Mini Project 1

By: Anirudh Chaudhary, Quan Le, Bob LeBow, Richard More, Hao Nguyen

# Purpose of the paper

Google Play, formerly known as Android Market, initially released on October 22, 2008, and currently serves as the main app store for the Android platform. It offers millions of applications to over 145 countries. Currently, Google Play has 6 main categories: Movies & TV, Apps, Music, Books, Devices, and Entertainment. Ratings are based on a 5-star scale, and users can submit an optional review/comment (McIlroy et al., 2017). Reviews contain a username, profile picture, rating, review, and time stamp. All of the reviews are public and may get replies from the developer of that particular app.

I personally posted a comment for a movie with random characters, and the review does go through the submission process. I’m not sure whether Google developers will go in and delete it later or not. This shows that there is a possibility that the review section of Google Play might contain many spams. Users also have the option to edit or delete the ratings provided by them for an app at any point of time.

Users are not allowed to delete other users’ comments. Google allows users to check out the historical record of a review's edits, like a comment, report the comments as unhelpful or spam, and can “link to the review” which pin a review to the top of the review section. There is no option to dislike a comment or reply to a comment.

According to the Google Play policy for comment posting, users are required to write an accurate review that does not include any confidential information, abusive and profanity. For developers, they are not allowed to manipulate the review system by providing the users some incentives such as discounts to get a high rating. Some developers have built in-app reviews, but Google Play senses the violations.

Review system in the past calculated stars based on user reviews. Hence, the stars were based on users' reviews over the entire lifetime of an app. Currently, the reviews are based on the most recent comments. In 2019, Google Play changes the policy that allows the developers to reset the rating for every new update. According to Milena Nikolic, an engineering director leading [Google](https://crunchbase.com/organization/google) Play Console, the average rating for Google Play apps will be recalculated and will give weight to the most recent user’s rating so the rating will be based on the current version of the app.

According to Google, the app store’s main purposes are allowing a communication bridge with the developers and providing parents to have an insight on potentially objectionable content that exists within an app.

# Related literature

This review and star-based rating system (McIlroy et al., 2017) sounds good, but does it work as it should? Well, based on a study led by Rahul Aralikatte, this combination causes some inconsistency between the star rating and the sentiment of the comments. One example from one of Instagram's reviews saying “Love instagram it’s the best in the world Love it it’s the best in the world” (sic) leading one to think the rating is 5 stars. That is not the case here. This reviewer gave only one star. This case is not unique, based on their study and investigation about 20% of the review-star combinations are inconsistent. (Aralikatte et al., 2018)

The real reason why there are frequent inconsistencies between ratings and reviews is not clear. Many researchers believe this mismatch occurs because of a lack of context within the rating system. Users leaving one star are often bitter about their experience and want to leave a sarcastically positive comment with a one star review, (such as the previous example) making a textual sensitivity analysis inaccurate. Another theorized reason is the ability for users to change their reviews after they receive a response from the developer. About 39% of the users change their reviews after they receive a comment from the developers of the app. (McIlroy et al., 2017) A possible mismatch occurs when the user changes their sentiment about the app but only updates the comment not the star rating or vice versa. In the end, these poorly interpreted reviews contribute to the 20% of inconsistent review-star combinations.

Research done by DPhil student Kerstin Frie in the Nuffield Department of Primary Care Health Sciences at the University of Oxford provides insight on how effective the Google Play rating system is for sentiment analysis on apps. (Frie et al., 2017) In her and her team's research, she studied weight loss apps. The premise of the research is that due to the long process of weight loss, developers must ensure user satisfaction towards their apps; users will not use their app for months if they are dissatisfied. Earning positive sentiment and therefore positive reviews is pertaninant in this app category. In her analysis, she discovered the apps with the best behavioural support for weight loss earned the highest rating score. We took this to mean the app review can be influenced by more than just the content of the app but also by these behavioural prompts.

