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 7, July 2016

Streamlining the Ranking Methodology for Mobile App Leader Boards

Budharaju Teja Sree, Dr.I.Hemalatha

PG Scholar, Dept. of IT, Sagi Rama Krishnam Raju Engineering College, Bhimavaram, India

Associate Professor, Dept. of IT, Sagi Rama Krishnam Raju Engineering College, Bhimavaram, India

ABSTRACT: Ranking fraud of mobile apps refer to the fraud activities with the purpose of raising apps in popular list. Indeed, 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. To this end, in this project, we provide a holistic view of ranking fraud and propose a ranking fraud detection system for mobile Apps. Considering the ranking frauds in the mobile market, we are proposing a ranking fraud detection system for mobile Apps to accurately locate the ranking fraud by mining the active periods, namely leading sessions, of mobile Apps. Such leading sessions can be leveraged for detecting the local anomaly instead of global anomaly of App rankings We also Investigate based on three types of evidences through statistical tests ranking based evidences, rating based evidences and review based evidences.

KEYWORDS: Mobile Apps, ranking fraud detection, evidence aggregation, historical ranking records, rating and review.

I. INTRODUCTION

The amount of convenient Apps has created at a stunning ate throughout late years. Case in point, as of the end of April 2013, there are more than 1.6 million Apps at Apple's App store and Google Play. To vivify the change of adaptable Apps, various App stores moved each day App pioneer sheets, which show the diagram rankings of most common Apps. As a general rule, the App pioneer board is a champion amongst the most key courses for progressing flexible Apps. A higher rank on the pioneer board generally prompts countless and million dollars in salary. The quantity of portable Apps has experienced childhood with a substantial scale in the course of recent years. For example, there are more than 1.6 million Apps at Apple's App store and Google Play toward the end of April 2013. To move the improvement of portable Apps, numerous App stores propelled every day App pioneer board, which demonstrates the diagram rankings of most well-known Apps. This sort of applications is the most vital routes for advancing portable Apps. A top rank on the pioneer board more often than not prompts countless and million dollars in income. Therefore, App designers grade to investigate different routes, for example, ad drive to elevate their Apps to get higher position in such App leader boards. The late pattern in business sector utilized by the exploitative App designers for App boosting is to utilize fake intends to deliberately support their applications. Finally, they additionally twist the graph rankings on an App store. This is typically executed by utilizing purported "web bots" or "human water armed forces" to raise the App downloads, appraisals and surveys in an almost no time. For instance, Venture Beat [1] reported that, when an App was advanced utilizing positioning control, it could be accelerated from number 1,800 to the highest 25 in Apple's sans top leader board and more than 50,000-100,000 new clients could be gained inside two or three days. In fact, such positioning misrepresentation elevates awesome worries to the portable App industry. For instance, Apple has advised of taking action against App designers who confer positioning extortion [2] in the App store. Driving occasions of portable Apps shapes distinctive driving sessions. The versatile Apps not generally positioned high in the leader boards, but rather it more often than not happens in the main sessions. Along these lines, recognizing positioning extortion of horde Apps is really the procedure to identify it inside the main session of the versatile Apps. Particularly, this paper proposes a basic and successful calculation to perceive the main sessions of every portable App in light of its chronicled positioning records. This is one of the extortion proof. Likewise, two sorts of misrepresentation proofs are proposed in light of Apps' appraising and audit history, which gives some oddity designs from Apps' authentic rating

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 7, July 2016

and survey records. Moreover, we propose an unsupervised evidence aggregation strategy to unite these three sorts of proofs for surveying the believability of driving sessions from versatile Apps.

