ISSN XXXX XXXX © 2017 IJESC

Research Article Volume 7 Issue No.5

Discovery of Ranking Fraud for Mobile Applications

AratiTule[1], Prof. Rahul Shahane[2]
M.Tech Student[1], HOD[2]
Department of Computer Science and Engineering[1], Department of Information Technology[2]
Wainganga College of Engineering and Management, Nagpur, Maharashtra, India

Abstract:
As we all know every person in the world are mobile users in fact smart-phone users with android applications. So, Due to this popularity and well-known concept there will be a rapid growth in mobile technology we have seen. As well as in data mining concept mining the needed data from a particular application is very difficult and crucial task. Merging these two concepts of ranking frauds in android market and mining needed data is gone very tough for us and this is challenging situation. We are using these concept in whole paper. As we know that the mobile Apps has grown at vast speed in some years; as for march 2017, there are nearby 2.8 million Apps at google play and 2.2 Apps at Apple Apps store. In addition, there are over 400,000 independent app developers all fighting for the attention of the same potential customers. The Apple App Store saw 128,000 new business apps alone in 2014 and the mobile gaming category alone has competition to the tune of almost 300,000 apps. Here the main need to make fraud search in Apps is by searching the high ranked applications up to 30-40 which may be ranked high in some days or the applications which are in those high ranked lists should be verified but this is not applied when we work for thousands of applications added per day. So, we go for broad view by applying some technique to every application to judge its ranking. In this paper of our project discovery of ranking fraud for mobile applications, we develop a need to make a flawless, fraud less and result that shows corrected application accordingly provide ranking; where we actually make it happen by searching fraud of applications. They make fraud of App by ranked high the App by methods using such as human water armies and bot farms; where they make fraud by downloading application through various devices and give fake ratings and reviews. So, as we said above here we have to mine crucial data relating particular application such as review which we said comments and also so many other information we have to mine and place algorithm to detect fakeness in application rank.

Keywords: Ranking, Review, Rating based evidences; Pattern Analysis, Semantic based analysis, Aggregation.

I. INTRODUCTION:

On daily basis, an app leaderboard can be updated by app store which displays chart rankings of most popular apps, also it is an inspirational thing to make encouraged the development of mobile apps. In fact, for promoting mobile Apps, leader board of apps is the most important way of upgradient in the market. An app should be ranked higher depend upon how its chart of development raise and progressively it can make number of downloads and ultimately high revenue in dollar. There were different ways to advertise Apps promotional drive in order to get top position in App leaderboards the legal one is white hat basis to promote their App to get famous and alternately more number of downloads. But there are also some illegal ways say black hat basis for bumping up the App by using some deceptive means used by corrupt App developers to get famous in some short time period. This technique usually implemented by using so-called “internet bots” or “human water armies” to raise the App downloads, ratings and reviews in a very little time. Some are necessary points that is to restrict fraud, showed as given two constraints. The first constraint is that an app can be rated only once from a user login and the second is implemented with the aid of IP address that limits the number of user login logged per day. Finally, the proposed system will be evaluated with real-world App data which is to be collected from the App Store for a long-time period called historical records. In the existing system, from the collected historical records, the leading event and leading session of an app is identified. There are two main steps for mining leading sessions. First, we need to discover leading events from the App’s historical ranking records. Second, we need to merge adjacent leading events for constructing leading sessions. Careful observation shows that the mobile Apps are not always at top most position in leader board. But only in some time period called leading event which is form different leading sessions means ranking fraud particularly occur in this leading session. Then from the user judgmental feedback, three different types of evidences are collected namely ranking based evidence, rating based evidence and review based evidence. As our project based on evidences collected from app data; the one of the mostly judgment by people is rating based evidences which can be used to rate the app while downloading it or we can rate it after seeing its performance. It is most important evidence to judge the app. But as discussed above there are some techniques with help of which the rating can get increases by doing fraud. So, another judged evidence based technique is review based evidence; finding to make the exact specification of app whether it is good or bad app to download. In Review Based Evidences, besides ratings, most of the App stores also allow users to write some textual comments as App reviews. So, people may sure shot about downloading that specific app by reading comments specified in review section and also give their opinion about that app. Due to the vast number of apps, it is hard to search ranking fraud for each apps; so, it is important to have a scalable way to automatically detect ranking fraud without using any benchmark information. So here are come the concept of algorithm used in our project. Especially, this paper proposes a simple and effective algorithm to recognize the leading sessions of each mobile App based on its historical ranking records. Here we are proposing some statistical test where the statistics give the exact demonstration of activities done by the app to rank itself. If rank is maintained over period and graph falls and so much fluctuations can be seen then those apps should want solid verification to place it in last position or make it out of the play store. Also, rating and review history, which gives some anomaly patterns from apps historical rating and reviews records. Also, we are making here semantic analysis test over those collected data; for example, here in this test we use review to find positive, negative and neutral effects on those comments and judging app up to the mark. The rest of the paper is organized as follows: Section II, presents the literature survey over the related work. In section III, proposed system is presented. In section IV, implementation for each module are presented. Finally, the section V concludes paper.

