INTERNATIONAL ONLINE CONFERENCE ON ADVANCE RESEARCH & DEVELOPMENT IN ENGINEERING AND TECHNOLOGY

IO-CARDET 2016

E-ISSN NO:-2349-0721

Impact factor : 2.25

MOBILE APP RECOMMENDATION & RANKING FRAUD DETECTION AMONG RATING & RANKING

Neha S. Hete
Agnihotri college of Engineering, Wardha
nehahete@gmail.com

Prof. Dhananjay Sable
Agnihotri college of Engineering Wardha
dhananjay.sable165@gmail.com

ABSTRACT:

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 addition, we propose an optimization based aggregation method to integrate all the evidences for fraud detection .The mobile app recommendation for Finally, we evaluate the proposed system with real-world App data collected from the iOS App Store for a long time period. In the experiments, we validate the effectiveness of the proposed system, and show the scalability of the detection algorithm as well as some regularity of ranking fraud activities.

Keywords: - Mobile Apps, Ranking Fraud Detection, Evidence Aggregation, Historical Ranking Records, Rating and Review, Recommendation apps, KNN algorithm.

I. INTRODUCTION

The number of mobile Apps has grown at a breath taking ate over the past few years. For example, as of the end of April 2013, there are more than 1.6 million Apps at Apple's App store and Google Play. To stimulate 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. However, as a recent trend, instead of relying on traditional marketing solutions, shady App developers resort to some fraudulent means to deliberately boost their Apps and eventually manipulate the chart rankings on an App store. This

is usually implemented by using so-called "bot farms" or "human water armies" to inflate the App downloads, ratings and reviews in a very short time. For example, an article from Venture Beat reported that, 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 leader board and more than 50,000-100,000 new users could be acquired within a couple of days. In fact, such ranking fraud raises great concerns to the mobile App industry. For example, Apple has warned of cracking down on App developers who commit ranking fraud in the Apple's App store.

Ranking fraud in the mobile App market refers to fraudulent or deceptive activities which have a purpose of bumping up 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 addition, we propose an optimization based aggregation method to integrate all the evidences for fraud detection. Finally, we evaluate the proposed system with real-world App data collected from the iOS App Store for a long time period. In the experiments, we validate the effectiveness of the proposed system, and show the scalability of the detection algorithm as well as some regularity of ranking fraud activities.

II. LITERATURE REVIEW

In this paper, developed a ranking fraud detection system for mobile apps that ranking fraud happened in leading sessions for each app from its historical ranking records.[1] In this method, we address the problem of review spammer detection, or ding users who are the source of spam reviews. Unlike the approaches for spammed review detections, our proposed review spammer detection approach is user centric, and user behavior driven. A user centric approach is preferred over the review centric approach as gathering behavioral evidence of spammers is easier than that of spam reviews. A review involves only one reviewer and one product. The amount of evidence is limited. A reviewer on the other hand may have reviewed a number of products and hence has contributed a number of reviews. The likelihood of ending evidence against spammers will be much higher. The user centric approach is also scalable as one can always incorporate new spamming behaviors as they emerge.[2] In this paper we first give a general framework for conducting Supervised Rank Aggregation. We show that we can define supervised learning methods corresponding to the existing unsupervised methods, such as Board Count and Markov Chain based methods by exploiting the framework. Then we mainly investigate the supervised versions of Markov Chain based

methods in this paper, because previous work shows that their unsupervised counterparts are superior. It turns out, however, that the optimization problems for the Markov Chain based methods are hard, because they are not convex optimization problems. We are able to develop a method the optimization of one Markov Chain based method, called Supervised MC2.Specifically, we prove that we can transform the optimization problem into that of Semi definite Programming.[3] we first give a general framework for conducting Supervised Rank Aggregation. We show that we can define supervised learning methods corresponding to the existing unsupervised methods, such as Board Count and Markov Chain based methods by exploiting the framework. Then we mainly investigate the supervised versions of Markov Chain based methods in this paper, because previous work shows that their unsupervised counterparts are superior. It turns out, however, that the optimization problems for the Markov Chain based methods are hard, because they are not convex optimization problems. We are able to develop a method the optimization of one Markov Chain based method, called Supervised MC2.Specifically, we prove that we can transform the optimization problem into that of Semi definite Programming.[4] In this paper, creator displayed different sorts of conventions to safeguard the protection or security of the information. This paper mulled over the issue of vitality sparing in MANETs in view of the method of system coding and demonstrated that Network-Coding is productive in calculation, and acquires less vitality utilization for encryptions/ decoding s. [5] In this study, we used app-usage as our metric. Given the characteristics of this data, we found that traditional memory-based approaches heavily favor popular apps contrary to our mission. On the other hand, latent factor models that were developed based on the Netflix data performed quite poorly accuracy-wise. We find that the Eigenapp model performed the best in accuracy and in promotion of less well known apps in the tail of our dataset.[6]

