IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 2, February 2023

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.118

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.118| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2023 ||

| DOI:10.15680/IJIRSET.2023.1202033 |

An Market Identifying Rating and Review Based Ranking Aggregation Method

Dr.P.Anitha, Thiyagaraj S

Professor & Head, Department of Computer Applications (MCA), K.S.R. College of Engineering (Autonomous), Tiruchengode, India

Department of Computer Applications (MCA), K.S.R. College of Engineering (Autonomous), Tiruchengode, India

ABSTRACT: Nowadays ranking fraud in mobile App market became more popular in the market in order to display their apps in the popularity list and to boost their sales .Therefore, there will be increase in ranking fraud in the upcoming time, as the number of Apps developers and applications will likely grow very significantly. Many traditional methods of fraud analysis have been used to detect fraud. But these methods are complex and time consuming. However there is more need to adopt some better techniques which can ensure the ranking fraud detection efficiently by data mining analysis. This paper explores the data mining methods to identify the fraud by using Rank Aggregation algorithm. Further three types of proofs are studied they are ranking, rating and review proofs. And an aggregation method is used to aggregate all the proofs and will produce an optimized report for fraud detection.

KEYWORDS*:* Mobile Apps market, fraud detection, aggregation method, records of the apps, rating and review proofs.

1. INTRODUCTION

For the past few years the number of mobile Apps increases day by day. As per the statistics taken in 2012, per day 1 millions of apps been downloaded in which 91% are free apps and the remaining 9% are paid one. And in order to increase the sales, each app is ranked and maintained in the leading board. Many fraudulent activities taking place in this board only. Thus, leading sessions is used for marketing the apps. Higher the rank usually lead to a huge downloads of the app and billion of rupees in gain. And they also follow various ways to move their app to higher position and will also undergo various marketing techniques including advertisements. Recently many trends are being used in order to boost the apps sales.
For example, Google found the developer who undergone the ranking fraud in the app store.

In the literature, while there are some related work, such as Latent Dirichlet allocation , A taxi driving fraud detection systemic in city taxis, Rank aggregation via nuclear norm minimization, An unsupervised learning algorithm for rank aggregation, Unsupervised rank aggregation with distance-based models the problem of detecting ranking fraud for mobile Apps is still under-explored. Thus, we proposed a paper for ranking mobile fraud detection.

To solve this problem, first local deviations are proposed later they are combined and proposed for the global deviation. Finally all this proofs are categorized and it is maintained in the database. And also the patterns which are followed by the fraudulent apps to rank their apps is unique that is it varies for each and every leading sessions of the apps stores. Ranking is calculated based on the user's review and rating. Thus, further two types of proofs are recorded which helps to detect the fraud in the leading board. Finally an unsupervised proof is produced and aggregation method to integrate these three types of proofs for evaluating the credibility of leading board from mobile Apps. Thus the proposed system is more scalable, efficient and it performance is high comparing the existing.

The rest of this paper is partitioned as follows. In section 2 describe the leader board and rank aggregation. Section 3 contains proposed method. Simulated results are discussed in section 4. Finally, section 5 explains the conclusions and future works.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.118| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2023 ||

| DOI:10.15680/IJIRSET.2023.1202033 |

II. RELATED WORKS

In this section, some of the related works about leader board session and rank aggregation are discussed.

2.1 Leader board session

It is observed that finding the historical records of the App are not always ranked higher in the leader board but only in few events .Thus, it is also found that there exist some adjacent leading events which are closer to one another to form a leading sessions. Further, to find the ranking fraud from several leading sessions, an effective approach is developed called as Evidence Aggregation based Ranking Fraud Detection (EA-RFD).specifically, this method is denoted by score based aggregation (i.e., Principle 1) as EA-RFD-1, and dealed with rank aggregation (i.e., Principle 2) as EARFD- 2, respectively.

Definition 1 (Leading Event): Given a ranking threshold K_ 2½; K, a leading event e of App a contains a time range Te ¼ ½testart; te end_ and corresponding rankings of a, which satisfies start _ K_ < start_1, and end _ K_ < end þ1. Moreover, 8 tk 2 start; te end p1, we have raked_ K.

Definition 2 (Leading sessions): A leading session s of App a contains a time range Ts ¼ ½ts start; ts end and n adjacent leading events fe1; eng, which satisfies ts start ¼ te1 start, ts end ¼ ten end and there is no other leading session s that makes Ts . While, 8i 3 ½1; np, we have teip1 start to tei endp < f, where f is a predefined threshold of time for merging leading events.

