International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

IJIEMR Transactions, online available on 28th Nov 2017. Link :http://www.ijiemr.org/downloads.php?vol=Volume-6&issue=ISSUE-11

Title: DETECTION AND ELIMINATION OF VULNERABILITY BEHAVIOR USING RATING AND REVIEW BASED AGGREGATION MINING METHOD

Volume 06, Issue 11, Pages: 387–391.

Paper Authors

JADI VASANTHA
VIF College Of Engineering And Technology

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

DETECTION AND ELIMINATION OF VULNERABILITY BEHAVIOR USING RATING AND REVIEW BASED AGGREGATION MINING METHOD

JADI VASANTHA
Assistant Professor, VIF College Of Engineering And Technology

Abstract:
Present day's smart phone became part of everyone life. There is need of various applications to be installed on smart Phone for the daily use. To download application smart phone user has to visit Apps store such as Google Play Store, Apples store etc. which is the major target of fraud applications. The detection and removal of these apps from android is the major problem in now days. In mobile Apps business ranking fraud alludes to false or tricky operations which have a purpose behind knocking up the Apps in the leader board chart. To be sure, it turns out to be more continuous for App developers to utilize doubtful means, for expanding their mobile Apps' business. The starting aim of this project is to enhance the prevention of ranking frauds in mobile apps. In this work the leading event and group of neighboring events that is leading session of an app is identified from the collected historical records of mobile Apps. Then three different types of evidences are collected from the user feedbacks like comments namely ranking, rating and review based. These three evidences are aggregated by using evidence aggregation method. The output of aggregation is the mobile app which decides the app is false or not. At last, we assess the proposed framework with certifiable App information gathered from the App Storeroom for quite a while interval and also ranking fraud detection method with different services related to Apps such as recommendation of Apps for user that is to preventing false apps to be recommended to user, to learning more powerful fraud evidences and latent relation analysis on reviews.

Keywords: Rank Based evidence, Ranking Fraud, Mobile Applications, Evidence Aggregation

Introduction:

Ranking fraud in the mobile app market refers to fraudulent or deceptive activities which have a purpose of bumping up the apps in the popularity 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 paper, we provide a holistic view of ranking fraud and propose a ranking fraud detection system for mobile apps. Specifically, we first propose 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. Furthermore, we investigate three types of evidences, i.e., ranking based evidences, rating based evidences and review based evidences, by modeling apps' ranking, rating and review behaviors through statistical hypotheses tests. In Rating 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. An App which has higher rating may attract more users to download and can also be ranked higher in the leader board. Thus, rating manipulation is also an important perspective of ranking fraud. In Review Based Evidences, besides ratings, most of the App stores also allow users to write some textual comments as App reviews. Especially, this paper proposes a simple and effective algorithm to recognize the leading sessions of each mobile App based on its historical ranking records. This is one of the fraud evidence. Also, rating and review history, which gives some anomaly patterns from apps historical rating and reviews records.

Evidence aggregation methodology implementation:

In this we are going to use *Leading Sessions algorithm.* In the leading session algorithm we are have to apply data mining technique on 3 clusters.

1. Review based cluster (comments)
2. Rating based cluster
3. Ranking based cluster

In the **review based** clustering we will maintain the leading session for the every user which mean we have set the threshold value to identify whether comment given by the fake user or not. Here threshold value is time (leading session time).For example user entered one comment for the app. If we are getting the comment from same system i.e., based on Mac or ip address within the leading session time (eg: 10 min) then we will consider that comment as a fake comment. After getting the both fake and authorized review we will apply **Naive Bayes** data mining algorithm to count positive and negative feedbacks. This is the Review based evidence.

For the **ranking based** evidence, based on the number of downloads we will provide the ranking for the app.

For the **Rating** we will have two type users to give rating local Anomaly and Global Anomaly (Registered user). For the two type users we will show statistical analysis.

Literature Survey:

- A flexible generative model for preference aggregation

Many areas of study, such as information retrieval, collaborative filtering, and social choice face the preference aggregation problem, in which multiple preferences over objects must be combined into a consensus ranking. Preferences over items can be expressed in a variety of forms, which makes the aggregation problem difficult. In this work we formulate a flexible probabilistic model over pairwise comparisons that can accommodate all these forms. Inference in the model is very fast, making it applicable to problems with

hundreds of thousands of preferences. Experiments on benchmark datasets demonstrate superior performance to existing methods .

