An improved analysis support system of open-ended questionnaire data using category-based dictionary

Norihisa Komoda, Keisuke Negoro, Masanori Akiyoshi,
Department of Multimedia Engineering
Graduate School of Information Science and Technology
Osaka University, Japan

Hiroaki Oiso
Codetoys, Japan
Contents

1. Background of the research
   “Who Wants to be a Millionaire?”

2. Unsubscribing questionnaire

3. Outline of the support system
   Conventional classification method
   New classification method
   - a method using category-based dictionary -
   User interface

4. Analysis examples

5. Conclusions
1. Background of the research(1)

- Target: Mobile quiz game content
  (monthly subscribe-type business)
  "Who Wants to be a Millionaire?"

- broadcasted on TV more than 56 countries.
- From 2001, Codetoys provides the mobile game
  in the world-wide under the license from Celador, UK.

- Established in 1998 in Finland

- In Japan, a web base mobile service started from 2002
  through the 3 domestic carriers.

  from 2002.4 au from 2002.10 vodafone
1. Background of the research(2)

- A content provider only gets a constant monthly content fee. (Carriers get packet charge.)
- To increase profit \( \rightarrow \) Prolong users’ subscription period
  Content provider needs to know why users unsubscribe.

- An analysis support system for grasping users’ opinions has been developed.
  + classification of opinions to several typical categories and atypical opinion category
  + user interfaces for typical and atypical opinions
1. Background of the research (3)

Problems of the conventional system

(1) Classification performance is insufficient.

- Needs many selected pairs of typical words
- Includes many typical opinions

(2) Setting and adjusting a dictionary DB is complex.

In this presentation, for solving above problems, a new classification method is proposed.
Questionnaire is given to customers who unsubscribe from the game.

2.1 Unsubscribing questionnaire

Open-ended questions
- To acquire unexpected ideas
- 2 questions
  (a) Demands of improvement
  (b) Other opinions
- The writing rate of opinions is around 10%.

Multiple-choice questions
- Asking reasons for unsubscribing
- 11 items
  ex) I could not win a prize.

The provider must read all answers to analyze open-ended questions.

Most opinions of open-ended questions reflect opinions that already known by the provider or duplicate the meaning of multiple-choice questions.

The provider wants to treat typical opinions statistically without reading them, and to read unexpected opinions.
## 2.2 Answers of open-ended questions

Opinions are categorized as 2 types.

<table>
<thead>
<tr>
<th>Type</th>
<th>Typical opinions</th>
<th>Atypical opinions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong></td>
<td>• Opinions having the same meaning as items of multiple-choice questions</td>
<td>• Opinions that are not typical</td>
</tr>
<tr>
<td></td>
<td>• Frequent opinions that the provider has already known</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Irrelevant opinions</td>
<td></td>
</tr>
<tr>
<td><strong>Analyst requirement</strong></td>
<td>◆ Unnecessary to be read all</td>
<td>◆ Necessary to be read</td>
</tr>
<tr>
<td></td>
<td>◆ Output of the number as graphs</td>
<td>◆ Output to find necessary opinions easily</td>
</tr>
</tbody>
</table>

Questionnaire analysis support system needs these functions:

- Classification into typical or atypical opinions,
- User interface that shows graphs about typical opinions,
- User interface that supports to read atypical opinions easily.
3. Outline of the support system

- Typical word DB reflects analyst’s background knowledge.
- Categories are defined beforehand.

Pre-process of Input Data
- Extracted open-ended questions
- Morphological analysis
- Word lists

Classification of opinions
- Graph represents the relationship between the number in each category and the unsubscription date.

Typical word DB
- Reflects analyst’s background knowledge.

Addition of typical word
- When analyst wants to move atypical opinion to typical opinion, he adds specific words in typical word DB.

Opinions are placed on the screen as cards.

- Quizzes, difficult, …
- Fun, children, …

Unsubscription date
3.2 Conventional classification method (1/3)

- Opinion meanings are grasped with word combination in typical word DB.

