ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

Headway of Discovery of Ranking Fraud for Mobile Apps

Dr. K. Praveen Kumar

Associate Professor, School of Electrical Engineering & Computing, Department of Computing, Adama Science &Technology University, Adama, Ethiopia

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. Ranking fraud refers to fraudulent or vulnerable activities which have a purpose of bumping up the Apps in the popularity list. In precise, we first propose to precisely locate the mining so one can position misrepresentation the dynamic intervals, to be unique riding sessions, of flexible Apps. Such riding periods can be applied for distinguishing the community oddity in preference to worldwide peculiarity of App ratings. Moreover, we research 3 forms of proofs, i.e., positioning based totally confirmations, modelling on the way to rate primarily based proofs and audit primarily based proofs, Apps' positioning, score and survey practices through measurable speculations tests. What's greater, we suggest a streamlining primarily based general approach to comprise each one of the proofs for misrepresentation detection. The versatile utility thought for ultimately, we investigate the proposed framework with true App records amassed from the iOS App Store for pretty some time duration. In the severities, we approve the adequacy of the proposed framework, and show the adaptability of the recognition calculation and also a few normality of positioning extortion sporting events.

KEYWORDS: Apps, ranking fraud detection, evidence aggregation,historical ranking records, Recommendation app, KNN.

I. 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 leader boards, 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.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

Then, with the analysis of Apps' ranking behaviours, 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. RELATED WORK

The related works of this study is grouped into three categories. The first category is about Web ranking spamdetection. Specifically, the Web ranking spam refers to any deliberate actions which bring to selected Web pages anunjustifiable favorable relevance or importance. In this, the problem of unsupervised web spam detection is studied.They introduce the concept of spamicity to measure howlikely a page is spam. Spamicity is more flexible and usercontrollable measure than the traditional supervisedclassification methods. They propose efficient online linkspam and term spam detection methods using spamicity.These methods do not need training and also cost effective. Areal data set is used to evaluate the effectiveness and theefficiency [1].

For example, Ntoulas et al. [2] have studied various aspectsof content-based spam on the Web and presented a numberof heuristic methods for detecting content based spam.In this paper, they continue investigations of "web spam":the injection of artificially-created pages into the web inorder to influence the results from search engines, to drivetraffic to certain pages for fun or profit. This paper considerssome previously-undescribed techniques for automaticallydetecting spam pages, examines the effectiveness of thesetechniques in isolation and when aggregated usingclassification algorithms.

Zhou et al [1] have studied the problem of unsupervised Web ranking spam detection. Specifically, they proposed anefficient online link spam and term spam detection methodsusing spamicity.Recently, Spirin et al. [3] have reported a survey on Webspam detection, which comprehensively introduces theprinciples and algorithms in the literature. Indeed, the workof Web ranking spam detection is mainly based on theanalysis of ranking principles of search engines, such asPageRank and query term frequency. This is different fromranking fraud detection for mobile Apps. They categorize allexisting algorithms into three categories based on the type ofinformation they use: content-based methods, link-basedmethods, and methods based on non-traditional data such asuser behavior, clicks, HTTP sessions. In turn, there is asub categorization of link-based category into five groupsbased on ideas and principles used: labels propagation, linkpruning and reweighting, labels refinement, graphregularization, and feature based.

The second category is focused on detecting online reviewspam. For example, Lim et al. [4] have identified severalrepresentative behaviors of review spammers and modelthese behaviors to detect the spammers. This paper aims todetect users generating spam reviews or review spammers.They identify several characteristic behaviors of reviewspammers and model these behaviors so as to detect thespammers. In particular, authors seek to model the followingbehaviors. First, spammers may target specific products orproduct groups in order to maximize their impact. Second,they tend to deviate from the other reviewers in their ratingsof products. They propose scoring methods to measure thedegree of spam for each reviewer and apply them on anAmazon review dataset. Authors then select a subset ofhighly suspicious reviewers for further scrutiny by userevaluators with the help of a web based spammer evaluationsoftware specially developed for user evaluationexperiments.

Wu et al. [5] have studied the problem of detecting hybridshilling attacks on rating data. The proposed approach isbased on the semi-supervised learning and can be used fortrustworthy product recommendation. This paper presents aHybrid Shilling Attack Detector, or HySAD for short, totackle these problems. In particular, HySAD introduces MCRelief to select effective detection metrics, and Semi supervised Naive Bayes (SNBλ) to precisely separateRandom-Filler model attackers and Average-Filler modelattackers from normal users.

Xie et al. [6] have studied the problem of singleton reviewspam detection. Specifically, they solved this problem bydetecting the co-anomaly patterns in multiple review basedtime series. Although some of above approaches can be

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

usedfor anomaly detection from historical rating and reviewrecords, they are not able to extract fraud evidences for agiven time period (i.e., leading session).Finally, the third category includes the studies on mobileApp recommendation. For example, Yan et al. [7] developeda mobile App recommender system, named Appjoy, whichis based on user's App usage records to build a preferencematrix instead of using explicit user ratings.

