## RESEARCH ARTICLE

**National Conference on Innovation in Science, Engineering and Management (NCISEM-2022)**
**Date of Conference 11-12 March 2022**

## PROPOSED SYSTEM FOR DETECTION OF FRAUD APPLICATION USING REVIEW ANALYSIS

**Divya Bhoyar[1], Mayur Bajait[1], Samiksha Mule[1] and Prof. A.R. Welekar[2]**

1. Final Year Students, Department of Computer Engineering, Bapurao Deshmukh College of Engineering, Sevagram, Wardha.
2. Assistant Professor, Department of Computer Engineering, Bapurao Deshmukh College of Engineering, Sevagram, Wardha.

……………………………………………………………………………………………………………....

*Manuscript Info*

…………………….

*Abstract*

………………………………………………………………

In today's era, thanks to rapid development within the mobile technology and mobile devices, the applications i.e., mobile apps are being very exciting and popular concept. As there's large number of mobile Apps, ranking fraud is that the challenge considers front of the mobile App market. Ranking fraud is that the term used for concerning fraudulent or suspicious activities having the intention using tricky means frequently for increasing their Apps sales. The foremost aim is to develop such system that find ranking, rating and review behaviors for investigating review based evidences, rating based evidences and ranking based evidences so aggregation supported optimization to mix all the evidences for detection of fraud.

……………………………………………………………………………………………………………....

## Introduction:-

The mobile industry is developing rapidly; therefore, the numbers of mobile applications are increasing day by day within the market. As there are many apps available in market users are in fuzzy state while downloading the apps for his or her use. Different App stores like Google play store and Apple store launched their leader board on each day to inspire the users to download most well liked applications by observing the ranking of applications. of course, to advertise a specific mobile Apps, leader board of apps is that the most significant way within the market. An app which is at the high on the leader board ends up in sizable amount of downloads and it'll gain maximum profit. so as to own their Apps ranked as high as possible, app developers promote their apps using various ways like advertising, offers etc. Such applications harm to phone and also may cause data thefts. Hence such applications must be identified, in order that they're going to be identifiable play store users. So, we are proposing an android application which can process the knowledge, comments and three reviews of the appliance with language processing to offer results. So, it'll be easier to choose fraud application.

**Corresponding Author:- Divya Bhoyar**
Address**:-** Final Year Students, Department of Computer Engineering, Bapurao Deshmukh College of Engineering, Sevagram, Wardha.

## Literature Survey:-

**Mobsafe Forensic Analysis For Android Applications And Detection of Fraud Apps Using Cloudstack And Processing ;**

During this paper, Patil Rohini, Kale Pallavi, Jathade Pournima, Kudale Kucheta proposed that the bulk folks use android Mobile nowadays and also uses the play store capability normally. Play store provide great numbers of applications but unfortunately few of those applications are fraud. Such applications dose damage to phone and also could even be data thefts. Hence such applications must be marked, so as that they're going to be identifiable for play store users.So it'll be easier to decide on which application is fraud or not. Multiple applications is processed at a time with the net application. Also, User cannot always get correct or true reviews about the merchandise on internet. So, we'll check for quite 2 sites, for reviews of same product [3].

**Discovery of Ranking Fraud for Mobile Apps;**

In above paper, Hengshu Zhu, Hui Xiong, Yong Ge, and Enhong Chen propose that ranking fraud detection system for mobile Apps. Specifically, they first showed that ranking fraud happened in leading sessions and provided some way for mining leading sessions for each App from its historical ranking records. Then, they identified ranking based evidences, rating based evidences and review based evidences for detecting ranking fraud [2].

**Survey on Fraud Ranking in Mobile Apps;**

In this paper, Monali Zende, Aruna Gupta proposed that fraud is happen any time during the complete life cycle of app, that the identification of the precise time of fraud is required. due to the large number of mobile Apps, it's difficult to manually label ranking fraud for each App, so it is vital to automatically detect fraud without using any basic information. Mobile Apps aren't always ranked high within the leaderboard, but only in some leading events ranking that's fraud usually happens in leading sessions. Therefore, main target is to detect ranking fraud of mobile Apps within leading sessions [6].