- Moreover, if a word combination included in an opinion that is the same as a word combination in the typical word DB, it is called typical element.

  ex) quiz, interesting, sorry, unsubscribe

- The longer the distance between the word combination, the weaker their relation.

  \[ d \]: The maximum distance can be considered to be typical element

  \[ word\ distance=1 \leq d (=2) \Rightarrow \text{The word combination is typical element.} \]

  ex) \( d = 2 \)

  \[ word\ distance=3 > d (=2) \Rightarrow \text{The word combination isn't typical element.} \]
3.2 Conventional classification method (2/3)

- This condition formula is based on the following concept.
  - If the ratio of typical elements to all keywords in an opinion is over a certain ratio, the opinion is classified typical.

**Condition formula for typical opinion**

\[
\text{(The number of keywords in typical elements)} + \alpha \times (\text{The number of typical elements}) > \text{(The number of all keywords)}
\]

**Example**

\[
2 \text{ typical keywords} + \alpha (=2) \times 2 \text{ typical elements} > 4 \text{ keywords}
\]

\[
2 \text{ typical keywords} + \alpha (=2) \times 1 \text{ typical element} < 6 \text{ keywords}
\]

- typical element:
  - packet fee, high, keep, unsubscribe
  - quiz, fun, keep, unsubscribe, category, better

- typical:

- atypical:
3.2 Conventional classification method (3/3)

- Two data sets from different carriers’ users were tested.
  - Data 1: 3263 opinions → 270 atypical opinions
  - Data 2: 1764 opinions → 175 atypical opinions (by the provider)
- About 8000 word pairs extended from Data 1 in *Typical word database*

<table>
<thead>
<tr>
<th></th>
<th>Number of opinions</th>
<th>Class by proposed method</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Typical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data 1</td>
<td>3263</td>
<td>2947</td>
<td></td>
<td>85.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>316</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Data 2</td>
<td>1764</td>
<td>1480</td>
<td>87.4</td>
<td>53.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>284</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\text{Recall} = \frac{\text{Atypical opinions correctly extracted by the method}}{\text{Atypical opinions originally identified by the provider}}
\]

\[
\text{Precision} = \frac{\text{Atypical opinions correctly extracted by the method}}{\text{Atypical opinions extracted by the method}} = \frac{270}{316}
\]

- Leakage of atypical opinions extraction was small.
- However, many typical opinions were classified to atypical opinions. Moreover, setting and adjusting typical word DB is complex.
3.3 New classification method (1/5)

Approach

1. Small set of opinions are classified by an analyst.
2. “Similarity” between categorized opinions and a target opinion are tested from two viewpoints:
   (a) Ratio of included representative words of a category (RW),
   (b) Coincidence of used words of the target opinion and words in all opinions in the category (CW).
3. If both similarity values of a category are larger than respective thresholds, the target opinion is classified to the corresponding category.

All set of similarity values are less than the thresholds, the target opinion is judged an atypical opinion.
3.3 New classification method (2/5)

(a) Ratio of included representative words of a category (RW)

\[
RW = \frac{\text{sum of TF-IDF of words in the target opinion}}{\text{number of words in the target opinion}}
\]

Before testing “Similarity” between categorized opinions and a target opinion, TF-IDF values of all words in the category are calculated.

The representative ratio of word (TF-IDF) = \( tf \times \log \left( \frac{N}{df} \right) \)

- \( tf \): number of appearance of a word \( w \)
- \( N \): total number of categories
- \( df \): number of categories used word \( w \)

The word-representative ratio table includes the following words:

- Simple
- Difficult
- Quiz
- Packet
- Easy
- Want

**Category:** “Quiz is difficult”.

**Pre-calculation**

**Word-representative ratio table**

<table>
<thead>
<tr>
<th>word</th>
<th>TF-IDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>simple</td>
<td>30</td>
</tr>
<tr>
<td>difficult</td>
<td>29</td>
</tr>
<tr>
<td>quiz</td>
<td>21</td>
</tr>
</tbody>
</table>

**Target opinion:** “Quiz is difficult”.

**Refer**

**Judgments**

\[ RW = \frac{21 + 29}{2} = 25 \]

\[ RW > \text{threshold} = k \times (\text{average RWs of opinions in the category}) \]

The target opinion may be classified into the category.
3.3 New classification method (3/5)

(b) Coincidence of used words of the target opinion and words in all opinions in the category (CW)

\[ CW = \frac{\text{sum of Jaccard coefficient of opinions in a category}}{\text{number of opinions in a category}} \]

