International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 18 (2015) pp 38690-38692
© Research India Publications. http://www.ripublication.com

Thor-Fd: The Hierarchial Online Ranking Fraud Detecting For Mobile Apps Using Sentiword Dictionary

Tamilmani.G

Assistant professor, Department of Computer Science and Engineering,
Vel Tech University, Avadi, Chennai – 62 taman.maya@gmail.com

K.Rajathi

Assistant Professor, Department of Computer Science and Engineering
Vel Tech University, Avadi, Chennai – 62 k.rajathimtech@gmail.com

K.Antony kumar,

Assistant Professor, Department of Computer Science and Engineering
Vel Tech University, Avadi, Chennai – 62 antonykmr32@gmail.com

Durai.S

Assistant Professor, Department of Computer Science and Engineering
Vel Tech University, Avadi, Chennai – 62 duraitrichy@gmail.com

Abstract
The primary aim of this project is to enhance the prevention of ranking frauds in mobile apps using the MAC address. In the existing system the leading event and leading session of an app is identified from the collected historical records. Then three different types of fact data are collected from the user feedbacks namely ranking based fact data, rating based fact data and review based fact data. These three fact data are aggregated by using fact data aggregation method. In the proposed system additionally, we are proposing two enhancements. Firstly, we are using Sentiword dictionary to identify the exact reviews scores. Secondly, the fake feedbacks by a same person for pushing up that app on the leader board are restricted. Two different constraints are considered for accepting the feedback given to an application. The first constraint is that an app can be rated only once from a user login. And the second is implemented with the aid of MAC address that limits the number of user login logged per day from a MAC address as five.

Keywords: Sentiword Dictionary, Sensor Node, Ranking Fraud Detection, MAC Address

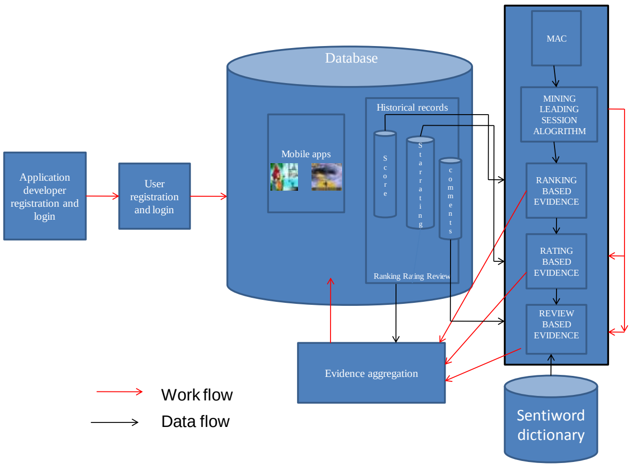
Introduction

Grade racketing in the mobile applications meant to deceitful or deceptive behavior that have a reason to knock up the application that has high popularity. Certainly, that becomes a popular among the app developers and the app marketers to perform some phony activity in the online market. The kind of activities may range from driving up their app's sales review or the by posting boosted up ratings to purposefully do ranking or grade racketing. Though the significance of the grade ranking frauds has been widely acknowledged. But the research work across in this field is limited. To overcome this factor we have provided a detailed view over the grading fraud and we have proposed a grade fraud detection especially for the mobile applications. The initial step of the proposed work lies on working and finding out the grading frauds by digging out the period of activity of the application. The activity is nothing but of the sessions that are in lead. Such lead session can be controlled for the local irregularity instead of going to larger irregularity of the application ranking system. Adding to the above facts, we probe another few types of face data. They are rank based data, rating based data, and review based data. By implementing the statistical data analysis of the taken app's rank, rate and reviews the conclusion shall be made. Once all the face data have been collected the optimized aggregation will be performed to integrate the face data.

SYSTEM ANALYSIS

Many mobile app stores launched daily app leader boards which shows the chart ranking of popular apps. The leader board is the important for promoting apps. Original application grade level decreases due to the duplication arrival in the mobile apps. In recent activities duplicate version of an application not burned or blocked. This is the major defect. Higher rank leads huge number of downloads and the app developer will get more profit.In this they Allow Fake Application also.User not understanding the Fake Apps then the user also give the reviews in the fake application.Exact Review or Ratings or Ranking

Percentage are not correctly Calculated.



