International Conference on Systems, Science, Control, Communication, Engineering and Technology 2016 [ICSSCCET 2016]

ISBN	978-81-929866-6-1	**VOL**	02
Website	icssccet.org	**eMail**	icssccet@asdf.res.in
Received	25 – February – 2016	**Accepted**	10 - March – 2016
Article ID	ICSSCCET101	**eAID**	ICSSCCET.2016.101

Survey of Developing a Model of Analysing and Detecting for Leaderboard Discrepancies of Mobile Application

Maria Varsha B[1], Kayalvizhi K[2], Hemalatha A[3], Loganya K[4]
[1,2,3,4] Student Scholar, Department of Information Technology, Karpagam Institute of Technology, Coimbatore.

Abstract: *Recently, many number of research has resulted that the opinion sources as product reviews, forum posts, and blogs. Ranking the fraud mobile App with more ratings is a fraudulent or deceptive activities which have a intension to indicate that Apps in the popularity list. It becomes an added advantage for the App developers to increasing their Apps' sales or posting higher ratings, to commit ranking fraud. Though the importance of preventing ranking fraud has been already recognized, there is only limited research and work has been done in this area. The existing research focuses on classification and summarization of opinions using online analytical processing and data mining techniques. In this proposed system a holistic view of ranking fraud and propose a fraud detection system for mobile Apps. Specifically, we first propose to locate the ranking fraud by mining the leading sessions, of mobile Apps. We further investigate three types of evidences, i.e., ranking, rating and review, based on mobile Apps' ranking, rating and review behaviours. Based on these records and analysis, we can generate the report on fraud detection. To aggregate these reviews, OLAP data aggregation method is used. In the proposed system, user feedbacks can be collected. These feedbacks can contain both the positive and negative feedbacks. User can select the url of the mobile app. From the user click, user can identify the original or fake links. And also we can analyse the user behaviour or user interest by analysing the session history. From this analysis we can conclude, the most frequently used apps by users. K-means clustering algorithm is used to cluster the fake profile comments.*

Keywords—*Mobile Apps, ranking fraud detection, evidence aggregation, historical ranking records, rating and review*

INTRODUCTION

Today, most of us use android phone these days and also uses the play store capability normally. For example, as of the end of 2015, there are more than million Apps were developed at Apple's store and Google Play. To manage the development of mobile Applications, App stores launched daily Application leader board, which shows the chart rankings of most popular applications. Leader board is one of the most important ways for promoting mobile Apps. The top application on the leader board usually leads to more number of downloads and million dollars in revenue. Therefore, app developer started advertising campaigns to promote the application ranked as high as possible in the leaderboard.

In recent trend, some shady application developers resort to fraudulent ways to deliberately boost their applications and eventually manipulate the chart rankings on an App store. Usually, this is implemented by using so-called "botfarms" or "human water armies" which increase the App downloads ratings and reviews in a very short time. Play store provide number of application but unluckily some of those applications are fraud. Such applications may cause damage to phone and also may be data thefts happen. Hence such applications must be identifiable, so that they will be useful for play store users. So we are developing a web application which will process the information based on ratings, ranking and reviews of the application using online analytic processing to give results in the

Cite this article as: Maria Varsha B, Kayalvizhi K, Hemalatha A, Loganya K. "Survey of Developing a Model of Analysing and Detecting for Leaderboard Discrepancies of Mobile Application". *International Conference on Systems, Science, Control, Communication, Engineering and Technology 2016*: 509-511. Print.

form of graph. So it is helpful for user to decide whether the application is fraud or not. Multiple applications can be processed at a time through the web application. At the same time, user cannot get correct or true reviews about the product on internet. So we can check for reviews of same product more than one user feedback. Hence we can get higher probability of getting real reviews.

Related Work

In existing system, many methods are introduced. Web ranking spam refers to any deliberate actions which bring to selected web pages an unjustifiable favourable relevance or importance. The work of web ranking spam detection is mainly based on analysis of ranking principles of search engines, such as page rank and query term frequency. In online review spam method, identified the behavior of review spammers and model these behaviors to detect the spammers. Generally speaking, the related works of this study can be grouped into three categories.

The first category is about web ranking spam detection. Specifically, the web ranking spam refers to any deliberate actions which bring to selected web pages an unjustifiable favorable relevance or importance. For example, Ntoulas et al have studied various aspects of content-based spam on the web and presented a number of heuristic methods for detecting content based spam. Zhou et al have studied the problem of unsupervised web ranking spam detection. Specifically, they proposed an efficient online link spam and term spam detection methods using spam city. Recently, Spirin and Han have reported a survey on web spam detection, which comprehensively introduces the principles and algorithms in the literature. Indeed, the work of web ranking spam detection is mainly based on the analysis of ranking principles of search engines, such as Page Rank and query term frequency. This is different from ranking fraud detection for mobile Apps.