**Jaccard coefficient** is a common used coefficient for evaluating similarity between two sentences. Average of **Jaccard coefficients** is calculated.

\[ \text{Jaccard coefficient} = \frac{\text{number of common words}}{\text{number of all words in both opinions}} \]

**Jaccard coefficient** = 1/3 = 0.33

**CW** = \( \frac{(0.50 + 0.67 + 0.25 + 0.33)}{4} = 0.44 \)

The target opinion may be classified into the category.
3.3 New classification method (4/5)

Evaluation (1)

For combination use of the conventional and new methods. The number of re-categorized opinions is large.

Target opinions
(atypical opinions judged by the conventional method from 8 month data)

1202

Typical opinions: 739
Atypical opinions: 463

52 categories,
pre-categorized opinions: 2044

Categorized opinions
(typical opinions)

Classification

355

Typical opinions: 431
(Atypical opinions: 416)

Atypical opinions

The half of the target opinions, which are difficult to classify, can be categorized.

Since another half still remain in the atypical opinions category, more improvement is desired.

K=0.1, J=0.1
3.3 New classification method (5/5)

Evaluation (2)

For single use of the new method. The pre-categorized opinions by hand are few.

pre-categorized opinions: 135

target opinions 2336

Typical opinions: 1931
Atypical opinions: 405

Typical opinions: 16 categories
Atypical opinions: 9 categories

Atypical opinions correctly extracted by the method
Atypical opinions originally identified by the provider

Recall = \frac{\text{Atypical opinions correctly extracted by the method}}{\text{Atypical opinions originally identified by the provider}}

Precision = \frac{\text{Atypical opinions correctly extracted by the method}}{\text{Atypical opinions extracted by the method}}

Performance is worse. Setting is quite easy.
So, we feel the new method can be improved easily.
3.4 User interface

- Typical opinions are shown as bar graph that represents the relationship between the number in each category and the unsubscription date.
- Atypical opinions are shown as card in order to grasp opinion content intuitively.
3.4 Function of the interface (1)

Interface for typical opinions

- The graph represents the relationship between the number in each category or group and the unsubscription date.

**Group:**
The set of categories that assembled by analyst

ex)  
- “Quiz is difficult.”
- “Quiz is interesting.”

→ “quiz”

Analyst can analyze in perspective.

Analysis example) The number of “quiz” opinions are increasing gradually.

→ The service provider should consider strategies about “quiz”.
3.4 Function of the interface(2)

An analyst can set two axes of opinion card area from properties of opinion and subscriber.

Analyst can grasp many correlations between two axes intuitively.

Examples of properties:
- Age
- Job
- Number of game play
- Similarity
4. Analysis examples (1/3)

- Horizontal axis: Singularity
- Vertical axis: Subscription period

Cards with low singularity are few at this rectangular.

It is possible to analyze that users who subscribe long term comparatively offer unexpected opinions.

By focusing on opinions of dissatisfaction in this circle, the service provider considers a strategy in which users who quit early can be made to subscribe long term.
4. Analysis examples (2/3)

- Horizontal axis: Singularity
- Vertical axis: Number of game play

The service provider found “The requirement differs from number of game plays”.

Light users request quick response quiz.

Ordinary users request categorized quiz.

Heavy users request the description of the quiz.
4. Analysis examples (3/3)

- Horizontal axis: Age
- Vertical axis: Unsubscribed date

From the card layout in this rectangular, young users did not answer the open-ended questionnaire.
5. Conclusions

- An improved analysis support system of open-ended questionnaire data has been developed.
  
  The new classification method is effective when the sufficient number of opinions are pre-classified.
  
  However, when the pre-classified opinions are few, the classification performance should be improved.

- Further researches;

  1. Improvement of the proposed classification method.
  
  2. New applications;
     - Problem analysis in software development projects.
       The individual problems occurred in plural projects are extracted from project reports.
     - View analysis of students
       From the students’ reports, the distribution of the interest and claims is visualized.
Thank you for your kind attention.
X.0 I/F developed in China

Problem analysis in offshore software development projects in Jinan, China

Project team (no meaning in order)

Academic hierarchy