**FairPlay Fraud and Malwares Detection in Google Play;**

In this paper, Mahmudur Rahman, Mizanur Rahman, Bogdan Carbunar and Duen Horng Chau proposed that FairPlay, a system to detect both fraudulent and malware Google Play apps. They study on a newly contributed Vol-3 Issue-2 2017 IJARIIE-ISSN(O)2395-4396 longitudinal app dataset, within which that that they had shown a high percentage of malware is involve in search rank fraud; both are accurately identified by FairPlay. additionally, they showed FairPlay's ability to urge many apps that have on Google Play's detection technology, including a replacement form of coercive fraud attack [7].

## Proposed System:-

The proposed ranking fraud detection system for mobile Apps showed that ranking fraud happened in leading sessions and provided a method for mining leading sessions for each App from its historical ranking records.An app with the higher ranking in the charts will have more downloads. Due to the huge number of mobile Apps, it is difficult to manually label ranking fraud for each App and due to the dynamic nature of chart rankings, it is not easy to identify and confirm the evidences.It is important to automatically detect fraud without using any basic information. Mobile Apps are not always ranked high in the leader board, but only in some leading events ranking that is fraud usually happens in leading sessions.Therefore, main target is to detect ranking fraud of mobile Apps within leading sessions.

The proposed system must detect fraudulent application, This system aims to identify the leading sessions of each App based on its historical ranking of records. The system represents the new novel approach for the development of a ranking fraud detection system for mobile apps. The system performs the aggregation of two suggestion to detect fraud apps. This method will offer considerable benefits and provides an opportunity to prevent fraudulent apps in the market.

The important modules include,
1. Rating Based Suggestions
2. Review Based Suggestions

Our proposed system as in Fig. 1 gives an overall flow of the process which is happening. It begins with the extraction of data that is the historical records of the applications and User details from the store.
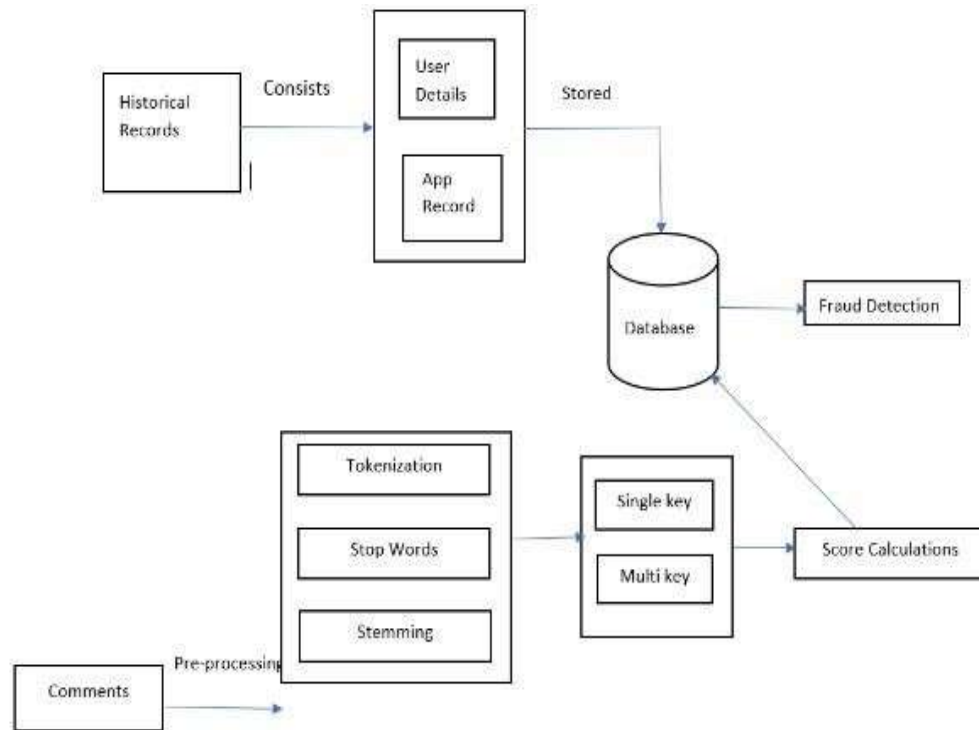
**Fig 1:-** System Architecture.