The second category is focused on detecting online review spam. For example, Lim et al have identified several representative behaviors of review spammers and model these behaviors to detect the spammers. Wu et al have studied the problem of detecting hybrid shilling attacks on rating data. The proposed approach is based on the semi supervised learning and can be used for trustworthy product recommendation. Xie et al have studied the problem of singleton review spam detection. Specifically, they solved this problem by detecting the co-anomaly patterns in multiple review based time series. Although some of above approaches can be used for anomaly detection from historical rating and review records, they are not able to extract fraud evidences for a given time period.

Finally, the third category includes the studies on mobile App recommendation. For example, Yan and Chen developed a mobile App recommender system, named Appjoy, which is based on user's App usage records to build a preference matrix instead of using explicit user ratings. Also, to solve the sparsity problem of App usage records, Shi and Ali studied several recommendation models and proposed content based collaborative filtering model, named Eigenapp, for recommending Apps in their website Getjar. In addition, some researchers studied the problem of exploiting enriched contextual information for mobile App recommendation. For example, Zhu et al proposed a uniform framework for personalized context-aware recommendation, which can integrate both context independency and dependency assumptions. However, to the best of our knowledge, none of previous works has studied the problem of ranking fraud detection for mobile Apps.

Conclusion

It is the easiest way of fraud ranking detection system for the mobile Apps. Specifically, we first showed that the ranking fraud happened within the leading sessions and also provided a method for mining leading sessions for each App from its provided historical ranking records. Then, we can identify the ranking, rating and the review based evidences for detecting the ranking fraud. Moreover, we proposed an optimized the aggregation method to the integrate all the evidences for evaluating the credit for the mobile Apps. An unique view for this approach is that all the evidences can also be designed by statistical hypothesis tests, thus it can also be easy to be extended with other evidences to the detect ranking fraud. Finally, we here validate the system with extensive experiments on a real-world datas collected from the App store. Experimental results showed the effectiveness of the proposed approach. In the future, we plan to study more effective fraud evidences and to analyze the latent relationship among the rating, review and the rankings. Moreover, we will also extend our ranking fraud detection approach with other mobile App related to services, such as mobile Apps recommendation, for enhancing the user experience.

References

1. Hengshu Zhu, Hui Xiong, Senior Member, IEEE, Yong Ge, and Enhong Chen, Senior Member, IEEE "Discovery of Ranking Fraud for Mobile Apps," IEEE Transactions On Knowledge And Data Engineering, Vol. 27, No. 1, January 2015
2. Pranjali Deshmukh, Pankaj Agarkar "Mobile Application for Malware Detection," International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 02 Issue: 02 May-2015 www.irjet.net
3. Anuja A. Kadam, Pushpanjali M. Chouragade "A Review Paper on: Malicious Application Detection in Android System," International Journal of Computer Applications (0975 – 8887) National Conference on Recent Trends in Computer Science & Engineering (MEDHA 2015

Cite this article as: Maria Varsha B, Kayalvizhi K, Hemalatha A, Loganya K. "Survey of Developing a Model of Analysing and Detecting for Leaderboard Discrepancies of Mobile Application". *International Conference on Systems, Science, Control, Communication, Engineering and Technology 2016*: 509-511. Print.

4. Muneer Ahmad Dar & Javed Parvez, University of Kashmir, India "Evaluating Smartphone Application Security: A Case Study on Android," Global Journal of Computer Science and Technology Network, Web & Security Volume 13 Issue 12 Version 1.0 Year 2013 Type: Double Blind Peer Reviewed International Research Journal Publisher: Global Journals Inc. (USA) Online ISSN: 0975-4172 & Print ISSN: 0975-4350.
5. D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," J. Mach. Learn. Res., pp. 993–1022, 2003.
6. Y. Ge, H. Xiong, C. Liu, and Z.-H. Zhou, "A taxi driving fraud detection system," in Proc. IEEE 11th Int. Conf. Data Mining, 2011, pp. 181–190.
7. D. F. Gleich and L.h. Lim, "Rank aggregation via nuclear norm minimization," in Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2011, pp. 60–68.
8. T. L. Griffiths and M. Steyvers, "Finding scientific topics," Proc. Nat. Acad. Sci. USA, vol. 101, pp. 5228–5235, 2004.
9. A. Klementiev, D. Roth, and K. Small, "Unsupervised rank aggregation with distance-based models," in Proc. 25th Int. Conf. Mach. Learn., 2008, pp. 472–479.
10. A. Klementiev, D. Roth, K. Small, and I. Titov, "Unsupervisedrank aggregation with domain-specific expertise," in Proc. 21st Int. Joint Conf. Artif. Intell, 2009, pp. 1101–1106.

Cite this article as: Maria Varsha B, Kayalvizhi K, Hemalatha A, Loganya K. "Survey of Developing a Model of Analysing and Detecting for Leaderboard Discrepancies of Mobile Application". *International Conference on Systems, Science, Control, Communication, Engineering and Technology 2016*: 509-511. Print.