Discovery of Ranking Fraud for Mobile Apps

[1]Mr. Y.G.Tamboli, [2]Prof. P. A. Satarkar

Department of Computer Science & Engineering, SVERI COE, Solapur University, Solapur, India

Abstract*- Ranking fraud within the mobile App market refers to deceitful or deceptive activities that have a purpose of bumping up the Apps within the quality list. Indeed, it becomes a lot of and a lot of frequent for App developers to use shady means that, like inflating their Apps' sales or posting phony App ratings, to commit ranking fraud. Whereas the importance of preventing ranking fraud has been well known, there's restricted understanding and analysis during this space. to the current finish, during this paper, we offer a holistic read of ranking fraud and propose a ranking fraud detection system for mobile Apps. Specifically, we have a tendency to initial propose to accurately find the ranking fraud by mining the active periods, specifically leading sessions, of mobile Apps. Such leading sessions are often leveraged for sleuthing the native anomaly rather than international anomaly of App rankings. what is more, we have a tendency to investigate 3 kinds of evidences, i.e., ranking primarily based evidences, rating based evidences and review based evidences, by modeling Apps' ranking, rating and review behaviors through applied math hypotheses tests.*

KEYWORDS- Ranking based evidences,, Rating based evidences, Review based evidences, Evidence Aggregation, Leading sessions.

1. INTRODUCTION

The number of mobile Apps has grown up at a panoramic rate over the past few years. for instance, as of the top of Apr 2013, there ar quite 1.6 million Apps at Apple's App store and Google Play. To stimulate the event of mobile Apps, several App stores launched daily App leader boards that demonstrate the chart rankings of preferred Apps. Indeed, the App leaderboard is one in all the foremost necessary ways that for promoting mobile Apps. A better rank on the leaderboard typically results in an enormous variety of downloads and million bucks in revenue. Therefore, App developers tend to explore varied ways that like advertising campaigns to push their Apps so as to own their Apps hierarchal as high as potential in such App leaderboards.

However, as a recent trend, rather than hoping on ancient selling solutions, shady App developers resort to some deceitful means that to deliberately boost their Apps Associate in Nursingd eventually manipulate the chart rankings on an App store. this is often typically enforced by mistreatment questionable "bot farms" or "human water armies" to inflate the App downloads, ratings and reviews in a very short time. for instance, a piece of writing from Venture Beat according that, once Associate in Nursing App was promoted with the assistance of ranking manipulation, it may well be propelled from number one, 800 to twenty five in Apple's top free leaderboard and quite 50000-100,000 new users may well be no inheritable inside some of days. In fact, such ranking fraud raises nice considerations to the mobile App trade. For instance, Apple has warned of cracking down on App developers United Nations agency commit ranking fraud within the Apple's App store. In the literature, whereas there are some connected work, like net ranking spam detection, on-line review spam detection and mobile App recommendation, the matter of detection ranking fraud for mobile Apps remains under-explored. To fill this important void, during this paper, we have a tendency to propose to develop a ranking fraud detection system for mobile Apps. on this line,we determine many vital challenges. First, ranking fraud doesn't continually happen within the whole life cycle of AN App, therefore we want to find the time once fraud happens. Such challenges are often thought to be detection the native anomaly rather than world anomaly of mobile Apps. Second, because of the massive variety of mobile Apps, it's troublesome to manually label ranking fraud for every App, therefore it's vital to own a scalable thanks to mechanically find ranking fraud while not victimization any benchmark data. Finally, because of the dynamic nature of chart rankings, it's harsh to spot and ensure the evidences coupled to ranking fraud, that motivates United States of America to find some implicit fraud patterns of mobile Apps as evidences.

2. LITERATURE SURVEY

2.1 Existing System

Generally speaking, the connected works of this study will be sorted into 3 classes.

1. The first class is regarding internet ranking spam detection. Specifically, the net ranking spam refers to associatey deliberate actions that bring around elect websites an unwarrantable favorable connexion or importance. for instance, Ntoulas et al. [1] have studied numerous aspects of content-based spam on the net and conferred variety of heuristic ways for detective work content primarily based spam. Zhou et al. [2] have studied the matter of unsupervised internet ranking spam detection. Specifically, they planned associate economical on-line link spam and term spam detection ways victimization spamicity. Recently, Spirin and Han dynasty [3] have reportable a survey on internet spam detection, that

comprehensively introduces the principles and algorithms within the literature. Indeed, the work of internet ranking spam detection is principally supported the analysis of ranking principles of search engines, like PageRank and question term frequency. this can be totally different from ranking fraud detection for mobile Apps.