IMPLEMENTATION

Module Descriptions

Mobile Apps & Historical Ranking Records:

The rank fraud in the mobile application environment mainly works over boosting up the app in the popularity chart. The main use of this fraud is to post falsified sales rating so that they can gain popularity across the app store that might results in large number of downloads. These phony frauds are happening widely across the network and there are certain measures are being taken to prevent this. Eventually the current preventive methods are limited to some extent and the research in this sector is growing at a slow pace.

The grading fraud usually happens in the primary sessions of the mobile application environments. So detecting meant finding the grading frauds in the primary session. The proposed solution is simple but efficient algorithm to recognize the primary sessions of the all the applications based on the history of grading. After analyzing the grading behavior of the various apps we can find out the deceitful ap. A simple logic applied here is by analyzing the grading pattern of the various apps in the primary sessions and the normal app.

Ranking Fraud Detection:

The holistic method of viewing the grading fraud and proposing a grade ranking detection system for the mobile app. Specifically, we first propose to accurately locate the grading fraud by mining the active periods, namely primary sessions, of mobile Apps. Such primary sessions can be boosted up to find the local deviation than going for the global deviation of the application ranking.

Rating and Review:

The proposed solution works on two patterns namely, app rating and history of the app review. Each will continue to show some negative patterns from the rating history and the record reviews. Adding to this pattern we also developed an aggregation of fact data method to adjoin all the collected fact data for evaluating the primary session credibility from the apps. Figure 1 explains the framework of Grading fraud detecting system for mobile Applications. It is good to see that all the fact data are collected using the modeling f the application grading and behavioral study on the review obtained through the suggestive study from the user in terms of review and star gradings. The proposed solution is scalable and can be compared using the other segment or network generated data facts in term of grading frauds.

Update Apps Details:

Prior to the development of the detecting system the identification of the fraud happening is the most important. The fraud will be taken place mostly in the primary sessions and then the face data will be mined from the same using the history of ranking. Then that data will be collected as the real time and can be categorized as rating based, grading based and review based.

Aggregation of Fact Data:

The novel fact data aggregation method which uses an automatic unsupervised method. It allows us to integrate all the collected fact data from various sources such as rating, grading and review for the further evaluation of the apps primary session credibility.

MAC ADDRESS:

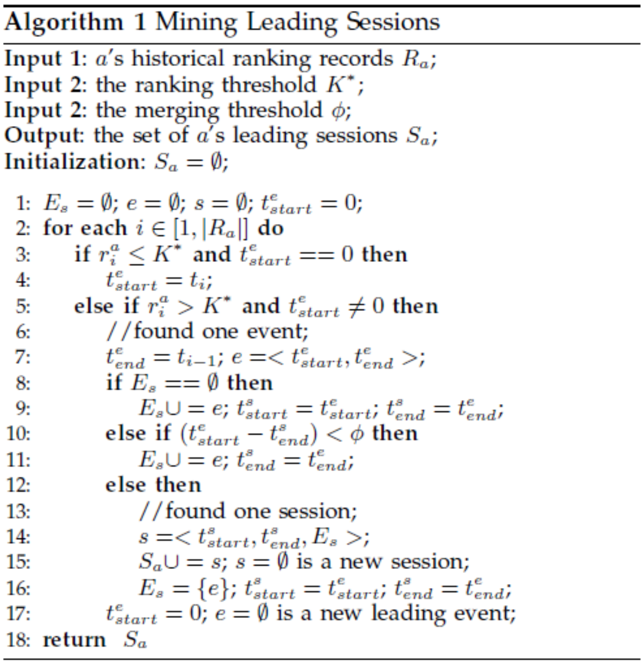
If the user gives ranking and rating many times for an app, then it will be identified by the admin using the MAC address. The user can not give more than five ranking or rating for an app a day from one MAC. MAC address cannot be changed. Using MAC, IP spoofing attack can be blocked.

Algorithm:

Mining Primary Sessions:

The steps involve in mining the primary session they are, first the lead event has to be discovered from the historical grading records from the application. Secondly, the lead events of the adjacent methods have to be constructed for the primary sessions. Algorithm 1 explains the alias code of mining lead sessions for the sample app.

