ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 11, November 2016

Detection of Fraud Ranking in mobile apps Using Outlier analysis of evidences

Shaik Abdul Jabbar, Dr.K.F.Bharathi

***Abstract*— This paper helps the people to be aware of genuine mobile apps .The number of mobile applications are growing day by day. So ranking of mobile applications plays a major in the decisions of users of mobile applications . However, from the experience and literature all ranks may not be genuine .Giving information about mobile applications can help mobile users to be alert while making important decisions.. In this project the detection of fraud ranking is improved by incorporating some more evidences that are related to download information and for extracting new evidences outlier technique is used and enhancing the aggregation method among the evidences .Atlast, we provide entire view of fraud ranking .we evaluate the proposed system with real world app data and give awareness about the fraud applications**

***Keywords*— Fraud ranking , Mobile apps , Outlier analysis , Evidences ,Download information.**

I. INTRODUCTION

As the mobile applications are growing day by day users of mobile applications are approaching daily App leaderboards, which are launched and are being maintained by mobile app stores.These leader boards shows the rankings of most popular Apps.The mobile app which acquires highest rank on the leaderboard usually tends to a large number of downloads and make profit in revenue for app developers.Therefore, App developers tend to explore various ways such as advertising campaigns,offering cashback and gift coupons to promote their Apps in such a way that to have their Apps ranked as high as possible in App leaderboards.These are the genuine methods to promote a mobile app.But to face the competition from the new apps some of the app developers are approaching fraudulent ways to stand on the top of the leaderboard. Some of these fraudulent ways are mobile application developers may try to manipulate rankings,they may give fraud ratings and even can manipulate reviews given by the users in order to make their applications popular

. One should remember that ranking fraud doesn't happen throughout the life cycle of an app ,so we need to detect the time when the fraud happens . The detection of fraud ranking is somewhat improved by incorporating some more evidences. So evidences relating to download information will be added and for extracting these new evidences outlier technique is used which usually means picking up abnormal object from a group of objects(apps in this case) .

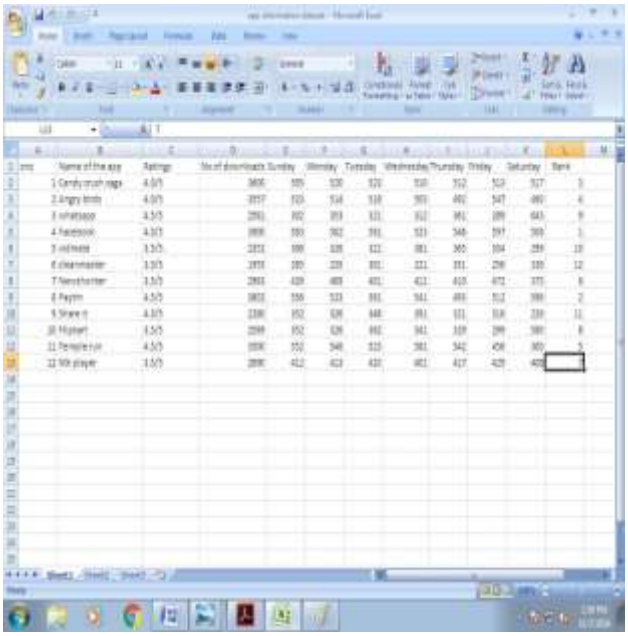


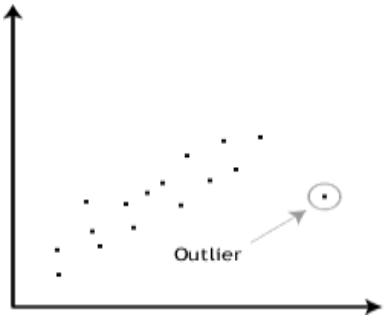
Fig.1dataset related to different mobile apps

II. EXTRACTING NEW EVIDENCES USING OUTLIER ANALYSIS

In this section we study how to extract new evidences as this is the primary task to find ranking fraud.For this purpose we have used outlier analysis technique .

A.Outlier analysis

Outlier analysis is a technique which is used to show things or phenomena that lie outside normal experience.It means outlier is a technique that is used to identify unusual or abnormal data from the dataset.For many situations we scattered plots to detect outlier.For example the following figure shows the outlier for certain data.



if tc-tb = 10 then assign p1=5 and and if tc-tb=20 then p1=4

if tc-tb =30 then assign p1-3 and so on

For certain data a graph is plotted with certain attributes and points are plotted according to it.Unexpectedly one point is far away from the usual set of points which shows abnormal data.This is indicated within a small circle .In this way outlier technique is used to find abnormal data in different situations such as weather fore casting which shows abnormal temprature, dataset of coronary illness patients which shows abnormal heart fuctioning,In banking with abnormal operations etc.,

B.Dataset

A dataset is having the data related to different apps.It contains Name of app,Ratings,No of downloads in different days of a week ,Rank.This data is a historical data of certain apps within a week.As this data changes from time to time we have to update it. We gather dataset of top 12 apps with various fields. we arrange the dataset into our required fashion. A Dataset of information about the applications is shown in Fig.1.By using this information we plot a graph of selected/suspicious apps with days of a week and number of downloads as attributes.