III. PROPOSED TECHNIQUE

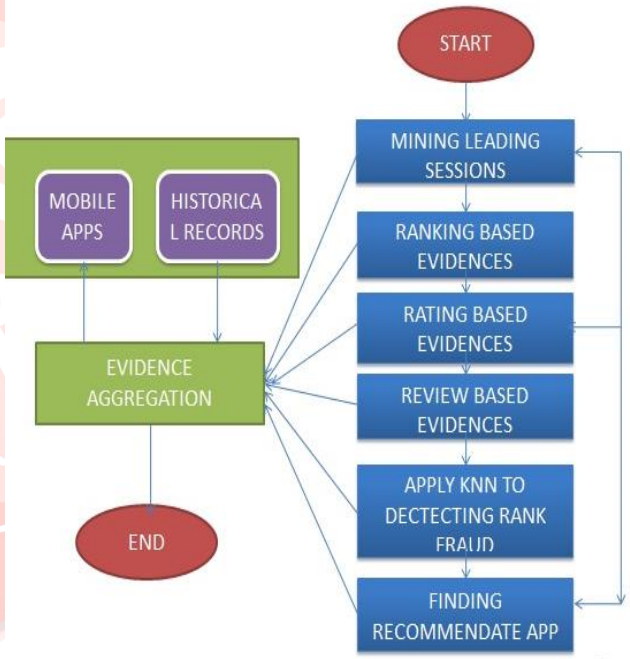


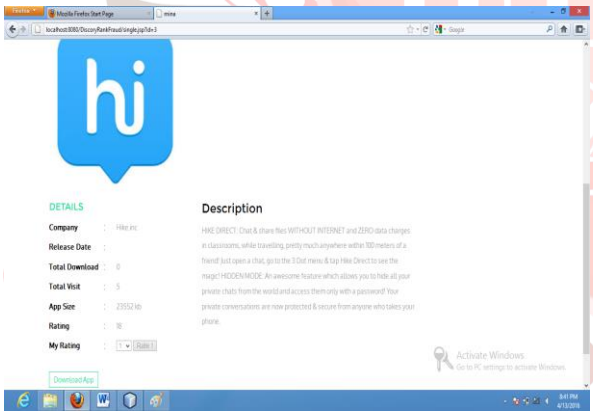
Fig 1: System Architecture

1. Mining leading sessions

First the mining leading sessions is used to discover leading events from the app's historical ranking records and then it merges adjacent leading events for constructing leading sessions .Then the ranking based evidence analyse the basic characteristics of leading events for extracting fraud evidences. The rating based evidence is used to rate by any user who downloaded it

2. Rating based Evidence

In rating based evidence, the user observes which app is leading from the sessions. If he wants to give high rank to the app which has low rate then he can give rate to this particular app in this way.



IV. CONCLUSION

In this paper we achieved the mining leading sessions and rating based evidence. This paper introduces more effective fraud evidences and analyze the latent relationship among rating ,review and rankings. We extended our ranking fraud detection approach with other mobile app related services ,such as mobile app recommendation for enhancing user experience.

REFERENCES

[1] Discovery of Ranking fraud for mobile apps. Hengshu Zhu,Hui Xiong,Senior members,IEEE,Yong Ge,and Enhong Chen,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 on World Wide Web.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] detection for mobile apps H. Zhu, H. Xiong, Y. Ge, and E. Chen. A holistic view. In Proceedings of the 22nd ACM international conference on Information and knowledge management,CIKM '13, 2013.

[9] 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.

[10] 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.