Complaints of an original version of application provider can be undertaken by using Mining Leading Session algorithm. The duplicate version is identified by the admin by means of Historical Records. The admin will also see the date of publication of the apps. When the apps is detected as fraudulently published by the admin then the respective app will be blocked. The user can give the feedback at only once. Sentiword dictionary is used for finding the exact reviews. The admin can block the fake application. The Review or Rating or Ranking given by users are Correctly Calculated. Hence, a new user who wants to download an app for some purpose can get clear view about the available applications.



Algorithm 1 Mining Leading Sessions

```
Input 1: $a$'s historical ranking records $R_a$;
Input 2: the ranking threshold $K^*$;
Input 2: the merging threshold $\phi$;
Output: the set of $a$'s leading sessions $S_a$;
Initialization: $S_a = \emptyset$;

$E_s = \emptyset$; $e = \emptyset$; $s = \emptyset$; $t^e_{start} = 0$;
for each $i \in [1, |R_a|]$ do
    if $r^a_i \leq K^*$ and $t^e_{start} == 0$ then
        $t^e_{start} = t_i$;
    else if $r^a_i > K^*$ and $t^e_{start} \neq 0$ then
        //found one event;
        $t^e_{end} = t_{i-1}$; $e =< t^e_{start}, t^e_{end} >$;
        if $E_s == \emptyset$ then
            $E_s \cup = e$; $t^s_{start} = t^e_{start}$; $t^s_{end} = t^e_{end}$;
        else if $(t^e_{start} - t^s_{end}) < \phi$ then
            $E_s \cup = e$; $t^s_{end} = t^e_{end}$;
        else then
            //found one session;
            $s =< t^s_{start}, t^s_{end}, E_s >$;
            $S_a \cup = s$; $s = \emptyset$ is a new session;
            $E_s = \{e\}$; $t^s_{start} = t^e_{start}$; $t^s_{end} = t^e_{end}$;
        $t^e_{start} = 0$; $e = \emptyset$ is a new leading event;
return $S_a$
```

Conclusion

Complaints of an original version of application provider can be undertaken by using Mining Leading Session algorithm. The duplicate version is identified by the admin by means of Historical Records. The admin will also see the date of publication of the apps. When the apps is detected as fraudulently published by the admin then the respective app will be blocked. The user can give the feedback at only once. Sentiword dictionary is used for finding the exact reviews. The admin can block the fake application. The Review or Rating or Ranking given by users are Correctly Calculated. Hence, a new user who wants to download an app for some purpose can get clear view about the available applications.

References

[1] L.Azzopardi, M. Girolami, and K. V. Risjbergen. Investigating the relationship between language model perplexity and irprecisionrecall measures. *In Proceedings of the 26th International Conference on Research and Development in Information Retrieval (SIGIR'03)*, pages 369–370, 2003.

[2] D. M. Blei, A. Y. Ng, and M. I. Jordan. Lantentdirichlet allocation. *Journal of Machine Learning Research*, pages 993–1022, 2003.

[3] Y. Ge, H. Xiong, C. Liu, and Z.-H. Zhou. A taxi driving fraud detection system. In *Proceedings of the 2011 IEEE 11th International Conference on Data Mining*, ICDM '11, pages 181–190, 2011.

[4] Opinion spam and analysis. In Proceedings of the 2008 International Conference on Web Search and Data Mining, WSDM '08, pages 219–230, 2008.

[5] P. Dangauthier, R. Herbrich, T. Minka, and T. Graepel. TrueSkill through time: Revisiting the history of chess. In Proceedings of the Neural Information Processing Systems, 2007.

[6] T. Bao, H. Cao, E. Chen, J. Tian, and H. Xiong. An unsupervised approach to modeling personalized contexts of mobile users. In ICDM'10, pages 38–47, 2010.

[7] D. M. Blei, A. Y. Ng, and M. I. Jordan. Lantent dirichlet allocation. In Journal of Machine Learning Research, pages 993–1022, 2003.

[8] H. Cao, T. Bao, Q. Yang, E. Chen, and J. Tian. An effective approach for mining mobile user habits. In CIKM'10, pages 1677–1680, 2010.

[9] T. L. Griffiths and M. Steyvers. Finding scientific topics.In Proceedings of National Academy of Science of the USA, pages 5228–5235, 2004.

[10] R. Kwapisz, Gary M. Weiss, and Samuel A. Moore. 2011. Activity recognition using cell phone accelerometers. SIGKDD Explor. Newsl. 12 (2011), 74–82. Issue 2.