C.Graphical representation:

The dataset consists of data related to different applications along with number of downloads.By using this data a graph should be plotted with days of a week and number of downloads as attributes. Fig.2 shows that particular graph.In Fig.2 ta shows starting of raising phase ,tb shows starting of maintaining phase,tc shows starting of recession phase and td shows ending of recession phase.We are calculating μ by tc-tb.There is more chance of fraud if this value is in decreasing order

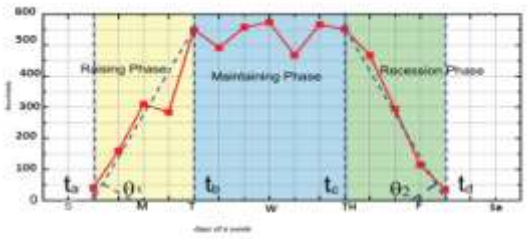


Fig.2 Graph showing different phases of particular app

D.Statistical Operations

Let μ =10,20,30,40,50 hours

tc-tb = maintaining phase period

Basing on this we have to calculate probability of fraud

Probability = [No. of possible outcomes/Total no. of outcomes]

For an example p = p1/5

III. EXPERIMENTAL RESULTS

We perform outlier analysis on the statistical data of download information of various apps to extract new evidences. To increase accuracy we are adding new evidences ,but extracting these new evidences is done by outlier analysis .For this purpose we are taking some historical data related to different applications which include ratings ,reviews and download information.This information is plotted on a graph with atttributes downloads and days of a week.Usually basing on the probability outcomes we will get an idea about the fraud apps.In this way, That we discover which apps are approaching fraudulent ways to get highest rank on the leaderboard position. To that extent we give information to the users about the fraud apps with some related information. At the end of this project we give awareness for the people about the genuine mobile apps.

Finally, we prepare a report is having total probabilities of different combinations which are mentioned in our analysis phase.

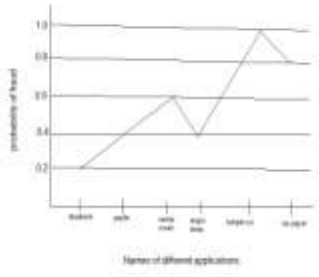


Fig. 3 Final Report of experimental results

IV. CONCLUSION AND FUTURE ENHANCEMENT

This project helps the people to be aware of genuine mobile apps.Statistical data about apps is taken and fraud detection is performed.To detect fraudulence in mobile applications several evidences are used to increase accuracy which include download information.To extract new evidences outlier technique is used on the dataset.Outlier technique plays a major role in this project

to find fraudulence.As we are increasing accuracy of the fraud detection by including new evidencs ,these evidences should be appropriate.For that purpose we are using outlier technique which easily pickups the abnormal data from the dataset.In future more number of datasets are taken and accuracy can be increased by adding some more evidences and to find out the new evidences outlier technique can be enhanced.

REFERENCES

[1]"GoogleSearchengine".Available:http://www.google.com

[2](2014).[Online].Available:http://en.wikipedia.org/wiki/information_retrieval

.

[3]Zhu, H., Ge, Y ., and Chen, E. (2013). Discovery of Ranking Fraud for Mobile Apps.IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING ,p 1-14.

[4]Klementiev, D. Roth, and K. Smal "Unsupervised rank aggregation with distance –based models," in proc. 25th Int. Conf.Mach. Learn ., 2008, pp. 472-479.

[5]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.

[6]Z. Wu, J. Wu, J. Cao, and D. Tao, "HySAD: A semi-supervisedhybrid shilling attack detector for trust worth product ,"in Proc. 18th ACM SIGKDD Int. Conf. Knowl. DiscoveryData Mining, 2012, pp. 985–993.

[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] H. Zhu, H. Cao, E. Chen, H. Xiong, and J. Tian, " Exploiting enriched contextual information for mobile app classification," in Proc. 21st ACMInt. Conf. Inform. Knowl. Manage., 2012, pp. 1617–1621

.

[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.

[10]A. Ntoulas, M. Najork, M. Manasse, and D. Fetterly, "Detectingspam web pages through content analysis," in Proc. 15th Int. Conf.World Wide Web, 2006, pp. 83–92.

[11] A. 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. DiscoveryData Mining, 2013, pp. 632–640

[12] Y.-T. Liu, T.-Y. Liu, T. Qin, Z.-M. Ma, and H. Li, Supervised rank aggregation," in Proc. 16th Int. Conf. World WideWeb, 2007,pp. 481–490.

Shaik Abdul Jabbar obtained B.Tech degree in Computer Science and Engineering from G.Pullareddy Engineering college, Kurnool Affiliated to Jawaharlal Nehru Technological University, Anantapur,A.P,India. Currently pursuing M.Tech in Computer Science from Jawaharlal Nehru Technological University Anantapur College of Engineering, JNT University, Anantapur, A.P, India, during 2014 to 2016. His research interests include Data Analytics and Data Mining.

Dr K.F.Bharati is currently working as an Assistant Professor in Computer Science at JNTUA College of Engineering, JNT University, Anantapur, A.P, India. She received her Ph.D degree in Computer Science and Engineering from JNT University, anantapur,A.P, India. she has around 10 years of experience as a Lecturer in Research and Development with strong analytical background in the education sector. Her research interests include data analytics and data mining