II. RELATED WORK

The related works of this study is assembled into three classes. The main classification is about Web positioning spam identification. In particular, the Web positioning spam alludes to any intentional activities which convey to chose Web pages a baseless ideal pertinence or significance. In this, the issue of unsupervised web spam location is contemplated. They acquaint the idea of spam city with measure how likely a page is spam. Spamicity is more adaptable and client controllable measure than the conventional regulated order strategies. They propose proficient online connection spam and term spam location strategies utilizing spamicity. This techniques needn't bother with preparing furthermore financially savvy. A genuine information set is utilized to assess the adequacy and the effectiveness [1]. For instance, Ntoulas et al. [2] have examined different parts of substance construct spam in light of the Web and introduced various heuristic techniques for identifying content based spam. In this paper, they proceed with examinations of "web spam": the infusion of falsely made pages into the web to impact the outcomes from web crawlers, to direct people to certain pages for the sake of entertainment or benefit. This paper thinks of some as beforehand un described procedures for naturally distinguishing spam pages, analyzes the viability of these methods in seclusion and when amassed utilizing grouping calculations. Zhou et al [1] have contemplated the issue of unsupervised Web positioning spam recognition. In particular, they proposed a productive online connection spam and term spam recognition strategies utilizing spamicity. As of late, Spirin et al. [3] have reported an overview on Web spam identification, which completely presents the standards and calculations in the writing. Without a doubt, the work of Web positioning spam recognition is predominantly taking into account the examination of positioning standards of web indexes, for example, Page Rank and inquiry term recurrence. This is not the same as positioning misrepresentation location for versatile Apps. They sort every single existing calculation into three classifications in view of the kind of data they utilize: content-based strategies, join based techniques, and techniques in light of non-customary information, for example, client conduct, clicks, HTTP sessions. Thusly, there is a sub categorization of connection based class into five gatherings in view of thoughts and standards utilized: marks spread, join pruning and reweighting, names refinement, diagram regularization, and highlight based. The second classification is centered around distinguishing online survey spam. For instance, Lim et al. [4] have distinguished a few agent practices of survey spammers and model these practices to identify the spammers. This paper plans to identify clients creating spam surveys or audit spammers. They distinguish a few trademark practices of audit spammers and model these practices in order to recognize the spammers. Specifically, creators look to show the accompanying practices. To start with, spammers may target particular items or item amasses keeping in mind the end goal to amplify their effect. Second, they tend to veer off from alternate commentators in their evaluations of items. They propose scoring techniques to quantify the level of spam for every commentator and apply them on an Amazon audit dataset. Creators then select a subset of profoundly suspicious commentators for further examination by client evaluators with the assistance of an online spammer assessment. Programming exceptionally created for client assessment tests. Wu et al. [5] have considered the issue of recognizing half and half shilling assaults on rating information. The proposed methodology depends on the semi-administered learning and can be utilized for reliable item suggestion. This paper displays a Hybrid Shilling Attack Detector, or HySAD for short, to handle these issues. Specifically, HySAD acquaints MCRelief with select compelling location measurements, and Semi directed Naive Bayes (SNBλ) to correctly isolate Random-Filler model aggressors and Average-Filler model assailants from typical clients. Xie et al. [6] have concentrated on the issue of singleton survey spam discovery. In particular, they tackled this issue by distinguishing the co-oddity designs in different survey based time arrangement. Albeit some of above methodologies can be utilized for abnormality identification from authentic rating and audit records, they are not ready to concentrate misrepresentation confirmations for a given time period (i.e., driving session). At long last, the third class incorporates the studies on versatile App proposal. For instance, Yan et al. [7] built up a portable App recommender framework, named Appjoy, which depends on client's App use records to assemble an inclination lattice as opposed to utilizing unequivocal client appraisals. Additionally, to tackle the sparsity issue of App use records, Shi et al. [8] contemplated a few suggestion models and proposed a substance based synergistic sifting model, named Eigen app, for prescribing Apps in their Web website Getjar. Also, a few specialists contemplated the issue of misusing enhanced logical data for portable App suggestion. For instance, Zhu et al. [9] proposed a uniform structure for customized setting mindful suggestion, which can incorporate both connection independency and reliance suppositions.

