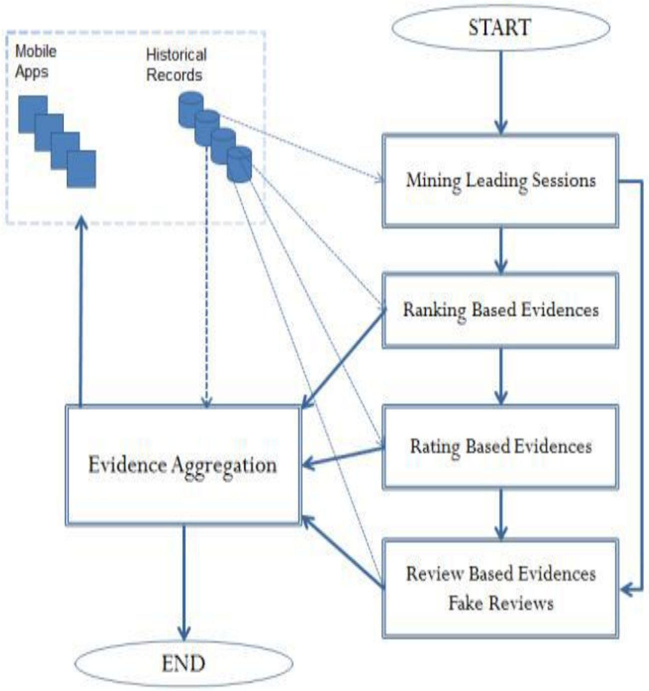

Fig -1: System Architecture

4.CONCLUSION

In this paper, gives the ranking fraud detection model for mobile apps. Now days many of mobile app developers uses various frauds techniques to increase their rank. To avoid this, there are various fraud detection techniques which are studied in this paper. We developed a ranking fraud detection system for mobile Apps. 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, we identified ranking based evidences, rating based evidences and review based evidences for detecting ranking fraud. Moreover, we proposed a mining Leading session algorithm for obtain mining leading session and aggregation method. In the future, we plan to study more effective fraud evidences and analyze the latent relationship among rating, review and rankings. Moreover, we will extend our ranking fraud detection approach with other mobile App related services, such as mobile Apps recommendation, for enhancing user experience. We detect the ranking fraud using actual fraud reviews. This paper proposes the time efficient system to detect the fraud Apps.

ACKNOWLEDGEMENT

I take his opportunity to express my hearty thanks to my guide Prof. V.S Gangwani for his guidance and sharing his findings for technical guidance and direction. Suggestions given by him were always helpful in this work to succeed. His leadership has been greatly valuable for me to work on this project and come with best out of it.

REFERENCES

[1] H. Zhu, H. Xiong, Y. Ge, E. Chen,"Discovery of Ranking Fraud for Mobile Apps", 2015 IEEE.

[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 8392, 2006.

[3] N. Spirin and J. Han. "Survey on web spam detection: principles and algorithms".SIGKDD Explor. News l., 13(2):5064, 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 939948, 2010.

[5] Z.Wu, J.Wu, J. Cao, and D. Tao. "Hysad: a semi-supervised 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 985993, 2012

[6] S. Xie, G. Wang, S. Lin, and P. S. Yu.”Review spam detection via temporal pattern discovery”.In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD 12, pages 823831, 2012.

[7] B. Yan and G. Chen.”Appjoy: personalized mobile application discovery”.In Proceedings of the 9th international conference on Mobile systems, applications, and services, MobiSys 11, pages 113126, 2011.

[8] K. Shi and K. Ali.”Getjar mobile application recommendations with very sparse datasets”.In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD 12, pages 204212, 2012.

[9] http://en.wikipedia.org/wiki/information retrieval.

[10]https://developer.apple.com/news/index.php?id=02062012a.