II. LITERATURE SURVEY:

As we know before us many great peoples worked on this android app ranking fraud detection through ads so we just go through their study work and take inspiration from their work and build our improved system. SabbineniPoojitha, Balineni Venkata Sai Mrudula and VemuriSindhura[1] in this paper, they give a extensively comprehensive point of view of situating trickery and propose a situating compulsion exposure structure for flexible Apps. Particularly, they first proposed to extremely situate the situating blackmail by mining the dynamic time periods, particularly motivating sessions, of flexible Apps. They scrutinize three sorts of pronouncements, i.e., situating based confirmations, rating based verifications and study based verification, by showing Apps' situating, rating and review hones through experimental hypotheses tests. Besides, propose a progression based accretion system to join each one of the pronouncement for compulsion characteristic proof. Ranjitha.R, Mathumitha.K, Meena.S, S.Hariharan [2] had proposed system additionally, they are proposing two enhancements using appreciation of keep a tally by the admin to recognize the exact reviews and rating scores. Secondly, the fake response as a feedback by a same person for pushing up that app on the leader board is restricted. Two different limitations are taking into account for accommodating the feedback given to an application as a part of their response toward the app whether it is good or bad. The first constraint is that an app can be rated only once from a one particular user login and the second are put into action with the id of IP address that limits the number of user login logged per day. Finally, the proposed system will be estimated with real-world App data which is to be composed from the App Store for a long-time period. R.Vinodharasi, P.Ramadoss[3] proposed to precisely situate the ranking fraud by mining the dynamic periods, namely leading sessions, of mobile Apps. Additionally, we examine 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 arithmetical mining based hypotheses tests. In addition, in this project and optimization based application used to incorporate all the evidences for fraud recognition based on EIRQ (efficient information retrieval for ranked query) algorithm. Finally, estimate the projected system with real-world App data collected from the IOS App Store for a long-time period. Experimentation was need to be done for authenticate the efficiency of the proposed system, and show the scalability of the recognition algorithm as well as some reliability of ranking fraud activities. Phopse P.E, Jondhale S.D[4] had provide a holistic view of ranking fraud and propose a ranking fraud appreciation system for mobile Apps. Additionally, to first propose to precisely 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 irregularity instead of global irregularity of App rankings. Furthermore, we consider three types of evidences, i.e., ranking based evidences, rating based evidences and review based evidences, by modeling Apps' ranking, rating and review nature through statistical hypotheses tests.In addition, to project an optimization based aggregation method to consolidate all the evidences for fraud recognition. Xiong and Zhu[5] had projected a ranking fraud detection system for android mobile apps. In this paper principally, they both demonstrated that ranking fraud take place in most important sessions for each app from its previous ranking accounts. Then, they recognized ranking based, rating based and review based confirmation for discovering ranking fraud. Moreover, they proposed an optimization based aggregation system to merge all the evidences for estimate the consistency of most important sessions from mobile apps. Priyanjai and Pankaj[6] planned techniques for assessment of investigation and invent pattern of android apps based on cloud computing and data mining. They developed system ASEF and SAAF for android apps to achieve protection. They also explain a tactic that performs apps security and provide user friendly interface on a mobile phone. Anuja A. Kadam, Pushpanjali M. Chouragade [7] make available a disciplined study on the different procedures of malicious application recognition in android mobiles. The examination of authorization induces possibility in Android apps on a large-scale in three stages. First upon position all the entity permissions with respect to their feasible risk with different processes. Secondly, classify subsets of risk permissions. Then using several algorithms identifies the suspected apps based on the recognized subsets of risky permissions.   Jakub Zilincan, Michal Gregus [8] had given the dedicated work on Search engine optimization techniques, often summarized to SEO, should lead to first situation in unprocessed search results. Some optimization techniques or procedures do not modify over time, yet still form the foundation of SEO. However, as the Internet and web design develop enthusiastically, new optimization procedures come in to account and sometime does not work. Thus, they have focused on most important features that can help to get better a pose in search outcome. It is important to accentuate, that none of the procedure can make sure it because search engines have complicated algorithms, which measure the superiority of Web pages and obtain their position in search results from.   Xiang Wang, Yan Jia, Ruhua Chen, Bin Zhou [9] in that they had told users can interpret themselves using free tags in microblogging website such as Sina Weibo. The tags of a user exhibit. The description of the user and are normally in a unsystematic direct without any significance or importance information. It restricts the usefulness of user tags in system suggestion and other applications. They also proposed a user tag ranking representation which is based on interactive and attractive dealings between users. Manipulate power between users is measured in our user tag ranking method. Significance scores between tags and users are also utilized to rank user tags.