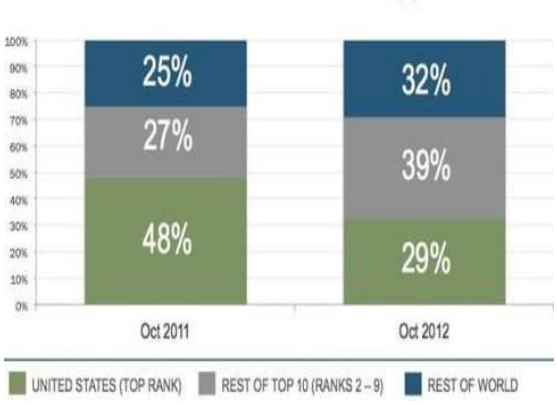


Figure 1 Distribution of ios and android apps in leading sessions

Fig 2 is a leader board of the apple apps store where it daily updates the ranking of each and every app in their store. And it also give the details of the app which losses its rank and also the app which moves up in the rank list. The fraud is happening mainly in these types of the leader boards and the proposed system helps to detect this fraud. Finally an optimized report is generated which helps to detect the fraudulent apps.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.118| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2023 ||

| DOI:10.15680/IJIRSET.2023.1202033 |

Figure 2 Example of leader board

2.2 Rank aggregation

If we are to cast the rank aggregation, we need to define our objective function. In this context, we would like to and a "super"-list which would be as "close" as possible to all individual ordered lists simultaneously. This is a natural requirement and the objective function, at least in its most abstract form, is very simple and intuitive.

$$\Phi(\delta) = \sum_{i=1}^{m} w_i d(\delta, L_i),$$

This is a proposed ordered list of length k = jLij, wi is the importance weight associated with list Li, d is a distance function.

III. PROPOSED METHOD

We find that the fraudulent Apps often have different ranking patterns in each leading session compared with normal Apps. And here the client can directly communicate with the server and can get the information about the rank of the app. And the user is also having rights to rate and give comments about the apps in the leader board. And this rating and review are aggregated to produce the rank of the app. Here the exact time when the rating and comments given is noted and maintained in the database. The request and response of the client and the server is secured using the private key.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.118| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2023 ||

| DOI:10.15680/IJIRSET.2023.1202033 |

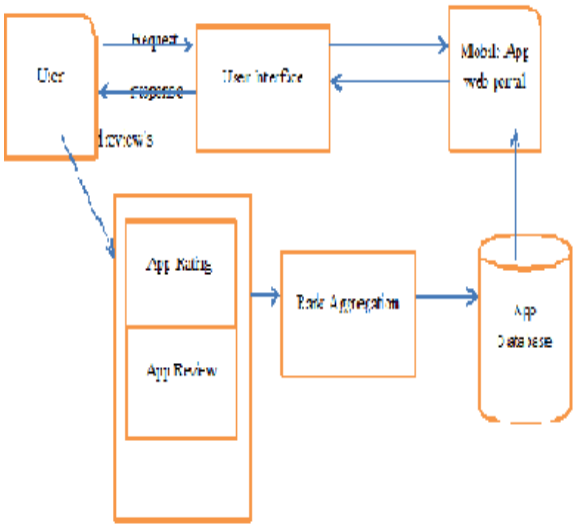


Figure 3 Architecture Diagram of Overall System

Figure shows the architecture of the overall system which consist of modules like User Interface, Identifying leading sessions, Ranking proof, Rating proof, Review proof and Evidence aggregation. We Design new Client side page for actively interact with server ,for improve secure purpose we generate unique Authentication scheme for each and every user(Application Data Owner)for check Competitor App Data Ranking History's . And also we design another Client side page's for user search, reviews, rating, the App.

Number of apps in Google play store	1,600,030
Number of apps in windows phone store	341,000
Number of apps in apple store	100bn

Table 1.Overview of the total number of apps

3.1 User interface:

In This We Design new Client side page for actively interact with server ,for improve secure purpose we generate unique Authentication scheme for each and every user(Application Data Owner)for check Competitor App Data Ranking History's .

And also we design another Client side pages for user search, reviews, rating, the App.

