Selecting a Characteristic Set of Reviews

Theodoros Lappas
Evimaria Terzi
Mark Crovella

Boston University
CUSTOMER REVIEWS

Major impact on the users’ purchase decisions

Crucial part of e-commerce, reviews on all types of products

Hundreds or thousands of reviews available on a single product.

Too many for the user to parse

High redundancy: same opinions on the same attributes

Varying Informativeness

User gets overwhelmed or misinformed
CURRENT SOLUTIONS

**Ranking** (e.g. date, helpfulness)

- Ignore *complementarity* among reviews
- Review-voting mechanisms are **biased**

**Statistical Summarization**

- Extract opinions from reviews
- Report a frequency-based summary of opinions
- Takes the narrative out of the picture
- Unintuitive. Users want to read real reviews, weary of “black box” approaches and numbers
- Demotivates reviewers
OUR APPROACH: REVIEW SELECTION

Select a compact and informative set of reviews, that accurately represents the entire corpus

- Small enough for the user to parse
- User friendly: users are shown actual reviews
- The selection mechanism can be reused to generate more sets at will
- Keeps reviews visible, motivates reviewers to submit high-quality content
CURRENT SELECTION METHODS

Make sure there is **at least one positive and at least one negative opinion for each attribute** in the set. [Tsaparas et al. - KDD 2011]

- ❌ Ignores the frequency of each opinion
- ❌ Does not represent the corpus, can be misleading to users

Make sure **only the most frequent** (positive or negative) **opinion on each attribute** is represented in this set. [Lappas and Gunopulos – PKDD 2010]

- ❌ Ignores minority opinion, which may still be significant
- ❌ Provides one-sided view of each attribute
- ❌ Does not represent the corpus, can be misleading to users
A MOTIVATING EXAMPLE

6 REVIEWS IN THE CORPUS

3 features:
• $f_1 \rightarrow 4/6^+ 2/6^-$
• $f_2 \rightarrow 4/6^+ 0/6^-$
• $f_3 \rightarrow 2/6^+ 4/6^-$

We want to select the 3 reviews that best represent the entire collection

{R4, R5, R6}:
• $f_1 \rightarrow 2/3^+ 1/3^-$
• $f_2 \rightarrow 2/3^+ 0/3^-$
• $f_3 \rightarrow 1/3^+ 2/3^-$

✓ Same ratio for all attributes
✓ Accurate & Compact Representation of the corpus
# THE ALGORITHMS

## THREE DIFFERENT SELECTION ALGORITHMS

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Greedy:</strong></td>
<td>Minimizes the incurred error at every step</td>
</tr>
<tr>
<td></td>
<td>- <strong>Myopic, may force itself to dead end</strong></td>
</tr>
<tr>
<td><strong>Iterative Random:</strong></td>
<td>Draw large number of random sample, choose the one with minimum error</td>
</tr>
<tr>
<td></td>
<td>- <strong>Unpredictable, may take too long to get a good solution</strong></td>
</tr>
<tr>
<td></td>
<td>- <strong>No way to know if there is a better solution</strong></td>
</tr>
<tr>
<td><strong>Integer Regression:</strong></td>
<td>Solves a continuous version of the problem via regression. It then transforms the continuous solution into the closest discrete one.</td>
</tr>
</tbody>
</table>
EXPERIMENTAL EVALUATION

6 Review Datasets from Amazon.com

- 233 Cameras
- 238 MP3 Players
- 62 Vacuum Cleaners
- 11,447 Books
- 40 Coffee Makers
- 112 Printers
## AVERAGE ERROR

<table>
<thead>
<tr>
<th></th>
<th>CAM</th>
<th>k = 5</th>
<th>k = 10</th>
<th>k = 15</th>
<th>k = 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greedy</td>
<td></td>
<td>0.1</td>
<td>0.01</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Integer-Regression</td>
<td></td>
<td>0.09</td>
<td>0.01</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Iterative-Random</td>
<td></td>
<td>0.03</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>COF</th>
<th>k = 5</th>
<th>k = 10</th>
<th>k = 15</th>
<th>k = 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greedy</td>
<td></td>
<td>0.14</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Integer-Regression</td>
<td></td>
<td>0.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Iterative-Random</td>
<td></td>
<td>0.06</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>BOOK</th>
<th>k = 5</th>
<th>k = 10</th>
<th>k = 15</th>
<th>k = 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greedy</td>
<td></td>
<td>0.3</td>
<td>0.03</td>
<td>0.01</td>
<td>0.0</td>
</tr>
<tr>
<td>Integer-Regression</td>
<td></td>
<td>0.27</td>
<td>0.05</td>
<td>0.02</td>
<td>0.0</td>
</tr>
<tr>
<td>Iterative-Random</td>
<td></td>
<td>0.3</td>
<td>0.05</td>
<td>0.013</td>
<td>0.0</td>
</tr>
</tbody>
</table>
A USER STUDY

- Ask users to rank the sets of reviews selected by:
  - Our approach
  - Helpfulness (top-k)
  - GroupCover [Tsaparas et al. KDD 2011]

- 10 different items (MP3 players)
- 40 annotators
- Report Avg. ranking per item

✓ Annotators are actually shown the frequency of each opinion in the corpus
USER STUDY - RESULTS

![Graph showing average rank for different items and categories]

- Integer-Regression
- Helpfulness
- GroupCover

Average Rank

Items 1 to 10
SUMMING UP

- A new review-selection paradigm
- Accurate representation of the opinion distribution in the corpus
- Superior to state-of-the-art for selection
THANK YOU!