III. PROPOSED SYSTEM:

There were so many investigations run on fraud finding areas such as for web ranking spam detection, online review spam detection and portable App awareness. Here we develop some fraud finding activity over fraud Apps, which are made by fraud App developers for leading their App ranking in leaderboard. Firstly, we take an account over time when fraud happens so that we can easily make sure about the fraudness of the App. So, detection of such App is conduct by making leading sessions of small leading events that shown when the Apps are in phases of achievement. Those are rising phase, maintaining phase and recession phase where we detect App ranking behavior from historical ranking records. Such a demonstration done over small start, we have called local anomaly detection strategy. When these phases run over App historical records the liable Apps can maintain their level of ranking constantly over long period, but the fraud Apps found fluctuation over that time and find fraud. But some App developers run the wrong way to make the other Apps developer to downfall by ranking wrongly to their App. So, there are more fraud finding evidences such as rating and review based evidences. The users who are newly logging to the app stores, they decide based on the existing ranking, rating, reviews for the individual apps.User do not understand about fake Apps and may download it. As review, also may be fake but we have conduct one semantic analysis test here on reviews by conducting test as Natural language processing to make the overall comments judge and get some final sure comment to make trust over the App by customer who downloaded it. So, exact review and rating finding and then calculated by matching the whole findings for sure is the main statement of the problem shown over how we implement our project.

Here is the proposed approach,

In our approach, we will read the dataset, and then pre-process it to separate out the textual reviews and the statistical reviews. The statistical reviews will then be mapped in sessions and each session will be checked separately. If we find that the sessions are evenly organized, then the chances of the review being fake will be less, but if the sessions are abruptly organized, meaning that if we find that for session S1 the mean reviews were very good, but for session S2 the mean reviews suddenly dropped then it means that the reviews in Session S1 were paid and might not be correct. Thus we will find out if the reviews were fake or genuine Once the statistical reviews are completed, then we will read all the textual reviews and apply NLP on the reviews. NLP process will consider of 2 parts, Parts of Speech (POS) tagging which will find out the parts of speech for each of the input words, and then Chunking which will remove all the unwanted POS from the reviews and give only the action words in the reviews. We will process all these action words and find out the type of review these action words produce, then again apply session based checking to find out if the reviews are fake or genuine Combining results from both of them will identify the true nature of the reviews and will generate the results. Advantages of Proposed system: The recommended framework is extensible and can be continued with other domain develop evidences for ranking fraud detection.

