ISSN 2395-1621

A Survey on: Mechanism of Ranking Fraud for Mobile Apps Applications

[#1]Rakesh Saini, [#2]Juber Shaikh, [#3]Harshal Shirole, [#4]Vilas Pawar

[1]rakeshkumarsaini0011@gmail.com,
[2]jubers1111@gmail.com,
[3]harshalsshirole@gmail.com
[4]vilaspawar0411@gmail.com

[#1234]Department of Computer Engineering,

JSPM's, ICOER, Wagholi.

ABSTRACT

In This project gives a whole perspective of positioning misrepresentation and describes a Ranking fraud identification framework for mobile Apps. This work is grouped into three category. First is web ranking spam detection, second is user review spam detection and last one is mobile app recommendation. The Web ranking spam refers to any deliberate actions which bring to selected Web pages an unjustifiable favorable relevance or importance. Review spam is designed to give unfair view of some products so as to influence the consumers' perception of the products by directly or indirectly influating or damaging the product's reputation.

Keyword: Mobile Apps, Ranking fraud detection, Evidence aggregation, Historical ranking records, Rating and Review.

ARTICLE INFO

Article History

Received: 1st November 2016

Received in revised form :

1st November 2016

Accepted: 3rd November 2016

Published online :

4th November 2016

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 leader board is one of the most important ways for promoting mobile Apps. A higher rank on the leader board 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 leader boards.

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.

II. RELATED WORK

This paper aims to detect users generating spam reviews or review spammers. We identify several characteristic behaviors of review spammers and model these behaviors so as to detect the spammers. In particular, we seek to model the following behaviors. First, spammers may target specific products or product groups in order to

maximize their impact. Second, they tend to deviate from the other reviewer in their ratings of products.

We finally show that the detected spammers have more significant impact on ratings compared with the unhelpful reviewers.
From this paper we have referred:-

- Concept of extracting of rating and ranking.
- Concept of extracting of review.

III. LITERATURE SURVEY

[1] B. Zhou, J. Pei, and Z. Tang. "A spamicity approach to web spam detection."2008, in this paper technique used: Link spam city, advantages of this system is do not need training, and finally get the result effective and efficient to detect spam pages.

[2] A. Ntoulas, M. Najork, M. Manasse, and D. Fetterly. "Detecting spam web pages through content analysis.",2006, technique used: Content based spam detection algorithm. Advantages: Classifier can correctly identify 86.2% of all spam pages, Result: detecting content based spam.

[3] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. "Lauw. Detecting product review spammers using rating behaviors." 2010, technique used: Rating behavioral approach to detect review spammers, advantages: Detect users generating spam reviews or review spammers, Result: Show that the detected spammers have more significant impact on ratings compared with the unhelpful reviewers.

[4] Z.Wu, J.Wu, J. Cao, and D. Tao. Hysad: "a semisupervised hybrid shilling attack detector for trustworthy product recommendation." 2012, technique used: Hybrid Shilling Attack Detector, advantages: precisely separate Random-Filler model attackers and effective against hybrid attacks, result: Effectively improve the accuracy of a collaborative-filtering based recommender system.

[5] S. Xie, G. Wang, S. Lin, and P. S. Yu. "Review spam detection via temporal pattern discovery." 2012, technique used: Hierarchical algorithm, advantages: Effective in detecting singleton review attacks, result: Robustly detect the time windows where such attacks are likely to have happened.

Xie etal. [6] have studied the problem ofsingleton review spam detection. Specifically, they solved this problem by detecting the co-anomaly patterns inmultiple review based time series. Although some ofabove approaches canbeused for anomaly detection from historical rating and review records, they are not able toextract fraud evidences for a given time period (i.e., leading session).

Finally, the third category includes the studies on mobile App recommendation. For example, Yan etal. [7] developed a mobile App recommender system, named App joy, 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 etal. [8] studied several recommendation models and proposed a content based collaborative filtering model, named Eigen app, for recommending Apps in their Web site Getjar. In addition, some researchers studied the problem of exploiting enriched contextual information for mobile App recommendation.

