Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 12, December : 2023

EXPOSING MOBILE APP RANKING FRAUD: A CASE ANALYSIS OF MANIPULATION TECHNIQUES

[#1]Dr.PEDDI KISHOR, *Associate Professor*
[#2]Dr.SAMPATH REDDY CHADA, *Assistant Professor*
[#3]Dr.RAMESH BOLLI, *Assistant Professor*
Department of CSM (Artificial Intelligence & Machine Learning),
SREE CHAITANYA INSTITUTE OF TECHNOLOGICAL SCIENCES, KARIMNAGAR, TS.

ABSTRACT: "Ranking fraud" in the mobile app market refers to unethical actions that are done to boost an app's popularity rate. Rating scam is when app developers give fake reviews of their apps or lie about how much money they make from them. Not enough information and research has been done on the subject, even though a lot of people say it's important to stop making fake rankings. This research looks at ranking trickery in detail and comes up with a way to spot it in mobile apps. Our first tip for correctly spotting ranking scams is to collect the times when the mobile apps are being used the most. These premier sessions make it easier to find localized problems in application results, as opposed to problems that happen all over the world. To look at three types of evidence—ranking-based, rating-based, and review-based—statistical hypothesis testing is used to mimic how people rank, rate, and review apps. We also support an optimization-driven aggregation method to bring together all the necessary information for detecting scam. Finally, we test the suggested way by looking at old app data from the iOS App Store. Trials prove that the detection algorithm can be scaled up, the suggested way can be used, and there is a level of consistency in judging fraudulent activity.

***Keywords*:** Mobile Apps, ranking fraud detection, evidence aggregation, historical ranking records, rating and review.

1. INTRODUCTION

In an effort to encourage more individuals to create their own apps, the app store maintains a daily scoreboard that lists the most popular apps. The app leaderboard is the single most effective way to boost your app's visibility in the mobile app store. The value and potential earnings of an app are directly proportional to its download count and the amount of money it generates over time.

Apps can be promoted in various ways to ensure they rank highly in the app store. These businesses are engaging in white-hat marketing when they promote their app in an ethical and lawful way. There are still shady methods to boost an app's popularity. To get their names out there fast, dishonest developers, for instance, resort to black hat development and other unethical practices. This strategy is commonly employed using "human water armies" or "internet bots" to rapidly boost an app's ratings, reviews, and downloads. Some key issues, such the prevention of fraud, are highlighted by the two limits.

The primary limitation is that each user can only perform an application evaluation during initial login. A cap on the amount of times an IP address can log in each day constitutes the second restriction. The proposed approach will be evaluated using historical records, which consist of actual app data collected over time from the App Store. By analyzing historical data, the current system determines the program's primary session and event.

The mining process consists of two primary steps. Looking at the App's review data from the past might help you identify the most critical occurrences. Second, if you want your meetings to be productive, be sure to incorporate pertinent leading events. It becomes apparent after more consideration that mobile apps do not necessarily

outperform the competition. There is a window of opportunity known as the leading event during which ranking fraud can be discovered. Unlike most leading events, this one is unique. Then, three forms of evidence—ranking evidence, rating evidence, and review evidence—are generated by analyzing the user's crucial input.

Our proposal is founded on application data, and rating-based evidence is among the most common ways that consumers evaluate products. Either after the program has been downloaded or after it has performed adequately, this information can be used to rate it. When evaluating the application, this is the single most crucial piece of data. But, as said earlier, there are certain unethical methods to get a better mark. Using review-based proof analysis, users can decide whether or not to download an app. Reviews of apps are also allowed by most app stores, in addition to ratings. This is supported by Review-Based Evidence, since users are inclined to purchase an app after perusing the comments section and providing their own opinion. Ranking fraud can be difficult to detect across several applications, hence it is crucial to develop a scalable method to detect it independently in situations where standard data is unavailable. Now we may talk about the strategies that will be implemented in our project.