IV. IMPLEMENTATION:

Ranking fraud detection have proofs for detection of fraud; such proofs are evidences called as ranking based, review and rating based evidences. These evidences are used to mine leading sessions. Ranking based, review and rating based evidences are applied step by step. The specific ranking pattern is fulfilled by app ranking behavior in ranking based evidences. In rating based evidences, rating pattern that is ratings given by user is used for ranking fraud detection. Rating is given by user at the time of downloading the App or after judging that App by using it after some time. If the ratings are high for that App in more quantity or more App users give almost high ranking then that App is attracted by more mobile users. In this there are more chances of make fraud by App developer by earn ratings performed in leading sessions. In review based evidences, reviews are comments given by mobile app users given after judging that app after downloading it. But here before mobile app users downloading that App they will definitely goes through those comments to get others view to clear their way to download that App or not. As the number of mobile Apps increases day by day, fraudulent Apps must be detected. So we have proposed a simple and effective algorithm for identifying the leading sessions of each App based on its historical ranking of records. We recognize that the fraudulent Apps often have different ranking patterns in each leading session, compared with normal Apps according to ranking behaviors of Apps. Some fraud evidences are identified from Apps historical ranking records, which results in development of three functions to detect likewise ranking based fraud evidence. So moreover, here two types of fraud evidences based on Apps rating and review history are proposed.

V. IDENTIFYING EVIDENCES FOR RANKING FRAUD DETECTION:

1. Identifying Leading Sessions: Leading sessions are the base for detecting fraud in mobile App as ranking fraud usually happens in leading sessions. And hence detecting ranking fraud is actually detecting ranking fraud within leading session of mobile Apps which we mine from mobile Apps historical ranking records. There are two main steps for mining primary sessions. First, we need to determine leading measures from the App's previous ranking records. Second, we need to collaborate neighboring leading events for developing leading sessions. Specifically, we first propose a simple yet effective algorithm to identify the leading events of each App based on its historical ranking records. Then, we merge adjacent leading events for constructing leading sessions. As per the observation the mobile apps do not always ranked high in the leader boards, in fact in some leading events only. With the analysis of Apps' ranking behaviors, the fraudulent Apps often have different ranking patterns in each leading session compared with normal Apps. Therefore, the problem of identifying ranking fraud is to find out vulnerable leading sessions.

2. Ranking based evidences: A leading session is composed of several leading events. Therefore, we should first analyze the basic characteristics of leading events for extracting fraud evidences. By analyzing the Apps' historical ranking accounts, Apps' ranking behaviors in a leading incident always assure a specific ranking pattern, which consists of three different ranking segments, expanding phase, maintaining phase and collapse phase. Mainly, in each leading event, an App's ranking first improve to a peak or extent position in the leaderboard (i.e., rising phase), then maintain such peak position for a phase (i.e., maintaining phase), and at last declines till the end of the event (i.e., recession phase). Definitely, such a ranking pattern confirms a significant consideration of leading event. In next section, we formally describe the three ranking phases of a leading event.

