DroidRP: A reputation system for Android apps and app developers - Final Report

CREU Final Reports 2015

Student Researchers: Theresa Guinard
Randall Van Why

Advisor: Jun Zheng

Institution: New Mexico Institute of Mining and Technology


Goals of the Project

The aim of this project is to develop a system that uses sentiment analysis to justify the quality of an Android application (app) based on its reviews and then evaluate the reputation of an app developer. Smartphones and mobile devices are becoming more popular mainly due to the large number of apps in various app markets such as Apple’s App Store and Google’s Play Store. Since an app could be developed by anyone from professional developers to programmers with very limited experience, it’s a big challenge for users to pick a high quality app from a large pool of candidates. The 5-star rating systems used in most app markets are too coarse to accurately describe app quality. On the other hand, it’s difficult for users to sift through reviews in order to determine how other users feel about an app. The goal of this work is to provide users with a better understanding of the reception of an app.

Process

Over 500 apps were collected from the Google Play Store. In the first phase of data collection, online lists of Top 100 Free and Paid apps were scraped and information of each app was collected. From there, we were able to create a list of developers from the apps already collected. In phase two, we iterated through the list of developers, collecting every app that a developer had written. The results of phases one and two were combined to form the final data set.

We implemented and compared two sentiment analysis approaches: SentiWordNet and sentiment classification using support vector machine (SVM).

SentiWordNet is a dataset that assigns a sentiment score to members of WordNet (a commonly used English lexical database). To obtain the sentiment score of an app, part of speech tagging was performed on each review of the app. Then the sum of SentiWordNet scores for adverbs, verbs, and adjectives was taken to determine whether the review was positive, negative, or neutral overall. Each review was assigned a sentiment score of 1, -1, or 0 (positive, negative, or neutral). The sentiment score of the app was calculated by the proportion of positive to negative reviews.
In the SVM approach, adjectives, adverbs, and verbs were considered, which were again determined using part of speech tagging. These words were then lemmatized using synonym sets defined by WordNet. The comments were then normalized via term frequency-inverse document frequency (tf-idf). To reduce dimensionality, we used latent semantic analysis (LSA) which is based on singular value decomposition. This process formed the input of the SVM. The SVM was trained using a subset of one-star and five-star comments from the set of all reviews. The sentiment score of the rest of the reviews was determined based on SVM output. The sentiment score of each review was binary: 1 denoted a positive review, and -1 denoted a negative review.

**Conclusions and Results**

It was found that the SVM approach was more accurate at classifying one-star and five-star comments in agreement with the star rating. The accuracy of the SVM approach was 0.735 while that of the SentiWordNet approach was 0.709. However, the sentiment score obtained by the SentiWordNet approach correlated with the average star rating better than the sentiment score obtained by the SVM approach (Pearson correlation coefficient, SentiWordNet: r = 0.49, SVM: r = 0.07).

We also used the sentiment score to evaluate the validity of Google Play’s “Top Developer” badge. For each developer in our dataset, we assigned each of the developer’s apps a sentiment score based on the SentiWordNet approach. The average sentiment score was used as the developer’s sentiment score. It was found that the average sentiment score of developers with Top Developer badge is significantly lower than that of other developers without badge. The results show that being a “Top Developer” does not indicate the developer can develop apps of good quality.

**Presentations and Publications**