.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.118| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2023 ||

| DOI:10.15680/IJIRSET.2023.1202033 |

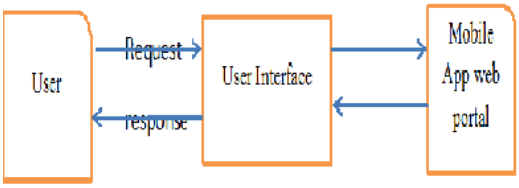


Figure 4 User interface diagram

3.2 *Identifying leading sessions*:

Raking fraud usually happens in leading session Therefore detecting ranking fraud is actually detect ranking fraud with in leading boards of mobile Apps. Specifically, we first propose a simple yet effective algorithm to identify the leading session of each application based on its ranking records. Then, with the analysis of Apps' ranking we find that the fraudulent Apps often have different ranking pattern in each leading session compared with normal Apps.

3.3 *Ranking Proof:*

A leading is composed of several leading events. Therefore, first we should analysis the basic features of leading events for bringing out the evidences. By analysing the Apps records.

Number of Mobile apps downloaded worldwide	102,082m
Projected number of apps downloads	258,692m
Number of free mobile apps	92.89bn
Number of paid mobile apps	9.29bn

Table2. Statistics of the data an overview

We found that App's ranking in a leading event always fulfil the specific pattern, which contain of three different phases, they are, rising phases, maintain phase and recession phase. Mainly, in each events, an App's ranking first rises to a peak position in the leader board (i.e., rising phase), then keeps such peak position for a period (i.e., maintaining phase), and finally decrease till the end of the events (i.e., recession phase).

3.4 *Rating proof*:

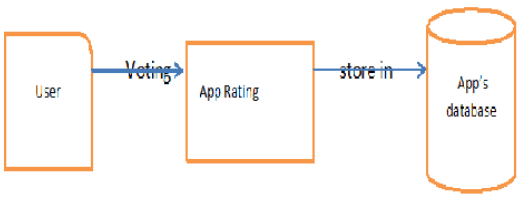


Figure 5 Rating based proofs diagram

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.118| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2023 ||

| DOI:10.15680/IJIRSET.2023.1202033 |

The Ranking based evidences are very well used for detecting ranking fraud. However, sometimes, it is not enough to only use raking based evidences. Mainly, after an App has been released, it can be rated by any user who downloads it. For App advertisement the very important features is app rating. An App with higher rating will attract more users to download and also ranked higher in the leader board. Thus, rating method is an important approach of ranking fraud. Certainly, if an App has ranking fraud in a leading session's, the rating during the time period of may have irregular patterns compared with its historical ratings in the board. Which can be used for constructing rating based evidences.

3.5 Review proof:

Figure 6 Review based proofs diagram

Apart from rating, most of the Apps stores also allow users to write some comments as app reviews. That reviews can reflect the personal viewpoint and usage of existing user for in use mobile Apps. Thus, review is one of the most importance points of ranking fraud in mobile app. Specifically, before downloading or purchasing a new mobile App, users often go through its historical reviews to easier their decision making, and a mobile App containing more good reviews will attract more user to download. Therefore, imposter often post fake review in the leading sessions of a specific App in order to inflate App download , and thus propel the App's raking position in the leading. Although some previous works on review detection have been reported in recent years, the problem of detecting the local anomaly of reviews in the leading session and taking them as evidences for ranking fraud detection are still under-explored.

3.6 Aggregating the proofs:

After identifying all the types of fraud evidences, the next thing is to combine them for raking fraud detection, however, there are many aggregation methods in the literatures, like permutation based models, score based models and Dempster-Shafer rules. However, many methods are there but proper method for detecting ranking fraud for new Apps is not found. Thus to solve this problem our methods are based on supervised learning techniques, which depend on the label training data and are difficult to be exploited. And also we propose an unsupervised approach which combines these evidences.

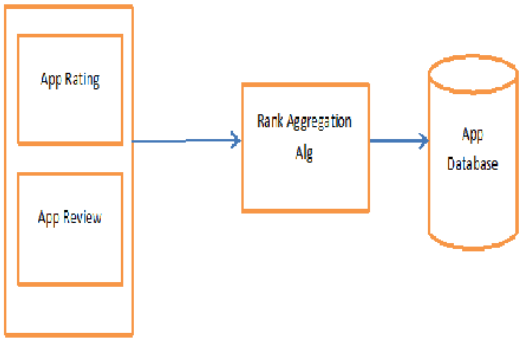


Figure 7 Aggregating the proofs diagram

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.118| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2023 ||

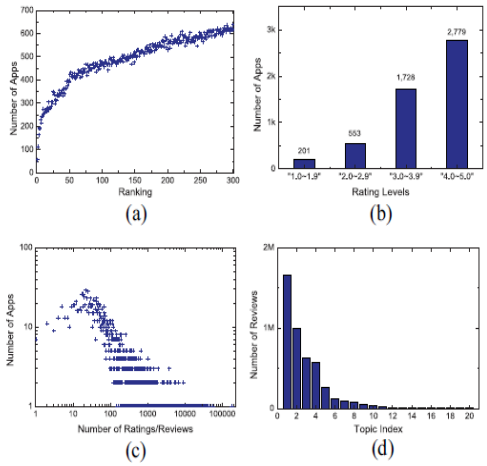
| DOI:10.15680/IJIRSET.2023.1202033 |

IV. DISCUSSION

Thus we focus on extracting evidences from Apps' historical ranking, rating and review records for ranking fraud detection.