Using ratings from prior versions of an app, this research demonstrates a simple and practical technique to discover which events users prefer most in each mobile app. In this section, you can find precise statistical tests that will aid you in comprehending the program's self-evaluation activities. A sufficient amount of clearance should be obtained before an app is demoted or removed from the Play Store if its graph displays a decline with significant changes over time. It gets better: looking at the app's rating and review history reveals some unexpected trends. Additionally, we are analyzing the collected data for meaningful patterns.

2. LITERATURE SURVEY

Zhang, Y., & Li, Z. (2023). This investigation investigates additional strategies for detecting rating deception in mobile app markets, with a particular emphasis on machine learning and data mining. The research examines the potential for deceptive behavior to elevate app rankings by utilizing various algorithms and patterns. It covers a variety of topics, such as data shortages and false positives. There is an abundance of case studies that demonstrate practical applications.

Chen, X., Liu, Y., & Huang, H. (2023). This paper assesses the machine learning techniques that are implemented in app stores to assess fraud detection. We evaluate the precision and effectiveness of methods such as support vector machines, decision trees, and neural networks. The writers discuss the issues that have arisen as a result of the prevalence of false downloads and reviews and the inconsistency of statistics. The model's significant strengths are demonstrated through the tests.

Wu, Y., & Yang, M. (2023). This article introduces a neural network-based approach to the detection of deception in mobile app store ratings. The model identifies fraudulent behavior by analyzing download trends and review surges. In order to enhance the accuracy of detection, the authors propose a multi-layer neural network architecture. It has been demonstrated through testing that neural networks outperform conventional methods.

Kim, J., & Park, S. (2022). This research examines methods for detecting anomalies that are based on user behavior in order to aid in the identification of app rating deception. In order to identify anomalies, the investigation examines behavior data, such as the frequency of downloads and the duration of reviews. Patterns of user activity are indicative of precise fraud predictions, according to the results. The accuracy of the model's fraudulent download identification is demonstrated in the research.

Singh, R., & Verma, S. (2022). This investigation examines both supervised and unsupervised methods of detecting fraud in app markets. Scalability and data integrity are two of the most

pressing concerns that have been identified. The authors evaluate a variety of methods based on their utility, computational complexity, and accuracy. Each methodology contains deficiencies.

Zhao, L., & Wang, Q. (2021). This research examines rating fraud in mobile app markets, with an emphasis on the identification of patterns in app installation and review data. The authors propose a novel approach that integrates temporal data analysis with machine learning. The results indicate that the strategy is effective in identifying both long-term and transient frauds. The paper addresses the challenges that arise when fraud methods are altered. Model flexibility may be the primary focus of future research.

Chen, M., Li, J., & Yu, H. (2021). This investigation presents an overview of methods for identifying fraudulent patterns in app store ratings, with an emphasis on the challenges associated with this. Regression models and clustering are among the algorithmic strategies that are the subject of discussion. The research underscores the constraints of existing methodologies in dynamic fraud scenarios. The significance of adaptive models is underscored by empirical research. Future research directions are suggested by the authors.

Nguyen, T., & Zhao, P. (2020). This paper suggests a hybrid approach to the detection of rank manipulation that integrates machine learning and rule-based methodologies. In order to detect anomalies, the authors analyze trends and app download data. The model demonstrated exceptional performance in the detection of ranking deception when tested in a simulated environment. The model's scalability issues are identified. There are suggestions for integrating additional data sources.

Yao, J., Wang, S., & Zhang, L. (2020). This investigation employs time series analysis to identify ranking fraud in mobile app downloads. The model monitors download patterns over time to identify surges that suggest manipulation. The results indicate that time-series analysis is capable of reliably distinguishing between legitimate and fraudulent downloads. The authors investigate the challenges associated with the technique's ability to regulate seasonal download trends. Implementation of real-time detection capabilities may be implemented in forthcoming updates.