Also, to solve the sparsity problem of App usage records,Shi et al. [8] studied several recommendation models andproposed a content based collaborative filtering model,named Eigenapp, for recommending Apps in their Web siteGetjar. In addition, some researchers studied the problem ofexploiting enriched contextual information for mobile Apprecommendation.For example, Zhu et al. [9] proposed a uniform frameworkfor personalized context-aware recommendation, which canintegrate both context independency and dependencyassumptions. However, to the best of our knowledge, noneof previous works has studied the problem of ranking frauddetection for mobile Apps.

III.PROPOSED SYSTEM

To start with the mining driving sessions is utilized to find driving occasions from the application's chronicledpositioning records and after that it blends nearby driving occasions for building driving sessions. At that point thepositioning based proof dissect the fundamental attributes of driving occasions for separating misrepresentationconfirmations. 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.

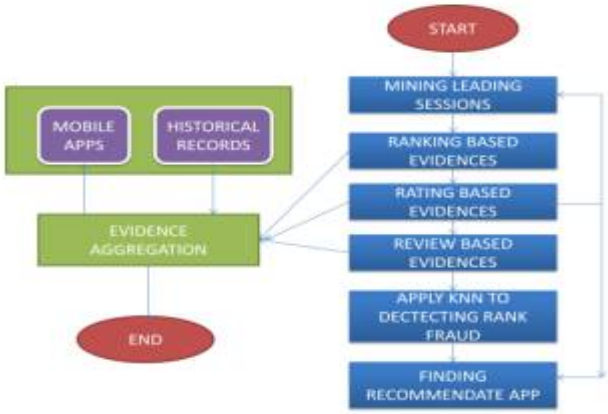

Fig.1 Basic System Architecture

Identifying Leading Sessions: Ranking fraud usually happens in leading sessions. Therefore, 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.

Ranking Based Evidences: A leading session is composed of several leadingevents. Therefore, we should first analyse the basic characteristics of leading events for extracting fraud evidences. By analysing theApps' historical ranking records, we observe that Apps' ranking behaviors in a leading event always satisfy a specific rankingpattern, which consists of three different ranking phases, namely, rising phase, maintaining phase and recession phase.

Rating Based Evidences: The ranking based evidences are useful for ranking fraud detection. However, sometimes, it is not sufficient to only useranking based evidences. Specifically, after an App has been published, it can be rated by any user who downloaded it. Indeed,user rating is one of the most important features of App advertisement.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

Review Based Evidences: Besides ratings, most of the App stores also allow users to write some textual comments as App reviews. Such reviews can reflect the personal perceptions and usage experiences of existing users for particular mobile Apps. Indeed, review manipulation is one ofthe most important perspectives of App ranking fraud.

IV. CONCLUSION

This paper, gives the ranking fraud detection model for mobile apps. Now days many ofmobile app developers uses various fraudstechniques to increase their rank. This paper presents more effective fraud evidences andanalyze the latent relationship among rating, review and rankings. We prolonged our ranking fraud detection approachwith other mobile app related services, such as mobile app recommendation for enhancing user experience.

REFERENCES

[1] B. Zhou, J. Pei, and Z. Tang. A spamicity approach toweb spam detection. In Proceedings of the 2008 SIAMInternational 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. InProceedings of the 15th international conference onWorld 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 ratingbehaviors. In Proceedings of the 19th ACMinternational conference on Information and knowledgemanagement, CIKM '10, pages 939–948, 2010.
[5] Z.Wu, J.Wu, J. Cao, and D. Tao. Hysad: a semi supervised hybrid shilling attack detector fortrustworthy product recommendation. In Proceedings ofthe 18th ACM SIGKDD international conference onKnowledge discovery and data mining, KDD '12, pages985–993, 2012
[6] S. Xie, G. Wang, S. Lin, and P. S. Yu. Review spamdetection via temporal pattern discovery. InProceedings of the 18th ACM SIGKDD internationalconference on Knowledge discovery and data mining,KDD '12, pages 823–831, 2012.
[7] B. Yan and G. Chen. Appjoy: personalized mobileapplication discovery. In Proceedings of the 9th international conference on Mobile systems,applications, and services, MobiSys '11, pages 113–126, 2011.
[8] K. Shi and K. Ali. Getjar mobile applicationrecommendations with very sparse datasets. InProceedings of the 18th ACM SIGKDD internationalconference on Knowledge discovery and data mining,KDD '12, pages 204–212, 2012.
[9] R. Agrawal and R. Srikant, "Fast algorithms for miningassociation rules," in VLDB, 1994.
[10] Spammers using behavioural Footprints A.Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu,M. Castellanos,and R. Ghosh. In Proceedings ofthe 19th ACM SIGKDD international conferenceon Knowledge discovery and data mining, KDD'13, 2013.

BIOGRAPHY

Dr.K.Praveen Kumarreceived the PhD in Computer Science & Engineering in 2015, M.Tech in Software Engineering from Kakatiya Institute of Technology & Science Warangal, Telangana, India in 2010 and B.Tech in Information Technology from Kakatiya Institute of Technology & Science Warangal, Telangana, India 2007. Presently working as Assistant Professor in Computer Science Department at Adama Science and Technology University, Adama, Ethiopia.