Another research suggests a different approach for the rating of the apps. Instead of just using a star rating as the overall score, this paper suggests a unified rating system that would use the star rating in combination with a numeric polarity generated by a sentiment analysis score of the review to generate a final rating for the app. This would remove the ambiguity between the star-rating and the review’s suggested rating. (Islam, 2014)

In an article written by Duc Cuong Nguyen at Saarland University, he and his team talk about the impact of user reviews on security and privacy. (Nguyen et al., 2019) After thorough investigation, the team found that mentioning security and privacy concerns in the app review prompted the developers to update the security of the app 60% of the time. The team further asserts, significantly more users make comments on security when the app asks for runtime permissions, implying that this request puts security in the mind of the users. Once the users are already thinking about security, they comment on the app which often leads to a security update. The second portion of Nguyen's research finds that half of the updates by developers on apps come from third party libraries meaning that the app developers have trouble meeting customer and regulatory expectations for privacy and security. The implication of this research is that users can use this review platform as a direct communication method with developers to show them that users have a vested interest in their security and privacy.

# Research problem

One problem that not only Google Play but also other reviewing systems are struggling is the identification of the spam reviews. Currently, Google Play allows its users to report a review as spam. However, not many users pay attention if a review is spam or not as they do not have any incentive to report spam. Users tend to spend time with reviews that they found helpful instead of finding and reporting spam reviews. In addition, some spam reviews can be well-written by a customer account who has an adequate profile, making customer reported spam an ineffective evaluation method by itself. There are also multiple types of spam that can be easy to notice like an advertisement or some more difficult to spot like a fake review. Each type of spam requires different levels of analysis in order to categorize it properly. Fortunately, based on account activity, account length, frequency of bad/good review, and the number of helpful/unhelpful click, Google can find an automatic way to filter all spam reviews.

## 1. Account Activity

Account Activity is the first factor that can be used to decide if the review comes from an account that is valid or not. Account Activity records the most recent clicked apps of an account which allows account users to find out which apps they already visited. In fact, if an account is not having many app clicks in a short period while providing the only review for each app it visited, the account was likely created in order to provide fake reviews. If the activity record of an account shows that it is visiting all the same types of apps and provides similar messages in reviews, then that Google Play account is likely serving advertisement purposes.

## 2. Account Length

Account Length is another important factor to decide if the review comes from a spam account or not. Normally, a fake review was created by an account that is no longer than a month old. For example, when a game app receives so many bad reviews about its content or gameplay, the app owner can create several new Google Play accounts to give his or her app several high stars and provide good reviews. As a result, those new accounts will only visit that gaming app and have no further activity since the time of posting the review. Similarly, when a company wants to lower the rating of its competitor’s app, the company can create a new account to provide negative reviews and low star ratings to its competitor’s app. All of these new accounts will be soon forgotten and remain inactive for a long time after completing their purposes.

## 3. Frequency of Bad/Good Reviews

One common way to find out if an account is used intentionally for review spam purposes is its frequency of bad or good reviews to specific apps. In fact, when a normal person gives reviews to some apps, the number of good/bad reviews will be likely equal or 40/60 proportion. Hence, if an account gives 100% good reviews or 100% bad reviews for several apps, it has likely been created purposefully for reviewing apps rather than using apps. Sometimes, when a spam account creates a fake review in apps, it is likely using the same content of review for all apps. Therefore, if 100% of good reviews have repeated the same content for many different apps, it can be concluded to be a review made by a bot.

## 4. Number of Helpful/Unhelpful Click

A helpful or unhelpful click can help identify the quality of a review. For instance, if all the good reviews of an account have content as “Great app! I highly recommend this app to everyone”, the content is not very helpful for other users as it is not detailed enough to give them an idea about how good or bad of an app it is. It won’t receive many helpful clicks from other users as well. As a result, the review must have been created to intentionally improve the rating of the targeted apps.

# Methodology to Address Problem

It is not possible to identify with 100% accuracy whether a review is good or bad. Although, processing the identified data in a specific manner and scrutinizing it deeply by putting it through various types of analysis can make the review system more reliable and meaningful. As the reviews are not a result of some multiple-choice objective questionnaire, this task requires the introduction of subjectivity combined with efforts of analytic algorithms and human intelligence.

The analysis of the review content as well as its source i.e. the reviewer profile, can give us meaningful insights to segregate the spam/fake reviews from genuine ones. Therefore, identifying genuine users becomes an integral part of identifying the genuine reviews. One important thing that should be kept in mind while designing the segregation algorithms is that only actual spam reviews are categorized as spam/fake. This will make sure that the actual experiences of the users are reflected through the reviews of an app.