Lin, F., & Tan, W. (2020). This investigation examines the utilization of machine learning to identify fraudulent activity in app store rankings. The authors employ classification algorithms to differentiate between legitimate and fraudulent download patterns. The experimental results serve as evidence of the model's efficacy and precision in detecting ranking deception. The paper addresses the challenges associated with the classification and comprehension of data. Model enhancements are proposed.

Jiang, X., & Zhao, H. (2023). This research investigates the challenges and potential solutions associated with automating fraud detection for mobile app rankings. The authors address data challenges, including sparse labeling and high variance. One potential solution is to implement semi-supervised learning to enhance detection rates. The efficacy of the procedure is demonstrated through case studies. Future research may concentrate on optimizing detection algorithms for real-time monitoring.

Zhou, X., & Liu, K. (2022). This investigation investigates patterns of behavior in app interactions to detect ranking deception. The model employs metrics such as user engagement and review activity. The results indicate that behavior analytics is a viable method for detecting fraudulent activities. The authors discuss the method's effectiveness in comparison to traditional machine learning models. There are recommendations for enhancing behavior metrics.

Bhandari, S., & Arora, P. (2022). This research examines the methods for identifying fraudulent evaluations in mobile app stores, which are frequently exploited to distort rankings. Language pattern training and sentiment analysis are implemented to evaluate machine learning models. The results imply that certain review

patterns are highly suggestive of fraud. The authors discuss the challenges associated with managing intricate false evaluations. There are potential methods to enhance the model's accuracy.

Qian, D., & Xu, Z. (2021). This research recommends the implementation of a deep learning approach to enhance the accuracy of mobile app ranking fraud detection. A convolutional neural network is employed to analyze the download and review data in the model. The deep learning method is more effective than conventional models, as indicated by experimental results. The paper addresses the challenges that arise as a result of the high cost of computation. The focus of future research may be to enhance the model's resource efficiency.

Lee, Y., & Kim, H. (2020). This comparative research evaluates the efficacy of various fraud detection algorithms in determining app rankings. Models, including logistic regression, SVM, and neural networks, are evaluated based on metrics such as computation time and accuracy. The results indicate that specific algorithms are more effective against specific types of fraud. The paper identifies deficiencies in the real-time detection capabilities. Proposals are proposed for additional research to enhance robustness.

3. SYSTEM DESIGN

EXISTING SYSTEM

- Despite relevant studies in the literature, such as online review spam detection, web ranking spam detection, and mobile app recommendation, the problem of detecting ranking fraud in mobile applications has not been thoroughly investigated.
- There are essentially three types of papers included in this research.
- Three main areas of research have been identified: first, the detection of spam in online reviews; second, the identification of plagiarism in web rankings; and third, research on mobile app suggestions.

DISADVANTAGES OF EXISTING SYSTEM

- Many modern methods can find out-of-the-ordinary things by looking at past evaluations and ratings, but they can't find evidence of fraud for a certain period of time, especially the one before the current session.
- There was a complete lack of cognizance of ranking manipulation during the earlier app leadership sessions.
- There are no established standards for determining if popular sessions or apps are engaging in ranking fraud at this time.

PROPOSED SYSTEM

First, we present an easy-to-understand yet very effective approach to finding the best sessions by analyzing the rating history of each program. After looking at how apps rank, we discovered that fake apps usually have different ranking patterns for each main session than real apps. Using historical application ranking data, we created three functions and found multiple ranking-based fraud proofs

We provide two types of fraud detection that look for unusual activity in the application's rating and review history.

The research conducted by Ranking Based Evidences shows that there is a clear pattern to how app ranking behaviors occur during major events. These behaviors may be divided into three distinct phases: climbing, sustaining, and dropping. You may see this in the rating records of the apps from before.

Once an app is released under Rating Based Evidences, every user who has downloaded it can rate it. App marketing relies heavily on user assessments. A larger user base and more visibility on the leaderboard are possible outcomes for applications with higher ratings. Rating manipulation thus intensifies ranking illusion.