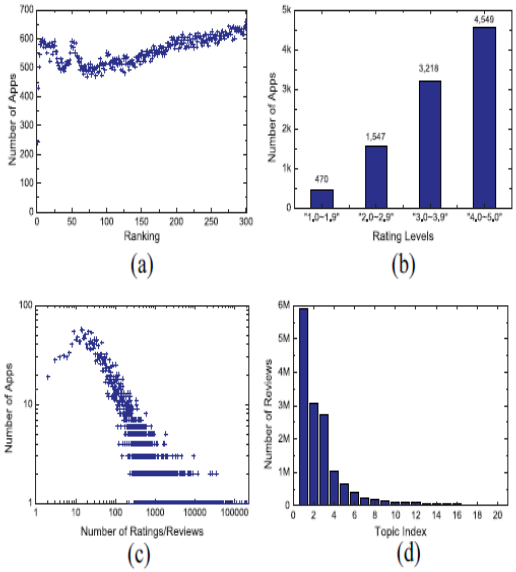
Our approach can discover the local anomaly instead of the global anomaly of mobile Apps. Thus, we should take consideration of such kind of local characteristics when estimating the credibility of Apps. To be specific, we define an App fraud score for each App according to how many leading sessions contains ranking fraud. Intuitively, an App contains more leading sessions, which have high fraud evidence scores and long time duration, will have higher App fraud scores.

Figure 8 Data distribution in top 300 paid apps



However, our approach is scalable for integrating other evidences if available, such as the evidences based on the download information and App developers' reputation. Second, the proposed approach can detect ranking fraud happened in Apps' historical leading sessions.

Figure 9 Data distribution in top 300 free apps



|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.118| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2023 ||

| DOI:10.15680/IJIRSET.2023.1202033 |

V. CONCLUSION AND FUTURE WORKS

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. In the future, we plan to study more effective fraud evidences and analyse the latent relationship among rating, review and rankings.

REFERENCES

[1] Y. Ge, H. Xiong, C. Liu, and Z.-H. Zhou, "A taxi driving fraud detection system," in Proc. IEEE 11th Int. Conf. Data Mining, 2011.
[2] D. F. Gleich and L.-h. Lim, "Rank aggregation via nuclear norm minimization," in Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2011, pp. 60–68.
[3] N. Jindal and B. Liu, "Opinion spam and analysis," in Proc. Int. Conf. Web Search Data Mining, 2008, pp. 219–230.
[4] Klementiev, D. Roth, and K. Small, "An unsupervised learning algorithm for rank aggregation," in Proc. 18th Eur. Conf. Mach. Learn., 2007, pp. 616–623
[5] Klementiev, D. Roth, and K. Small, "Unsupervised rank aggregation with distance-based models," in Proc. 25th Int. Conf. Mach. Learn., 2008, pp. 472–479.
[6] Klementiev, D. Roth, K. Small, and I. Titov, "Unsupervised rank aggregation with domain-specific expertise," in Proc. 21st Int. Joint Conf. Artif. Intell., 2009, pp. 1101–1106.
[7] 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
[8] Y.-T. Liu, T.-Y. Liu, T. Qin, Z.-M. Ma, and H. Li, "Supervised rank aggregation," in Proc. 16th Int. Conf. World Wide Web, 2007, pp. 481–490
[9] Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu, M. Castellanos, and R. Ghosh, "Spotting opinion spammers using behavioral footprints,"in Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2013, pp. 632–640.
[10] 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.
[11] G. Shafer, A Mathematical Theory of Evidence. Princeton, NJ, USA: Princeton Univ. Press, 1976
[12] K. Shi and K. Ali, "Getjar mobile application recommendations with very sparse datasets," in Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2012, pp. 204–212.
[13] 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.
[14] M. N. Volkovs and R. S. Zemel, "A flexible generative model for preference aggregation," in Proc. 21st Int. Conf. World Wide Web,2012, pp. 479–488

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.118

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details