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 7, July 2016

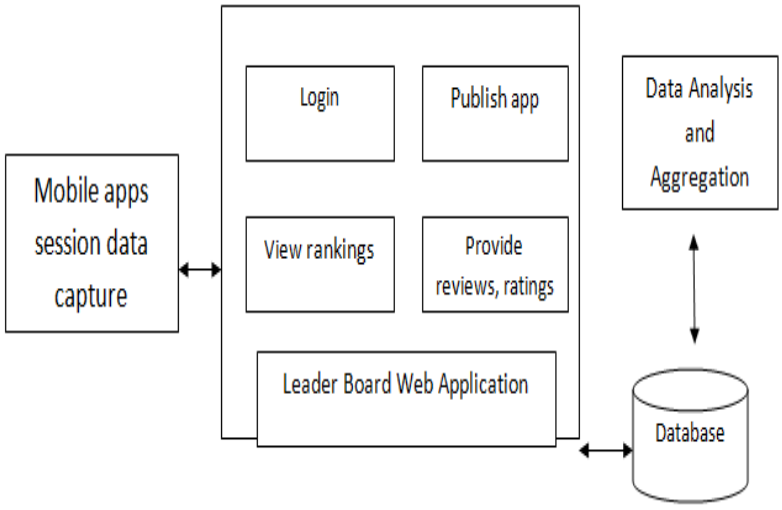
In any case, to the best of our insight, none of past works has considered the issue of positioning extortion discovery for versatile Apps.

III. PROPOSED SYSTEM

1. Ranking frauds in the mobile App market

- Fraudulent activities for bumping up their Apps in the popularity list and inflating their Apps sales.
- Posting phony App ratings
- 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 leaderboard and more than 50,000-100,000 new users could be acquired within a couple of days.
- Such ranking fraud raises great concerns to the mobile App industry and the detection of fraud ranking is under exploration.
- In the proposed system we will detect the ranking fraud for the mobile applications. Detecting the ranking fraud for the mobile apps is to actually detect the ranking fraud within the leading sessions of the mobile apps. Leading sessions of the mobile apps represents its period of popularity.
- For this proposed system in order to get the leading sessions we perform the Mining leading sessions technique .In this we perform two tasks we will discover the leading events of app from apps history and merge the adjacent leading events to construct the leading sessions.
- The proposed system is scalable and can be extended with other domain generated evidences for ranking fraud detection.
- Use of the mining leading sessions gives more accurate results.
- Provides the app information and can be used to know the original app ranking.

Architecture:



There are two fundamental stages for distinguishing the positioning misrepresentation:
i) Identifying the main sessions for versatile applications
ii) Identifying confirmations for positioning extortion discovery

i). Distinguishing the main sessions for portable applications: Primarily, mining driving sessions has two sorts of steps worried with versatile extortion applications. To start with, from the Apps chronicled positioning records, revelation of driving occasions is done and afterward second converging of nearby driving occasions is done which showed up for building driving sessions. Surely, some particular calculation is exhibited from the pseudo code of mining sessions of given portable App and that calculation can recognize the specific driving occasions and sessions by checking verifiable records one by one.

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 7, July 2016

ii). Recognizing confirmations for positioning extortion discovery

1) Ranking based confirmations:

It infers that driving session contains different driving occasions. Subsequently by examination of fundamental conduct of driving occasions for discovering extortion confirmations furthermore for the application chronicled positioning records, it is been watched that a particular positioning example is constantly fulfilled by application positioning conduct in a main occasion.

2) Rating based proofs:

Past positioning based proofs are valuable for recognition reason however it is not adequate. Determining the "limit time exhaustion" issue, misrepresentation confirmations acknowledgment is arranged because of application chronicled rating records. As we realize that rating is been done subsequent to downloading it by the client, and if the rating is high in leaderboard impressively that is pulled in by the greater part of the versatile application clients. Suddenly, the evaluations amid the main session offers ascend to the oddity design which happens amid rating extortion. These chronicled records can be utilized for creating rating based proofs.

3). Review based proofs:

We are acquainted with the survey which contains some printed remarks as audits by application client and before downloading or utilizing the application client for the most part want to allude the audits given by the majority of the clients. Consequently, albeit because of some past takes a shot at audit spam recognition [13] there still issue on finding the nearby irregularity of surveys in driving sessions. So in light of applications survey practices, misrepresentation confirmations are utilized to identify the positioning extortion in Mobile App.