IV.DATA MINING TECHNIQUE

Partition-based clustering:
It is centroid based clustering in which data points splits into k partition and each partition represents a cluster. Different methods of partitioning clustering are k-means, bisecting k-means method, Medoids method, Portioning Around Medoids (PAM), CLARA (Clustering Large Applications) and the Probabilistic centroid.

K-means clustering:
K-means clustering technique is a technique of clustering which is widely used. This algorithm is the most popular clustering tool that is used in scientific and industrial applications. It is a method of cluster analysis which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean.

K-Nearest Neighbor:
K-Nearest Neighbor algorithm that is being widely used for classification and regression and also it is a non-parametric method. Every training set that is being present in the multidimensional feature space are the vectors with the specific class labels specified.

Support Vector Machines:
SVM was introduced by Boser, Guyon and Vapnik and widely being used for classification, regression and pattern recognition.SVM has capability to classify indeed

of the dimensions or size of the input space. It acquires the major advantage because of its high generalization performance with indeed of the much prior knowledge. The goal of the SVM lies in finding the best classification function and also it aims to distinguish between members of the two classes in training data.

Naive Bays classification:
The naive bays classifiers are a family of simple probabilistic classifiers based on applying bays theorem with strong independence assumptions between the features. Bayesian classifier is working on the dependent events and the probability of an event occurring in the future that can be detected from the previous occurring of the same event. The naive bays classifier is a simple statistical algorithm providing amazingly better results.

Here we user the Naive Bays classification mining technique for detecting fraud mobile ranking application. Naïve bays is gives better output from other mining technique.

Advantages of naïve bays:

- To improves the classification performance by removing the irrelevant Features.
- Good performance.
- It is short computational time.

V. CONCLUSION

In this paper, we study a ranking fraud detection system for mobile Apps. Specifically, we first showed that ranking fraud happened in leading sessions and provided a method for mining leading sessions for each App from its historical ranking records. Then, we identified ranking based evidences and rating based evidences for detecting ranking fraud. Moreover, we proposed an optimization based aggregation method to integrate all the evidences for evaluating the credibility of leading sessions from mobile Apps.

REFERENCES

[1] 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.

[2] A. Ntoulas, M. Najork, M. Manasse, and D. Fetterly. Detecting spam web pages through content analysis. In Proceedings of the 15th international conference on World Wide Web, WWW '06, pages 83–92, 2006.

[3] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw. Detecting product review spammers using rating behaviors. In Proceedings of the 19th ACM international conference on Information and knowledge management, CIKM '10, pages 939–948, 2010.

[4] Z.Wu, J.Wu, J. Cao, and D. Tao. Hysad: a semisupervised hybrid shilling attack detector for trustworthy product recommendation. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '12, pages 985–993, 2012.

[5] S. Xie, G. Wang, S. Lin, and P. S. Yu. Review spam detection via temporal pattern discovery. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '12, pages 823–831, 2012.

[6] S.Xie, G.Wang, S.Lin, and P.S.Yu. Review spam detection via temporal pattern discovery. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD'12,pages 823–831, 2012.

[7] B.Yanand G.Chen. Appjoy: personalized mobile application discovery. In Proceedings ofthe 9th international conference on Mobile systems, applications, andservices, MobiSys '11,pages 113–126,2011.

[8] K.Shi and K.Ali. Getjar mobile application recommendations with very sparse datasets. In Proceedings ofthe 18th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD'12,pages 204–212, 2012.

[9] R.Agrawal and R.Srikant, "Fastalgorithms for mining association rules,"inVLDB, 1994.

[10]H.Zhu, E.Chen, K.Yu, H.Cao, H.Xiong, and J.Tian. Mining personal context-aware preferences for mobile users. In Data Mining (ICDM), 2012 IEEE12th International Conference on, pages 1212–1217, 2012.

[11]Hengshu Zhu, Hui XiongDiscovery ofRanking Fraud for Mobile Apps. IEEETRANSACTIONS ON KNOWLEDGE ANDDATA ENGINEERING,2013.