3. Rating based evidences: The ranking based evidences are first step towards ranking fraud recognition. However, sometimes, it is not satisfactory to only use ranking based evidences. Take an example, some Apps formed by the legendary developers, such as Gameloft, may have some leading events due to the developers' trustworthiness and the "word-of-mouth" advertising effect. Moreover, some of the permissible marketing services, such as "limited-time discount", may also consequence in significant ranking based evidences. To solve this matter, we also study how to extort fraud evidences from Apps' historical previous rating records. Indeed, user rating is one of the most important features of App advertisement. A higher rated App 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. Intuitively, if an App has ranking fraud in a leading session, the ratings during the time period of that leading session may have drastically changed patterns if seen from its previous historical ratings, which can be used for constructing rating based evidences. Rating to app is given by the user who downloaded it. Hence rating is one of the main evidence in ranking fraud of apps. In this module it performs preprocessing of ratings that is it removes ratings that are less than or equal to two in number given as star to that App that is if 5 star given to the App is one in number among 100 users given other rating but not 5 star then it should be deleted and thus calculates rating score by summing all the ratings class collected and decision is taken on the basis of rating which scores high amongst all.

4. Review based evidences: Including ratings, most of the App stores also allow users to write some textual comments as App reviews to submit to the developer. Such reviews can reflect the personal observations and usage understanding of breathing users for particular mobile Apps. Indeed, review management is one of the most important base of finding App ranking fraud. Specifically, before downloading or purchasing a new mobile App, users often first read its previous historical reviews to simplify their conclusion making and a mobile App includes more encouraging reviews may attract more users to download. Therefore, imposters often place counterfeit reviews in the leading sessions of a specific App in order to inflate the App downloads, and thus boost the App's ranking position in the leaderboard. Therefore, manipulation and detection of reviews is one way used over shady app developers to expertise the app. Hence reviews are used to detect the ranking fraud in Mobile App industry is the foremost viewpoint to find ranking fraud. On semantic analysis level review rechecking can be done to show the concluded review to user of app to make them easy to judge that app. As the Sentiment Analysis is a natural language processing task that deals with finding orientation of opinion in a piece of text with respect to a topic. To determine the semantic orientation of the sentences a dictionary based technique of the unsupervised approach is adopted. To determine the opinion words and their synonyms and antonyms WordNet is used as a dictionary. This module performs pre-processing of reviews and then performs sentiment analysis on pre-processed reviews. As the growing market of internet brought to the conclusion of product reviews as it made easy our decision about that product and as Internet is used by everyone the numbers of reviews that a product receives grow rapidly. To read all of comments is very time taking task for a potential customer and make a decision on whether to buy that product or not. Thus, mining this data about reviews, preprocessing that data, and classify them is an important task to make the reviews result corrected as shown below on stepwise proposing of such work: Gathering data for reviews from app store, and other sources: To determine the polarity of the sentences, based on aspects, large numbers of reviews are collected from the Web. There are lots of websites on the Internet where the large numbers of customer reviews are available. Amazon website (www.amazon.com) and also play stores like google play are used to collect the reviews. Pre-processing data to remove any missing entries (using filtering technique): To determine the semantic orientation of the sentences a dictionary based technique of the unsupervised approach is adopted. To determine the opinion words and their synonyms and antonyms WordNet is used as a dictionary; also, it plays a vital role in detecting any missing entries using filtering technique. Semantic matching for finding quality of review (Positive, Negative or Neutral): A large amount of reviews of users are collected on the Web that needs to be explored, analyze and organized for better decision making. Opinion Mining or Sentiment Analysis is a Natural Language Processing and Information Extraction task that identifies the user's views or opinions explained in the form of positive, negative or neutral comments and quotes underlying the text. Aspect based opinion mining is one of the level of Opinion mining that determines the aspect of the given reviews and classify the review for each feature. Semantic Matching: Algorithms and Implementation - Semantic Scholar

The system performs this task in several steps as follows: -

4.1 Data Collection: To determine the polarity of the sentences, based on aspects, large numbers of reviews are collected from the Web. There are lots of websites on the Internet where the large numbers of customer reviews are available. Amazon website (www.amazon.com) is used to collect the reviews.