- Getjar mobile application recommendations with very sparse datasets

The Netflix competition of 2006 has spurred significant activity in the commendations field, particularly in approaches using latent factor models However, the near ubiquity of the Netflix and the similar MovieLens datasets1 may be narrowing the generality of lessons learned in this field. At GetJar, our goal is to make appealing recommendations of mobile applications (apps). For app usage, we observe a distribution that has higher kurtosis (heavier head and longer tail) than that for the aforementioned movie datasets. This happens primarily because of the large disparity in resources available to app developers and the low cost of app publication relative to movies. In this paper we compare a latent factor (PureSVD) and a memory-based model with our novel PCA-based model, which we call Eigenapp. We use both accuracy and variety as evaluation metrics. PureSVD did not perform well due to its reliance on explicit feedback such as ratings, which we do not have. Memory-based approaches that perform vector operations in the original high dimensional space over-predict popular apps because they fail to capture the neighborhood of less popular apps. They have high accuracy due to the concentration of mass in the head, but did poorly in terms of variety of apps exposed. Eigenapp, which exploits neighborhood information in low dimensional spaces, did well both on precision and variety, underscoring the importance of dimensionality reduction to form quality neighborhoods in high kurtosis distributions.

- Detecting spam web pages through content analysis

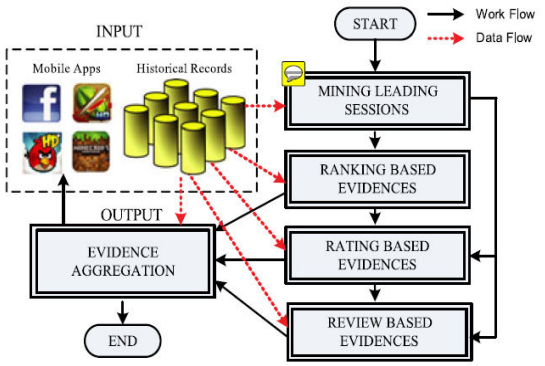
In this paper, we continue our investigations of "web spam": the injection of artificially-created pages into the web in order to influence the results from search engines, to drive traffic to certain pages for fun or profit. This paper considers some previously-undescribed techniques for automatically detecting spam pages, examines the effectiveness of these techniques in isolation and when aggregated using classification algorithms. When combined, our heuristics correctly identify 2,037 (86.2%) of the 2,364 spam pages (13.8%) in our judged collection of 17,168 pages, while misidentifying 526 spam and non-spam pages (3.1%).

Evidence Aggregation Architecture:

After extracting three types of fraud evidences the next challenge is how to combine them for ranking fraud detection. In addition, there are many methods of ranking and evidence aggregation in the literature, such as permutation based models score based models and Dempster Shafer rules. However, some of these methods focus on learning a global ranking for all applicants. This way is not proper for detecting ranking fraud for new Apps. Distinct methods are based on supervised learning techniques, which rely on the labelled training data and are hard to be exploited. Rather, we suggest an

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

unsupervised approach based on fraud similarity to combine these evidences. 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. Thus, we characterize some fraud evidences from Apps' historical ranking records, and develop three functions to extract such ranking based fraud evidences.



effectiveness of the proposed approach.

Conclusion:

In this paper, we develop ranking fraud detection system for mobile apps. It reviews various existing strategies used for internet or web spam detection, which is associated with the rating fraud for mobile Apps. Also, we've seen references for online review unsolicited mail detection and mobile App advice. By using mining the main sessions of mobile Apps, we aim to locate the ranking fraud. The leading classes works for detecting the nearby anomaly of App ratings. The machine targets to locate the ranking frauds based on three styles of evidences, including rating based evidences, ranking based evidences and comment based evidences. In addition, an optimization based totally aggregation method combines all of the three evidences to hit upon the fraud. Ultimately, we validate the proposed system with extensive experiments on real world App data collected from the google play store.

References:

- D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," J. Mach. Learn. Res., pp. 993–1022, 2003.
- Y. Ge, H. Xiong, C. Liu, and Z.-H. Zhou, "A taxi driving fraud detection system," in Proc. IEEE 11th Int. Conf. Data Mining, pp. 181– 190, 2011.
- D. F.Gleich and L.-h. Lim, "Rank aggregation via nuclear norm minimization," in Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, pp. 60–68, 2011. T. L. Griffiths and M. Steyvers, "Finding scientific topics," Proc
- . Nat. Acad. Sci. USA, vol. 101, pp. 5228–5235, 2004.
- G. Heinrich,Parameter estimation for text analysis, " Univ. Leipzig, Leipzig, Germany, Tech. Rep., http://faculty.cs.byu.edu/~ringger/ CS601R/papers/Heinrich-GibbsLDA.pdf, 2008.

- 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.

JADI VASANTHA
vasantha803@gmail.com
Qualification:m.tech-2009 2011(computer science and engineering) Residential address: flat no:508, rank one towers, langer house,mehdipatnam,hyd-500008.
Working as a assistant professor at VIF College of engineering and technology