The following steps needs to be processed for the analysis of the collected data in order to segregate fake/spam reviews from the genuine reviews:

* For **Account Activity**, initially we need to list down all the users who have put their reviews for a targeted app. For each user, the *Number of Apps Installed* and the *Categories of Apps Installed* needs to be identified. If the only installed app is the one for which it has put down the review, then that user’s review will be categorized as fake. Another criterion, if the user has installed multiple apps but only of a specific Category like ‘Medical’ and has posted reviews with the same content as promoting a particular entity then also it will be categorized as spam. Such users mostly use the review platform to advertise a specific commodity.
* For **Account Length**, initially we need to list down all the users who have put their reviews for a targeted app. Then, we need to list down the *Duration of Existence* of the user along with the *Number of Reviews* from that user. If the duration is less than a week from the review date and the number of reviews put down by that user in Google Play Store is alarmingly high (100+) then that user’s review will be categorized as potential fake/spam.
* For **Frequency of Review type**, we need to list down all the users who have put their reviews for targeted apps. For each user, the *Number of Reviews* and *Review Type* needs to be identified. Based on the usage of words which the user has used in his review, it will be categorized as a positive or negative review along with the star-rating associated to the review. If for a user, the number of reviews is 100+ and all the reviews get categorized as positive or negative then it raises a flag of it being a potential spam review.
* For **Length of Negative Reviews**, we find out the *Number of Words* used while writing a review by a user. *Review Type* needs to be identified too as positive or negative based on the usage of words in the review along with the star-rating associated to the review. If for a negative review with one-star rating, the number of words used is less than five then it is potentially a fake review. A user who is unsatisfied after using an app takes out his or her time to write a review rather than just uninstalling the app usually won’t be using just 3-4 words to describe its experience.
* For **Spam Reported by User**, we need to list all the reviews for a targeted app. Then we need to list down the *Count of Spam Reports* for all the reviews. A review can be reported as spam by clicking the link – ‘Flag as spam’ in the Google Play Store. If the count of spam reports is more than five, the review will be categorized as actual spam. It should not be categorized as spam for less five reports to maintain the impartiality of the review system as the developer of an app can be motivated to mark bad review as spam or the counterparts can be motivated to mark good reviews as spam.

# Sources

McIlroy, S., Shang, W., Ali, N. and Hassan, A. (2017). Is It Worth Responding to Reviews? Studying the Top Free Apps in Google Play. *IEEE Software*, [online] 34(3), pp.64-71. Available at: https://ieeexplore.ieee.org/document/7325189 [Accessed 23 Oct. 2019].

Islam, M. (2014). Numeric rating of Apps on Google Play Store by sentiment analysis on user reviews. *2014 International Conference on Electrical Engineering and Information & Communication Technology*. [online] Available at: https://ieeexplore.ieee.org/abstract/document/6919058 [Accessed 23 Oct. 2019].

Aralikatte, R., Sridhara, G., Gantayat, N. and Mani, S. (2018). Fault in your stars. *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data - CoDS-COMAD '18*. [online] Available at: https://arxiv.org/abs/1708.04968 [Accessed 23 Oct. 2019].

Frie, K., Hartmann-Boyce, J., Jebb, S., Albury, C., Nourse, R. and Aveyard, P. (2017). Insights From Google Play Store User Reviews for the Development of Weight Loss Apps: Mixed-Method Analysis. *JMIR mHealth and uHealth*, [online] 5(12), p.e203. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5756316/ [Accessed 23 Oct. 2019].

Nguyen, D., Derr, E., Backes, M. and Bugiel, S. (2019). Short Text, Large Effect: Measuring the Impact of User Reviews on Android App Security & Privacy. 2019 IEEE Symposium on Security and Privacy (SP). [online] Available at: https://nguyenduccuong.de/papers/userreviews\_impact\_sp2019.pdf [Accessed 23 Oct. 2019].

Play.google.com. (2019). *Comment posting policy*. [online] Available at: https://play.google.com/about/comment-posting-policy/ [Accessed 23 Oct. 2019].

Play.google.com. (2019). *User Ratings, Reviews, and Installs | Store Listing and Promotion - Developer Policy Center*. [online] Available at: https://play.google.com/about/storelisting-promotional/ratings-reviews-installs/ [Accessed 23 Oct. 2019].