4.2 POS Tagging: After collecting the reviews, they are sent to the POS tagging module where POS tagger tag all the words of the sentences to their appropriate part of speech tag. POS tagging is an important phase of opinion mining, it is necessary to determine the features and opinion words from the reviews.POS tagging can be done manually or with the help of POS tagger. Manual POS tagging of the reviews take lots of time. Here, POS tagger is used to tag all the words of reviews. 4.3 Feature Extraction: All the features are extracted from the reviews and stored in a database then its corresponding opinion words are extracted from these reviews. It will find out whether the comment is positive, negative or neutral. If word is positive then it will add plus one to score; if word is negative it will minus one from score. Sometimes it is unable to find sentiment of some reviews, that time it makes the use of Naive Bayes classifier. In this way it will find final score by analyzing sentiment of each review and determine whether app is fraud or not on the basis of review evidences.

Algorithm:

```
1. Read all feedback information
2. Divide the information into sessions
3. For each session find the feedback obtained, to get the list
S1 F1
S2 F2
S3 F3
Sn Fn
Where Si is the session, and Fi is the feedback from that session
```

4. Check if the feedbacks have a common trait,

```
if(F1 = F2 and F2 = F3 and .... Fn-1=Fn)
Then it means the review is genuine
else
if there is a abrupt shift in the pattern, then the feedback might be non-genuine
For NLP based technique,
1. Read all feedback information
2. For each feedback, find action words using POS Tagging and Chunking process
3. Evaluate the sentiment from the feedback and mark the feedback as Good or Bad
4. Divide the feedback into sessions
5. For each session find the feedback obtained, to get the list
S1 F1
S2 F2
S3 F3
.
.
Sn Fn
Where Si is the session, and Fi is the feedback from that session
6. Check if the feedbacks have a common trait,
if(F1 = F2 and F2 = F3 and .... Fn-1=Fn)
Then it means the review is genuine
else
if there is a abrupt shift in the pattern, then the feedback might be non-genuine Combine results from both the algorithms to conclude if the given feedback is genuine or not.
```

5. Pattern analysis using machine learning:

There are two main steps for mining leading sessions. First, we need to discover leading events from the App's historical ranking records. Second, we need to merge adjacent leading events for constructing leading sessions. 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). Indeed, such a ranking pattern shows an important understanding of leading event. An App has several impulsive leading events with high ranking positions. In contrast, the ranking behaviors of a normal App' sleading event may be completely different. For example, ranking records from a popular App "Angry Birds: Space", which contains a leading event with a long-time range (i.e., more than one year), especially for the recession phase. In fact, once a normal App is ranked high in the leaderboard, it often owns lots of honest fans and may attract more and more users to download. Therefore, this App will be ranked high in the leaderboard for a long time. Based on the above discussion, we propose here some ranking based signatures of leading sessions to construct fraud evidences for ranking fraud detection.

6. Result analysis based on the matching: After extorting three types of fraud evidences, the next dare is how to merge them for ranking fraud detection. Indeed, there are many ranking and evidence association techniques in the literature that we have studied before, such as transformation based models, achieve based models, and Dumpster-Shafer rules. However, some of these methods spotlight on learning a worldwide ranking for all contenders.

VI. CONCLUSION:

Here developed a ranking fraud detection system for mobile Apps. Specifically, here 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, here identified ranking based evidences, rating based evidences and review based evidences for detecting ranking fraud. Moreover, here proposed an optimization based aggregation method to integrate all the evidences for evaluating the credibility of leading sessions from mobile Apps. A unique perspective of this approach is that all the evidences can be modeled by statistical hypothesis tests, thus it is easy to be extended with other evidences from domain knowledge to detect ranking fraud. Finally, here validate the proposed system with extensive experiments on real-world App data collected from the App store. Experimental results showed the effectiveness of the proposed approach.

VII. REFERENCES:

[1]. Ranjitha.R, Mathumitha.K, Meena.S, S.Hariharan, "Discovery of Ranking of Fraud for Mobile Apps", International Journal of Innovative Research in Engineering & Management (IJIREM) ISSN: 23500557, Volume-3, Issue-3, May-2016.