IV. CONCLUSION

This paper surveys different existing techniques utilized for web spam recognition, which is identified with the positioning misrepresentation for versatile Apps. Additionally, we have seen references for online audit spam discovery and versatile App recommendation.By mining the main sessions of portable Apps, we expect to find the positioning misrepresentation. The main sessions works for identifying the nearby irregularity of App rankings. The framework intends to identify the positioning cheats in light of three sorts of confirmations, for example, positioning based proofs, rating based proofs and audit based confirmations. Further a streamlining based conglomeration technique joins all the three confirmations to distinguish the extortion.

V. FUTURE WORK

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.

REFERENCES

[1] Discovery of ranking fraud for mobile apps. HengshuZhu, HuiXiong, Senior members, IEEE, YongGe, and EnhongChen, Senior member, IEEE, IEEE transactions on knowledge and data engineering,vol .27, No.1, January 2015.
[2] 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.
[3] Supervised rank aggregation. Y.-T. Liu, T.-Y. Liu, T. Qin, Z.-M. Ma, and H. Li In Proceedings of the 16th international conference onWorld Wide Web.
[4] Z. Wu, J. Wu, J. Cao, and D. Tao, "HySAD: A semi supervised hybrid shilling attack detector for trustworthy product recommendation," in Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2012, pp. 985–993.
[5] S. Xie, G. Wang, S. Lin, and P. S. Yu, "Review spam detection via temporal pattern discovery," in Proc. 18th ACM SIGKDD Int.Conf. Knowl. Discovery Data Mining, 2012, pp. 823–831.
[6] K. Shi and K. Ali, "Getjar mobile application recommendations with very sparse datasets," in Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2012, pp. 204–212.
[7] B. Yan and G. Chen, "AppJoy: Personalized mobile application discovery," in Proc. 9th Int. Conf. Mobile Syst., Appl., Serv., 2011,pp. 113–126.
[8] An unsupervised learning algorithm for rank aggregation, A. Klementiev, D. Roth, and K. Small In Proceedings of the 18th European conference on Machine Learning, ECML '07, pages 616–623, 2007.
[9] Getjar mobile application recommendations with very sparse datasets. K. Shi and K. Ali. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '12, pages 204–212, 2012.

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 7, July 2016

[10] 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.
[11] 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.
[12] 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.
[13] 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.
[14] 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.
[16] N. Spirin and J. Han, "Survey on web spam detection: Principlesandalgorithms," SIGKDD Explor. Newslett., vol. 13, no. 2, pp. 50–64, May 2012.
[17] B. Zhou, J. Pei, and Z. Tang, "A spamicity approach to web spamdetection," in Proc. SIAM Int. Conf. Data Mining, 2008, pp. 277–288.
[18] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw,"Detecting product review spammers using rating behaviors," inProc. 19thACMInt. Conf. Inform. Knowl. Manage., 2010, pp. 939–948.
[19] Z. Wu, J. Wu, J. Cao, and D. Tao, "HySAD: A semi-supervisedhybrid shilling attack detector for trustworthy productrecommendation," in Proc. 18th ACM SIGKDD Int. Conf. Knowl.Discovery Data Mining, 2012, pp. 985–993.
[20] S. Xie, G. Wang, S. Lin, and P. S. Yu, "Review spam detection viatemporal pattern discovery," in Proc. 18th ACM SIGKDD Int.Conf. Knowl. Discovery Data Mining, 2012, pp. 823–831.
[21] H. Zhu, H. Xiong, Y. Ge and E. Chen. Discovery of Ranking Fraud for Mobile Apps in IEEE, 2015.

BIOGRAPHY

Budharaju Teja Sree is currently pursuing her M.Tech(IT) in Information Technology Department, Sagi Rama Krishnam Raju Engineering College, West Godavari, A.P. She received her B.Tech in Information Technology Department from Sagi Rama Krishnam Raju Engineering College, Bhimavaram.

Dr.I.Hemalatha is currently working as an Associate Professor in Information Technology Department, Sagi Rama Krishnam Raju Engineering College, West Godavari. Her research includes networking and data mining.