The admin adds an application to the database along with the rating details. From here, it will mine the leading session where it is calculated on the basis of evidence observed for that particular Application. For this, the mining leading session algorithm is used which is able to identify the leading session and events. After that, the evidences of rating, ranking and reviews are looked into one by one.

The system is analyzing the review involving user comments, and with the help of this comments we can determine wheather the app is genuine or not.  Due to the fast growth of usage of mobile devices, mobile apps are essential in day-to-day activities of most of the people. This system also detects the fraud application and ensure mobile security as well. To identify the fraud is the critical challenge in front of the mobile App market because there is a large number of mobile Apps. App developers are using delicate means more and more frequently for increasing their Apps sales. For mobile application, we develop review fraud detection system for this first discover the review fraud occur in leading session and provide a method for mining leading session. This system also detects the fraud application and ensure mobile security as well. For mobile application, we develop review fraud detection system for this first discover the review fraud occur in leading session and provide a method for mining leading session.

## Methodology:-

Fraudulent Application must be detected, as there is an increase in the number of mobile applications. This project aim is practical algorithm for identifying the leading sessions of each Application based on its historical ranking of records. With the analysis of ranking behaviours of Application, this system recognizes that the fraud that the fraudulent application often has different review patterns in their every leading session compared with usual application. Some fraud suggestion identifies from application historical ranking records resulting in the development of three functions to detect likewise review and rating based fraud suggestion. Moreover, two types of fraud suggestion based on Application ranking and review history are proposed. This project represents the approach for the development of a review and rating fraud detection system for mobile Application. Initially, a rating-based proposal is identified. Then, following the identification of review-based suggestions, ranking fraud suggestions are collected by leading mining sessions. And finally, the system performs the aggregation of all three suggestion to detect fraud Application. This method will offer considerable benefits and provides an opportunity to prevent fraudulent Application in the market. The important modules include,

The system comprises of 2 major modules,

**A. Rating based analysis**
**B. Review based analysis**

**A. Rating Based Evidences:**
After downloading app users generally rate the app. The rating given by the user is one of the important factors for the popularity of the app. An app having higher rating always attracts large number of users to download it and naturally it can also be ranked higher in the chart rankings. As a result, rating-based evidences are a significant aspect in app ranking fraud, and they must be evaluated. For that we are implementing here average rating formula that can help and generate result on the basis of ratings.

**B. Review Based Evidences:**
Along with rating users are allowed to write their reviews about the app. Such reviews demonstrate unique usage experiences for specific mobile Apps. The review given by the user is one of the most important factors for the popularity of the app. As the reviews are given in natural language so pre-processing of reviews is performed. The system will find sentiment of the review which can be positive/negative. If a review is positive, it adds one point to the positive score; if it is negative, it subtracts one point from the negative score. As a result, it will assess if the app is fraudulent or not based on the scores of each of the reviews.

**i ) Data Cleaning :**
Data cleaning is the process of fixing or removing incorrect, irrelevant, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.
Steps to clean the data :-
✦ Duplicate or irrelevant observations
✦ Fix structural errors
✦ Filter unwanted outliers

**ii ) Data Analysis :**
Data analysis is a process of inspecting, cleansing, transforming, and modelling data with the goal of discovering useful information, informing conclusions, and supporting decision-making.
Steps to clean the data:-
1. Checking the context
2. Pooling data from sources
3. Searching for data
4. Categorizing qualitative data

**iii ) Natural language processing (NLP) :**
NLP is a field that focuses on making natural human language usable by computer programs. A large portion of the data you may be examining is unstructured and comprises human-readable text. Natural Language Toolkit (NLTK) is a Python package.
Steps to Analyse by NLTK:-
1. Tokenization
2. Stop word removal
3. Stemming
4. Porter Stemmer algorithm

**iv) Text2Emotion:**
Emotion is a biological state associated with the nervous system brought on by neurophysiological changes variously associated with thoughts, feelings, behavioural responses, and a degree of pleasure or displeasure.