[2]. SabbineniPoojitha, Balineni Venkata Sai Mrudula and VemuriSindhura, "A Novel Method To Identify False Apps Through Data Mining", International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE) ISSN: 0976-1353 Volume 23 Issue 5 –SEPTEMBER 2016.

[3]. R.Vinodharasi, P.Ramadoss, " Efficient Retrival Of Mobile Apps Using EIRQ", International Journal Of Engineering And Computer Science ISSN: 23197242 Volume 5 Issues 6 June 2016, Page No. 1683016835.

[4]. Phopse P.E, Jondhale S.D, "Discovery of Ranking &Rating Fraud for Mobile Application", International Journal of Research in Science & Engineering e-ISSN: 2394-8299 Volume 2 Issue 4.

[5]. Hengshu Zhu, Hui Xiong, Senior Member, IEEE, Yong Ge, and Enhong Chen, Senior Member, IEEE Discovery of Ranking Fraud for Mobile Apps‖ IEEE Transactions On Knowledge And Data Engineering, Vol. 27, No. 1, January 2015.

[6]. Pranjali Deshmukh, Pankaj Agarkar ―Mobile Application For Malware Detection‖ International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 02 Issue: 02 | May-2015

[7]. Anuja A. Kadam ,Pushpanjali M. Chouragade ―A Review Paper on: Malicious Application Detection in Android System‖ International Journal of Computer Applications (0975 – 8887) National Conference on Recent Trends in Computer Science & Engineering (MEDHA 2015).

[8] Jakub Zilincan ,MichalGregus "Improving Rank of a Website in Search Resuts – a Experimental Approach"2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet

[10] App Analytic: A Study on Correlation Analysis of App Ranking Data Sun-Young Ihm; Woong-KeeLoh; Young-Ho Park Cloud and Green Computing (CGC), 2013 Third International Conference on Year: 2013 Pages: 561 ▪ 563, DOI: 10.1109/CGC.2013.95 IEEE Conference Publications

[11]. Jakub Zilincan, MichalGregus "Improving Rank of a Website in Search Resuts – a Experimental Approach"2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing978-1-4673-9473-4 /15 $31.00 © 2015 IEEE

[12]. L. Azzopardi, M. Girolami, and K. V. Risjbergen, "Investigating the relationship between language model perplexity and ir precision-recall measures," in Proc. 26th Int. Conf. Res. Develop. Inform. Retrieval, 2003, pp. 369–370.

[13]. D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," J. Mach. Learn. Res., pp. 993– 1022, 2003.

[14]. 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, pp. 181–190.

[15]. 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.

[16]. T. L. Griffiths and M. Steyvers, "Finding scientific topics," Proc. Nat. Acad. Sci. USA, vol. 101, pp. 5228–5235, 2004.

[17]. G. Heinrich, Parameter estimation for text analysis, "Univ. Leipzig, Leipzig, Germany, Tech. Rep., http://faculty. cs.byu. edu/~ringger/CS601R/papers/Hei nrich-GibbsLDA.pdf, 2008.

[18]. N. Jindal and B. Liu, "Opinion spam and analysis," in Proc. Int. Conf. Web Search Data Mining, 2008, pp. 219–230.

[19]. J. Kivinen and M. K. Warmuth, "Additive versus exponentiated gradient updates for linear prediction," in Proc. 27th Annu. ACMSymp. Theory Comput., 2005, pp. 209–218.

[20]. E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, "Detecting product review spammers using rating behaviors," in Proc. 19th ACMInt. Conf. Inform. Knowl. Manage. 2010, pp. 939–948.

[21].Caruana, R. (1997). Multitask Learning. Machine Learning, 28, 41–75.

[22]. A.-M. Popescu and O. Etzioni,(2005) "Extracting product features and opinions from reviews," presented at the Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, Vancouver, British Columbia, Canada.

[23]. B. Pang, L. Lee, and S. Vaithyanathan,(2002),"Thumbs up? Sentiment classification using machine learning techniques" In Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 79–86.

[24]. Bing Liu,(2012), "Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers".