International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 04 Issue 01
January 2017

A Ranking Fraud Detection System For Mobile Apps

Sayeedakhanum Pathan (Mtech(CSE))
Asst Professor
pathan.sayeeda@gmail.com
Aurora's Scientific Technological and Research Academy

B. SHRUTHI
M.TECH (CSE)
Asst Professor
badugu.shruthi1@gmail.com
Aurora's Scientific Technological and Research Academy

Abstract: *Ranking fraud in the mobile App market refers to false or deceptive activities which have a reason of bumping up the Apps in the popularity list. Certainly, 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. A ranking fraud detection system for mobile Apps was developed. Specifically, this ranking fraud happened in leading sessions and provided a method for mining leading sessions for each App from its historical ranking records and 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. An unique perspective of this approach is that all the evidences can be modelled by statistical hypothesis tests, In this paper we want to propose 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.*

Keyword: Mobile Apps, Ranking Fraud Detection, Evidence Aggregation, Historical Ranking Records, Rating and Review, Recommendation app, KNN.

1. INTRODUCTION

The quantity of mobile Apps has developed at an amazing rate in the course of recent years. For instances, the growth of apps were increased by 1.6 million at Apple's App store and Google Play. To increase the development of mobile Apps, many App stores launched daily App leaderboards, which demonstrate the chart rankings of most popular Apps. Indeed, the App leaderboard is one of the most important ways for promoting mobile Apps. A higher rank on

the leaderboard usually leads to a huge number of downloads and 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 "bot farms" or "human water armies" to inflate the App downloads, ratings and reviews in a very short time[10]. There are some related works, for example, web positioning spam recognition, online survey spam identification and portable App suggestion, but the issue of distinguishing positioning misrepresentation for mobile Apps is till under investigated. The problem of detecting ranking fraud for mobile Apps is still underexplored. To overcome these essentials, in this paper, we build a system for positioning misrepresentation discovery framework for portable apps that is the model for detecting ranking fraud in mobile apps. For this, we have to identify several important challenges. First, fraud is happen any time during the whole life cycle of app, so the identification of the exact time of fraud is needed. 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 automatically detect fraud without using any basic information. Mobile Apps are not always ranked high in the leaderboard, but only in some leading events ranking that is fraud usually happens in leading sessions. Therefore, main target is to detect ranking fraud of mobile Apps within leading sessions. First propose an effective algorithm to identify the leading sessions of each App based on its historical ranking records. Then, with the analysis of Apps' ranking behaviors, find out the fraudulent Apps often have different ranking patterns in each leading session compared with normal Apps. Thus, some fraud evidences are characterized from Apps' historical ranking records. Then three functions are developed to extract such ranking based fraud evidences. Therefore, further two types of fraud evidences are proposed based on Apps' rating and review history, which reflect some anomaly patterns from Apps' historical rating and review records. In addition, to integrate these three types of evidences, an unsupervised evidence-aggregation method is developed which is used for evaluating the credibility of leading sessions from mobile Apps.

II. LITERATURE SURVEY

In this paper, built up a positioning extortion identification framework for versatile applications that positioning misrepresentation happened in driving sessions for each application from its verifiable positioning records.[1] In this technique, we address the issue of survey spammer recognition, or ding clients who are the wellspring of spam audits. Dissimilar to the methodologies for spammed survey recognitions, our proposed audit spammer location methodology is client driven, and client conduct driven. A client driven methodology is favored over the survey driven methodology as social occasion behavioral proof of spammers is less demanding than that of spam audits. An audit includes one and only commentator and one item. The measure of proof is constrained. An analyst then again may have checked on various items and

consequently has contributed various surveys. The probability of closure proof against spammers will be much higher. The client driven methodology is likewise adaptable as one can simply consolidate new spamming practices as they emerge.[2] In this paper we first give a general system for directing Supervised Rank Aggregation. We demonstrate that we can characterize directed learning techniques relating to the current unsupervised strategies, for example, Board Count and Markov Chain based routines by abusing the system. At that point we predominantly research the administered forms of Markov Chain based techniques in this paper, in light of the fact that past work demonstrates that their unsupervised partners are unrivaled. Things being what they are turns out, on the other hand, that the streamlining issues for the Markov Chain based routines are hard, in light of the fact that they are not curved improvement issues. We have the capacity to add to a system the enhancement of one Markov Chain based technique, called Supervised MC2.Specifically, we demonstrate that we can change the advancement issue into that of Semi positive Programming.[3] We first give a general structure for leading Supervised Rank Aggregation. We demonstrate that we can characterize administered learning routines relating to the current unsupervised systems, for example, Board Count and Markov Chain based strategies by abusing the structure. At that point we principally examine the administered variants of Markov Chain based techniques in this paper, in light of the fact that past work demonstrates that their unsupervised partners are predominant. Things being what they are turns out, in any case, that the enhancement issues for the Markov Chain based strategies are hard, in light of the fact that they are not arched advancement issues. We have the capacity to add to a technique the enhancement of one Markov Chain based strategy, called Supervised MC2.Specifically, we demonstrate that we can change the advancement issue into that of Semi positive Programming.[4] In this paper, maker showed diverse sorts of traditions to defend the insurance or security of the data. This paper thought about the issue of essentialness saving in MANETs in perspective of the strategy for framework coding and exhibited that NetworkCoding is beneficial in figuring, and gets less imperativeness usage for encryptions/decodings.[5] In this study, we utilized application use as our metric. Given the attributes of this information, we found that customary memory-based methodologies vigorously support mainstream applications as opposed to our central goal. Then again, inert variable models that were created in light of the Netflix information performed very ineffectively exactness savvy. We find that the Eigenapp model performed the best in precision and in advancement of less understood applications in the tail of our dataset.[6]