Steps to Analyse by Text2Emotion:-
1. Text Pre-Processing
2. Emotion Investigation
3. Emotion Analysis

With the help of NLTK and text2emotion package we separate emotions into 5 types and give score (0 to 1),
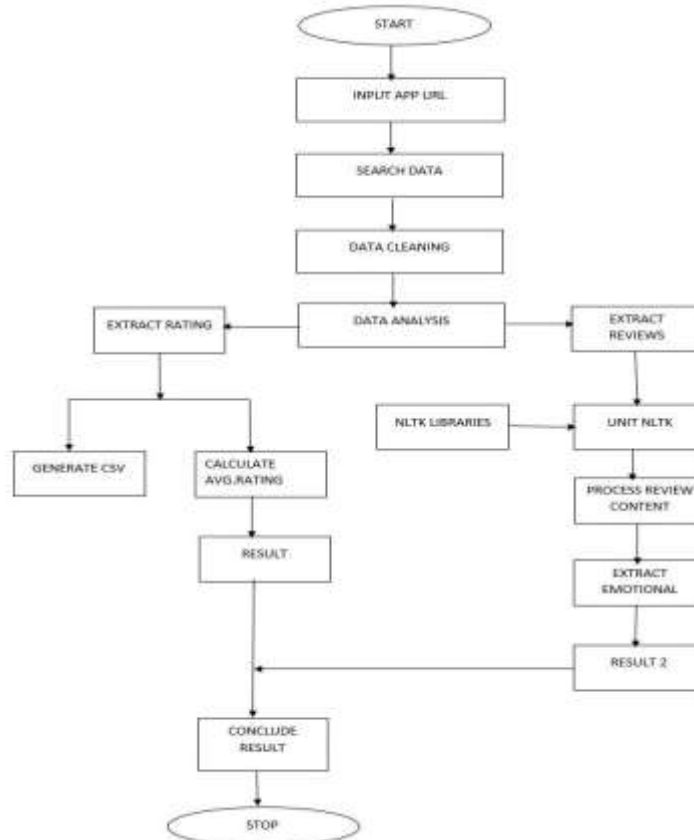1) Angry   2) Sad   3) Fair   4) Happy   5) Surprise

**Fig 2:-** System Flowchart.

Let's discuss the flowchart of system, Start the system and then enter the input app URL to search the application details from HTML page. Data cleaning is the process of correcting are Removing inaccurate records from the database. Data analysis helps us to read data from Sources file CSVs and SQL and now to us libraries like NumPy, pandas and many more. After data analysis the process divided into two Extract, 1st is Exact Rating, which is used to generate CSV and calculate average rating to find the result and then conclude the result. The 2nd is extract Reviews, by the use of NLTK library we can build python program that work with human language for the process Review content to Extract emotional and then also conclude the Result 2, Afterwards stop the system.

**Application :**
1. In this system we collect many datasets contains the reviews for various applications like Social, Games, Education, Finance, News, Food that seems user secured with their activity.
2. The manipulation of review is one of the key aspects of App ranking fraud. By using Natural language processing and data mining.
3. By analysis of reviews and ratings, system can help to determine the correct application for Android.

## Conclusion:-
After successful completion of our system.We conclude that this system might be able to detect the rating frauds supported two forms of evidences, like, rating based evidences and review based evidences. Further, an optimization based aggregation method combines all the two evidences to detect the fraud.

## References:-
1. E. Siegel, "Fake reviews in Google Play and Apple App Store," Appentive, Seattle, WA, USA, 2014.
2. Hengshu Zhu, Hui Xiong, Yong Ge, and Enhong Chen, "Discovery of Ranking Fraud for Mobile Apps", IEEE Transactions on Knowledge and Data Engineering, Vol.27, No. 1, January 2015.
3. Patil Rohini, Kale Pallavi, Jathade Pournima, Kudale Kucheta, Prof.Pankaj Agarkar, "MobSafe: Forensic Analysis for Android Applications and detection Of Fraud Apps Using CloudStack and Data Mining",

International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 Issue 10, October 2015.

4. Daniel Roberts, How to spot fake apps on the Google Playstore. Fortune, 2015.

5. Rejin Idiculla Johnson1, John Prakash 21 PG Scholar, Fraud Detection of Mobile App Department of Computer Science & Engineering, Mangalore Institute of technology and engineering, Mangalore, Karnataka (India) June 2016.