2. The second class is concentrated on sleuthing on-line review spam. as an example, Lim et al. [4] have known many representative behaviors of review spammers and model these behaviors to find the spammers. Wu et al. [5] have studied the matter of sleuthing hybrid shilling attacks on rating information. The planned approach relies on the semi supervised learning and might be used for trustworthy product recommendation. Xie et al. [6] have studied the matter of singleton review spam detection. Specifically, they solved this drawback by sleuthing the co-anomaly patterns in multiple review based mostly statistic. though a number of higher than approaches is used for anomaly detection from historical rating and review records, they're not capable to extract fraud evidences for a given fundamental measure (i.e., leading session).

3. Finally, the third class includes the studies on mobile App recommendation. as an example, Yan and Chen [7] developed a mobile App recommender system, named Appjoy, that relies on user's App usage records to create a preference matrix rather than victimization specific user ratings. Also, to resolve the poorness downside of App usage records, Shi and Ali [8] studied many recommendation models and planned a content primarily based cooperative filtering model, named Eigenapp, for recommending Apps in their web site Getjar.

In addition, some researchers studied the matter of exploiting enriched discourse info for mobile App recommendation. as an example, Zhu et al. [9] planned a consistent framework for customized context-aware recommendation, which might integrate each context independence and dependency assumptions. However, to the most effective of our data, none of previous works has studied the matter of ranking fraud detection for mobile Apps.

2.2 Proposed System

OBJECTIVES

The key objectives of current work are as given below:

1. To provide a holistic view of ranking fraud and propose a ranking fraud detection system for mobile Apps.

2. To improve the fraud detection efficiency

3. METHODOLOGY

Methodology with relevancy every objective is concisely given below:

1. To provide a holistic view of ranking fraud and propose a ranking fraud detection system for mobile Apps. First propose to accurately find the ranking fraud by mining the active periods, specifically leading sessions, of mobile Apps.

- Identifying leading sessions for mobile apps:

During this section, we have a tendency to 1st introduce some preliminaries, so show a way to mine leading sessions for mobile Apps from their historical ranking records.

- Mining Leading Sessions

There ar 2 main steps for mining leading sessions. First, we'd like to find leading events from the App's historical ranking records. Second, we'd like to merge adjacent leading events for constructing leading sessions.

2. To improve the fraud detection efficiency

For this purpose first study how to extract and combine fraud evidences for ranking fraud detection.

- Ranking Based Evidences

  By analyzing the Apps' historical ranking records, we observe that Apps' ranking behaviors in a leading event always satisfy a specific ranking pattern, which consists of three different ranking phases, namely, rising phase, maintaining phase and recession phase. Specifically, in each leading event, an App's ranking first increases to a peak position in the leaderboard (i.e., rising phase), then keeps such peak position for a period (i.e., maintaining phase), and finally decreases till the end of the event (i.e., recession phase).

- 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 leaderboard. Thus, rating manipulation is also an important perspective of ranking fraud. Intuitively, if an App has ranking fraud in a leading session s , the ratings during the time period of S may have anomaly patterns compared with its historical ratings, which can be used for constructing rating based evidences.

- Review Based Evidences

1. Besides ratings, most of the App stores also allow users to write some textual comments as App reviews. Such reviews can reflect the personal perceptions and usage experiences of existing users for particular mobile Apps. Indeed, review manipulation is one of the most important perspective of App ranking fraud.

Specifically, before downloading or purchasing a new mobile App, users often first read its historical reviews to ease their decision making, and a mobile App contains more positive reviews may attract more users to download. Therefore, imposters often post fake reviews in the leading sessions of a specific App in order to inflate the App downloads, and thus propel the App's ranking position in the leaderboard. Although some previous works on review spam detection have been reported in recent years, the problem of detecting the local anomaly of reviews in the leading sessions and capturing them as evidences for ranking fraud detection are still under-explored. To this end, here we propose two fraud evidences based on Apps' review behaviors in leading sessions for detecting ranking frauds.

2. Strengths of this evidences:
   If there are equal number of users for an particular application who give both positive and negative reviews then it's difficult to find out correct results of that application.
   For this type of application different kind of attributes are provided like design, price, use, durability, availability, etc. For example design, use, availability are medium to high then application gets positive results.

- Evidence Aggregation

After extracting three types of fraud evidences, the next challenge is how to combine them for ranking fraud detection. Indeed, there are many ranking and evidence aggregation methods in the literature, such as permutation based models, score based mode and Dempster-Shafer rules. However, some of these methods focus on learning a global ranking for all candidates. This is not proper for detecting ranking fraud for new Apps. Other methods are based on supervised learning techniques, which depend on the labeled training data and are hard to be exploited. Instead, we propose an unsupervised approach based on fraud similarity to combine these evidences.