To start with the mining driving sessions is utilized to find driving occasions from the application's chronicled positioning records and after that it blends nearby driving occasions for building driving sessions. At that point the positioning based proof dissect the fundamental attributes of driving occasions for separating misrepresentation confirmations. The rating based confirmation is utilized to rate by any client who downloaded it. Audit based confirmation is utilized to check the surveys of the application. The KNN calculation is utilized

to enhance effectiveness and precision of the application. These all proofs are consolidated for recognizing the extortion applications.

III.SYSTEM ARCHITECTURE

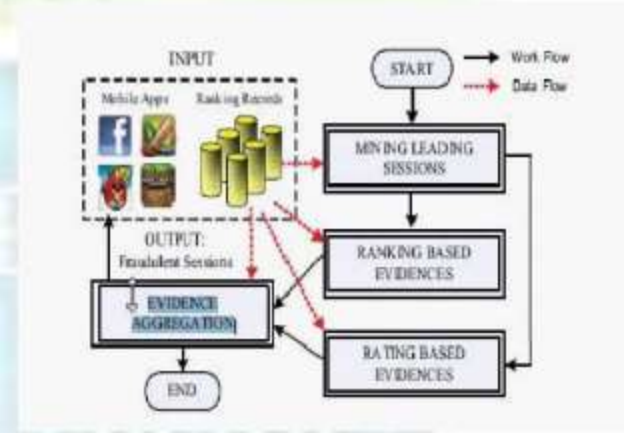

Fig 1. The frame work of the Ranking fraud detection system for Mobile Apps

With the increase in the number of web Apps, to detect the fraud Apps, this paper proposes a simple and effective system. Fig.1 shows the Framework of Fraud ranking discovery in mobile app.

Module 1: Leading events Given a positioning limit K _ 2 [1, K] a main occasion e of App a contains a period range also, relating rankings of a, Note that positioning edge K * is applied which is normally littler than K here on the grounds that K may be huge (e.g., more than 1,000), and the positioning records past K _(e.g., 300) are not exceptionally helpful for recognizing the positioning controls. Moreover, it is finding that a few Apps have a few nearby driving even which are near one another and structure a main session.

Module 2: Leading Sessions 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.

Module 3: Identifying the leading sessions for mobile apps Basically, mining leading sessions has two types of steps concerning with mobile fraud apps.Firstly, from the Apps historical ranking records, discovery of leading events is done and then secondly merging of adjacent leading events is done which appeared for constructing leading sessions.Certainly, some specific algorithm is demonstrated from the pseudo code of mining sessions of given mobile App and that algorithm is able to identify the certain leading events and sessions by scanning historical records one by one.

Module 4: Identifying evidences for ranking fraud detection Ranking Based Evidence It concludes that 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.

Rating Based Evidence Previous ranking based evidences are useful for detection 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.

Review Based Evidence We are familiar with the review which 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 behaviors, fraud evidences are used to detect the ranking fraud in Mobile app.

4. CONCLUSION

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 detect the ranking fraud using actual fraud reviews. This paper proposes the time efficient system to detect the fraud Apps.

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] Getjar mobile application recommendations with very sparse datasets. K. Shi and K. Ali. In Proceedings of the 18th ACM SIGKDDinternational conference on Knowledge discovery and data mining, KDD '12, pages 204–212, 2012.

[7] Ranking fraud Mining personal contextaware 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.

[8] detection for mobile apps H. Zhu, H. Xiong, Y. Ge, and E. Chen. A holistic view. In Proceedings of the 22nd ACMinternational conference on Information and knowledge management,CIKM '13, 2013.

[9] 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.

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

B. Shruthi received her B.Tech degree in Information Technology from Holy Mary Institute Of Technology and Science affiliated to JNTU-Hyderabad in 2013. Now she is doing her M.Tech in Aurora Scientific Technology And Research Academy with Software Engineering under Computer Science Department. Her Area Of interest including C, C++, JAVA and web technologies.

Author's Profile:

Sayeedakhanum Pathan B.E, M.Tech, received her M.Tech from Shadan Women's College of Engineering and Technology affiliated to JNTU Hyderabad University and Received the B.Tech Computer Science and Engineering from Vishwanath Rao Deshpande Rural Institute of Technology Haliyal Karnataka affiliated to VTU Belguam University.Now she is working as Assistant professor in Aurora's Scientific Technological and Research Academy Bandlaguda.Her Area Of interest is data mining.