6. Monali Zende, Aruna Gupta, "Survey on Fraud Ranking in Mobile Apps", IJSR Volume 5 Issue 2, February 2016.

7. Mahmudur Rahman, Mizanur Rahman, Bogdan Carbunar and Duen Horng Chau, "FairPlay: Fraud and Malware Detection in Google Play" 2016.

8. Esther Nowroji., Vanitha., "Detection Of Fraud Ranking For Mobile App Using IP Address Recognition Technique", Vol. 4, International Journal for Research in Applied Science & Engineering Technology, 2016.

9. Vivek Pingale, Laxman Kuhile, Pratik Phapale, Pratik Sapkal, Prof. Swati Jaiswal, "Fraud Detection & Prevention of Mobile Apps using Optimal Aggregation Method", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 8, March 2016.

10. Jitendra Kumar Rout, Anmol Dalmia, Kim-Kwang Raymond Choo (Senior Member, IEEE), Sambit Bakshi1 (Member, IEEE), And Sanjay Kumar Jena (Senior Member, IEEE) Department of Computer Science and Engineering, National Institute of Technology Rourkela, Odisha 769 008, India, IEEE Trans. January 18, 2017.

11. sMahmudur Rahman, Mizanur Rahman, Bogdan Carbunar, and Duen Horng Chau, "Search Rank Fraud and Malware Detection in Google Play", IEEE Transactions on Knowledge and Data Engineering, Vol. 29, June 2017.

12. Tejaswini Shingare : Fraud Application Detection Using Data Mining Techniques, BE Students, Computer Engineering, SRES COE, Kopargaon, Maharashtra, India (IJARIIE-ISSN), 2017.

13. Iker Burguera, Urko Zurutuza, and Simin NadjmTehrani. Crow-droid: Behavior-Based Malware Detection System for Android. In Proceedings of ACM SPSM, 2011.

14. Ahmad Firdaus, Nor Badrul Anuar, Ahmad Karim, Mohd Faizal Ab Razak, "Discovering optimal features usin Mandava Firdaus, static analysis and a genetic search based method for Android malware detection" Frontiers of Information Techonology and Electronic Engineering, 2018.

15. Neha M. Puram, Kavita R. Singh, "Semantic Analysis of App Review for Fraud Detection using Fuzzy Logic", International Journal of Computer & Mathematical Sciences,Vol. 7, January 2018.

16. Sidharth Grover, "Malware detection: developing a system engineered fair play for enhancing the efficacy of stemming search rank fraud", International Journal of Technical Innovation in Modern Engineering & Science, Vol. 4, October 2018.

17. Rami Mohawesh : Fake Reviews Detection: A Survey, Computer Science Department, Jordan University of Science and Technology, Irbid 22110, Jordan IEEE Trans. April 26, 2021.

18. Gauri Rao, Detection of Fraud Apps using Sentiment Analysis, International Journal of Innovative Science and Research Technology, May - 2018.

19. Avayaprathambiha P, Bharathi M, Sathiyavani B, Jayaraj S "To Detect Fraud Ranking For Mobile Apps

20. Using SVM Classification" International Journal on Recent and Innovation Trends in Computing and Communication, Vol. 6, February 2018.

21. Vivek Pingale, Laxman Kuhile, Pratik Phapale, Pratik Sapkal, Prof. Swati Jaiswal, "Fraud Detection &

22. Prevention of Mobile Apps using Optimal Aggregation Method", International Journal of Advanced Research in Computer Science and Software Engineering,Vol. 8, March 2016.

23. Tahura Shaikh, Dr. Deepa Deshpande, "Feature Selection Methods in Sentiment Analysis and Sentiment

24. Classification of Amazon Product Reviews", International Journal of Computer Trends and Technology (IJCTT), Vol. 36, June 2016.

25. Mandava Rama Rao, Nandhini Kannan, V S Nihanth. Detecting Fraud Apps using Sentiment Research International Journal of Recent Technology and Engineering (IJRTE). ISSN: 2277-3878, Volume-8, Issue-

26. 2S3, July 2019.