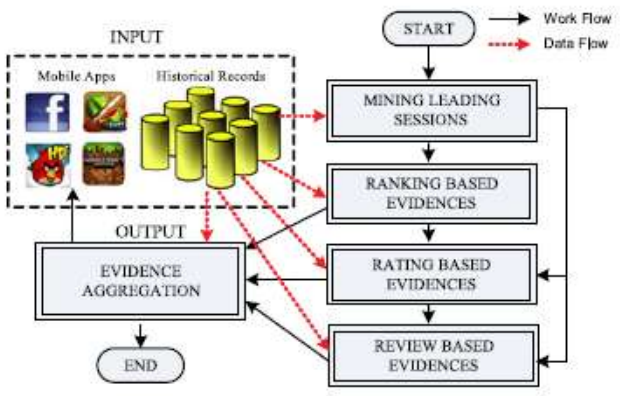

Fig1: The Framework of our ranking fraud detection system for mobile Apps

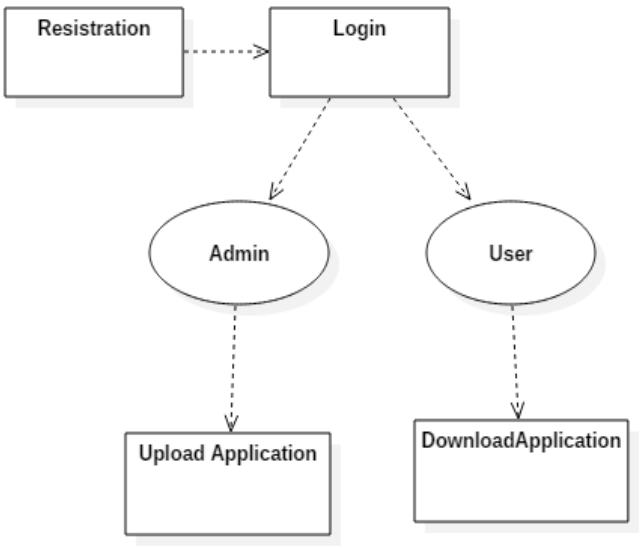

Fig 2: Data flow Diagram

4. CONCLUSIONS

In the existing system different methods are used such as web ranking spam detection, on-line review spam, mobile App recommendation etc. However, to the most effective of our data, none of previous works has studied the matter of ranking fraud detection for mobile Apps.

In proposed system, provide a holistic read of ranking fraud and propose a ranking fraud detection system for mobile Apps. First propose to accurately find the ranking fraud by mining the active periods, particularly leading sessions of mobile Apps. For achieving this goal following process are used:

Distinctive leading sessions for mobile apps during this section, we tend to 1st introduce some preliminaries, so show a way to mine leading sessions for mobile Apps from their historical ranking records.

Mining Leading Sessions

There ar 2 main steps for mining leading sessions. First, we'd like to get leading events from the App's historical ranking records. Second, we'd like to merge adjacent leading events for constructing leading sessions.

Then to improve the fraud detection efficiency we use different methods like Ranking Based Evidences, Review Based Evidences, Rating Based Evidences and Evidence Aggregation.

REFERENCES

1. Ntoulas, M. Najork, M. Manasse, and D. Fetterly, "Detecting spam web pages through content analysis," in Proc. 15th Int. Conf. World Wide Web, 2006, pp. 83–92.

2. B. Zhou, J. Pei, and Z. Tang, “A spamicity approach to web spam detection,” in Proc. SIAM Int. Conf. Data Mining, 2008, pp. 277–288.
3. N. Spirin and J. Han, “Survey on web spam detection: Principles and algorithms,” SIGKDD Explor. Newslett., vol. 13, no. 2, pp. 50–64, May 2012.
4. E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, “Detecting product review spammers using rating behaviors,” in Proc. 19thACMInt. Conf. Inform. Knowl. Manage., 2010, pp. 939–948.
5. Z. Wu, J. Wu, J. Cao, and D. Tao, “HySAD: A semi-supervised hybrid shilling attack detector for trustworthy product recommendation,” in Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2012, pp. 985–993.
6. S. Xie, G. Wang, S. Lin, and P. S. Yu, “Review spam detection via temporal pattern discovery,” in Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2012, pp. 823–831.
7. B. Yan and G. Chen, “AppJoy: Personalized mobile application discovery,” in Proc. 9th Int. Conf. Mobile Syst., Appl., Serv., 2011, pp. 113–126.
8. K. Shi and K. Ali, “Getjar mobile application recommendationsnwith very sparse datasets,” in Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2012, pp. 204–212.
9. H. Zhu, E. Chen, K. Yu, H. Cao, H. Xiong, and J. Tian, “Mining personal context-aware preferences for mobile users,” in Proc. IEEE 12th Int. Conf. Data Mining, 2012, pp. 1212–1217.