A COMPREHENSIVE RANKING AND SELECTION OF APPLICATIONS

Nahida Sultana Chowdhury, Rajeev R. Raje
Department of Computer & Information Science, School of Science

Objectives

Traditional Feedbacks (# of downloads, # of installs/updates, # of ratings, Average rating score, user reviews) are not suitable!!!
- Self-selection bias
- Poorly written reviews
- Flood app stores with positive reviews before even launching their apps.
- Developers urge users and friends to downgrade competing Apps.
- Changes to an App Store can kill the review scores.

Internal Evidence based Trust Rating and Ranking (TRR) algorithm

Input: Evidences generated by FindBugs
For each App:
  - For each Evidence (1 to 20):
    - Calculate b, D, U, t tuples using formulae 1 to 3
  - Aggregate trust tuples by the use of the weighted consensus operator and generate a single b, D, U, t tuple
Use ordering operator to generate the app rating and normalize the rating to be out of 5
Based on the rating order them 1 to N
Output: Ordered Ranking based on internal evidence-based rating Apps

Results

Four different kinds of rankings Schemas:
I. Ranking based on average star rating of collected reviews
II. Ranking based only on the trust tuples computed for the internal view
III. Ranking based only on the trust tuples associated with the external view
IV. Ranking based on the tuples generated after combing the internal and external views

<table>
<thead>
<tr>
<th>App Categories</th>
<th>External &amp; Average rating (Distance %)</th>
<th>Internal &amp; Average rating (Distance %)</th>
<th>Internal &amp; External (Distance %)</th>
<th>Combined ranking &amp; Average rating (Distance %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication</td>
<td>9%</td>
<td>56%</td>
<td>56%</td>
<td>51%</td>
</tr>
<tr>
<td>Photography</td>
<td>11%</td>
<td>48%</td>
<td>51%</td>
<td>43%</td>
</tr>
<tr>
<td>Entertainment</td>
<td>0%</td>
<td>24%</td>
<td>24%</td>
<td>19%</td>
</tr>
<tr>
<td>Games</td>
<td>16%</td>
<td>31%</td>
<td>29%</td>
<td>29%</td>
</tr>
<tr>
<td>Tools</td>
<td>14%</td>
<td>33%</td>
<td>36%</td>
<td>28%</td>
</tr>
</tbody>
</table>

 Proposed Approach

Dataset Details

- Target Distributed Platform - Google Play Store
- 250 Apps from 5 different categories
[dataset is available here: https://github.com/nahida-aan/google-play-store-app-reviews-dataset]

- Communication – Messenger
- Entertainment – Fake call
- Games – 2048 Puzzle
- Photography – Photo Editor
- Tools – Scientific Calculator

Sentiment score of User’s Reviews Map to External Evidences

Findings

- Our experiments indicate that proposed combined ranking that encompasses both the internal and external views is a better alternative than any ranking that focuses only on the internal or external view.

Future Investigation

To incorporate different evidences (e.g., collusion checking, Ad presence checking, system specification, number of downloads, user and developers’ reputation) to generate more meaningful evidences.

References


https://www.cs.iupui.edu/~nschowdh/TRR/