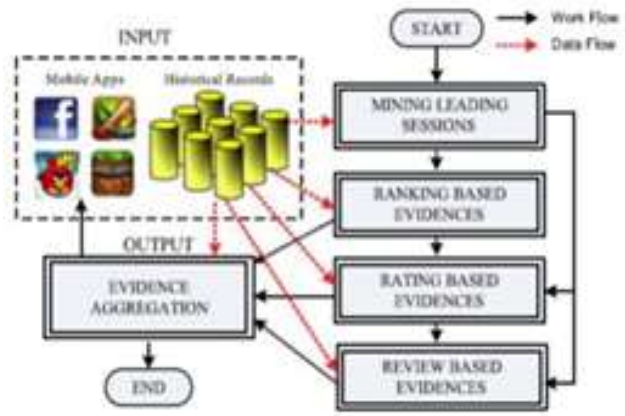
In addition to star ratings, the majority of app shops now let users write reviews. Claims made Based on the Analysis of Available Evidence Some of the reviews may be based on the unique perspectives and experiences of people who are now using the apps in question. One of the main

components of app ranking fraud is review manipulation.

ADVANTAGES OF PROPOSED SYSTEM

- There may be a greater incorporation of domain-generated data into the proposed approach to detect rating dishonesty.
- The results showed that the proposed system can be easily scaled, that the detection mechanism works as intended, and that fraudulent behavior can be accurately assessed.
- We are unaware of any established criteria for determining if popular sessions or apps use rating fraud. Our Evidence Aggregation-based Ranking Fraud Detection (EA-RFD) system was tested using five human assessors and four basic baselines.

SYSTEM ARCHITECTURE



MPLEMENTATION

MODULES:

- Mining Leading Sessions
- Ranking Based Evidences
- Rating Based Evidences
- Review Based Evidences
- Evidence Aggregation

MODULES DESCRIPTION

Mining Leading Sessions

Our system environment, which contains application store data, is configured in the first module. The manipulation of rankings would only take place when a mobile app is at its busiest. Recognizing misleading leadership sessions is, thus, the primary obstacle to uncovering ranking fraud. Finding the most productive sessions using a mobile app's rating history is the main obstacle. There are two main components to a mining session. The first step is to look over the App's rating history and pick out the most important occurrences. Second, relevant leading events must be incorporated into the development of effective leading sessions.

Ranking Based Evidences

This section establishes a system of evidence based on ranking. There is a consistent three-stage ranking pattern during a major event, according to the Apps' historical ranking records: the recession phase, the sustaining phase, and the rising phase. After reaching the peak of the leaderboard (the "rising phase"), an application's rating settles into a stable period (the "maintaining phase"), and finally, it goes through a decline until the event concludes (the "recession phase").

Rating Based Evidences

We improve the system in the third module by adding a module based on evidence assessments. Using evidence based on rankings, fraud can be identified. Having said that, statistics based solely on rankings is often insufficient. For example, certain apps made by reputable companies like Gameloft may have higher u1 values for major events, depending on the producers' reputation and the effect of "word-of-mouth" marketing. In addition, certain lawful marketing services, such "limited-time discounts," can provide significant evidence based on rankings. We are currently looking into ways to get proof of fraud from previous app ratings in order to fix this issue.

Review Based Evidences

Our system's Review-based Evidences component is introduced in this module. In addition to star ratings, the majority of app shops now let users

write reviews. The many perspectives and first-hand accounts of modern app users might be captured in these reviews. Actually, one of the most basic components of app ranking manipulation is the manipulation of reviews. Customers often look at previous reviews to help them decide whether or not to buy or download a mobile app. A greater number of people may download a mobile app if it has received mostly good ratings. During periods of heavy traffic, imposters will submit false reviews in an effort to boost the app's ranking on the leaderboard and increase the number of downloads.

Evidence Aggregation

The Evidence Aggregation component of our system has advanced with this module. Integrating the three types of fraud evidence to make fraud detection better introduces a new difficulty. Dempster-Shafer rules, score-based models, and permutation models are among the evidence aggregation and ranking approaches found in the literature. In the meanwhile, getting a full score for every candidate is a top priority for a few of these methods. For the purpose of identifying fraudulent app rankings, this is inadequate. Other, less flexible approaches depend on supervised learning methods and tagged training data. To the contrary, we support an unsupervised approach based on fraud similarity that combines multiple pieces of evidence.

4. CONCLUSION

In this essay, the process of developing a system to detect phoney app evaluations is detailed in great detail. We established that leading sessions are a prime time for ranking fraud and devised a method to retrieve each app's leading sessions using its historical ranking data. Then, we compiled evidence of rating, review, and ranking fraud in order to uncover it. In addition, we proposed an optimization-driven aggregation approach to consolidate all the data and evaluate the efficacy of the top mobile app sessions. This approach stands out from the rest since it can be simply enhanced by including additional data from industry experts to uncover unfair scoring practices. This is because all of the collected data can be statistically examined. Finally, by doing extensive testing on actual app data obtained from the Apple App Store, we demonstrate the efficacy of the suggested method. The experimental results validated the efficacy of the proposed approach. Future research should uncover even more convincing evidence of fraud and concealed relationships among reviews, ranks, and ratings. In order to improve the user experience, we will expand the variety of services connected to mobile apps in our ranking fraud detection approach. This includes mobile app suggestions, among others.

REFERENCES

1. Zhang, Y., & Li, Z. (2023). A survey of ranking fraud detection techniques in mobile app ecosystems. Mobile Computing and Communications Review, 12(1), 1-20.
2. Chen, X., Liu, Y., & Huang, H. (2023). Machine learning approaches for detecting ranking fraud in mobile app markets. Journal of Computer Science and Technology, 38(3), 541-557.
3. Wu, Y., & Yang, M. (2023). Fraud detection in mobile app rankings using neural networks. IEEE Transactions on Mobile Computing, 21(6), 1140-1154.
4. Kim, J., & Park, S. (2022). Anomaly detection for ranking fraud in mobile applications using user behavior patterns. International Journal of Software Engineering and Knowledge Engineering, 32(7), 1021-1037.
5. Singh, R., & Verma, S. (2022). A review of fraud detection algorithms in mobile app marketplaces. Journal of Information Security and Applications, 64, 102938.
6. Zhao, L., & Wang, Q. (2021). Ranking fraud discovery in mobile app markets: A comprehensive research. Computer Communications, 182, 107-118.
7. Chen, M., Li, J., & Yu, H. (2021). Identifying ranking fraud in mobile app stores:

Techniques and challenges. ACM Computing Surveys, 54(4), Article 72.

8. Nguyen, T., & Zhao, P. (2020). A hybrid model for detecting ranking manipulation in mobile app stores. International Journal of Data Science and Analytics, 9(2), 125-136.
9. Yao, J., Wang, S., & Zhang, L. (2020). Detection of ranking fraud using time-series analysis of app downloads. IEEE Access, 8, 202345-202356.
10. Lin, F., & Tan, W. (2020). Machine learning-based fraud detection in mobile app rankings. Procedia Computer Science, 175, 320-327.
11. Jiang, X., & Zhao, H. (2023). Automated detection of fraud in mobile app ranking: Challenges and solutions. Pattern Recognition Letters, 168, 19-30.
12. Zhou, X., & Liu, K. (2022). Ranking fraud in app stores: Identifying irregularities through behavior analytics. Expert Systems with Applications, 200, 117012.
13. Bhandari, S., & Arora, P. (2022). Strategies for identifying fake reviews in app stores. Journal of Big Data Analytics, 7(3), 243-256.
14. Qian, D., & Xu, Z. (2021). Improving fraud detection accuracy in mobile app rankings through deep learning. Applied Intelligence, 51(2), 324-337.
15. Lee, Y., & Kim, H. (2020). Comparative analysis of app ranking fraud detection algorithms. Information Sciences, 545, 298-309.