DEIM Forum 2019 G6-2

# Query Keyword Extraction from Complaint Data for Collecting Advice

Liang YANG†, Daisuke KITAYAMA††, and Kazutoshi SUMIYA†

† Kwansei Gakuin University 2-1 Gakuen, Sanda-shi, Hyogo, 669-1337 Japan
†† Kogakuin University 1-24-2 Nishi-Shinjuku, Shinjuku-ku, Tokyo, 163-8677 Japan
E-mail: †{dui93794,sumiya}@kwansei.ac.jp, ††kitayama@cc.kogakuin.ac.jp

**Abstract** Nowadays, there are a large number of users who post complaints about a certain service on the Internet. Because users have various values and views, even if they receive the same service, they may complain in different ways. However, it is quite difficult to respond to various user demands for service in real time and there are almost no direct solutions when users feel dissatisfied with a certain service. Therefore, in this paper, we propose a method that generates query keywords according to various user complaints about a certain service for collecting advice. To generate query keywords, we extract the company names, complaint topic words and advice topic words from complaint data. Then suitable web pages containing advice are recommended from the results of the query. This advice could address users' dissatisfaction and respond to their different demands in a comprehensive way.

**Key words** query extraction, complaints, advice.

## 1. Introduction

Nowadays, there are a large number of users who post complaints about a certain service on the Internet. Because users have various values and views, even if they receive the same service, they may complain in different ways. However, it is quite difficult to respond to various user demands for service in real time and there are almost no direct solutions when users feel dissatisfied with a certain service. What's more, it is difficult to determine the expected solutions if the query keywords are not appropriate, especially for certain users who are not used to online search, like the elderly and foreigners. For instance, users might complain about an online retailer's cardboard box using different phrases such as "the size is too big," "low quality," or "over-wrapped." It is difficult to resolve users' complaints in a direct way because items such as a cardboard box are chosen by retailers and users cannot determine their size or quality by themselves. Moreover, the expected solutions would not be found if users search with keywords such as "cardboard box, too big." Hence, we suggest resolving users' complaints in an indirect way by offering them advice like "how to modify a cardboard box" or "how to utilize a cardboard box for storage."

To offer these kind of advice, in this paper, we propose a method that generates query keywords according to various user complaints about a certain service. To generate query keywords, we first analyze complaint data from the Fuman Kaitori Center website and extract company names from the data. Second, from the negative reviews for each company, we extract complaint topic words that represent users' dissatisfactions. Third, we extract advice topic words that can help users to collect advice according to their complaint. Finally, we generate the query based on the company name, complaint topic word, and advice topic word. Moreover, we use an OR-based search to acquire suitable web pages containing advice by using the result of the query. This advice could address users' dissatisfaction and respond to their different demands in a comprehensive way.

The remainder of this paper is structured as follows. Section II presents a brief summary of related work. Section III introduces the dataset we use for research and explains the proposed method. Section IV discusses the experimental results and the evaluation of the proposed method. Finally, Section VI concludes this paper and discusses future work.

## 2. Related Work

### 2.1 FKC dataset

The FKC dataset has been used for several studies in recent years. Mitsuzawa et al. [1] presented the FKC dataset which is from Fuman Kaitori Center (FKC). "Fuman" means dissatisfaction in Japanese. The FKC is a Japanese consumers' negative opinion data collection and analysis service. In our work, we used and analyzed the FKC dataset.

Hasegawa et al. [2] analyzed and visualized the contents of the FKC dataset such as the distribution of users' ages, jobs, and gender. In our work, we determine the aims of the experiment based on their results.

Hayashi et al. [3] proposed a system to recommend appropriate products for users according to their complaints. This

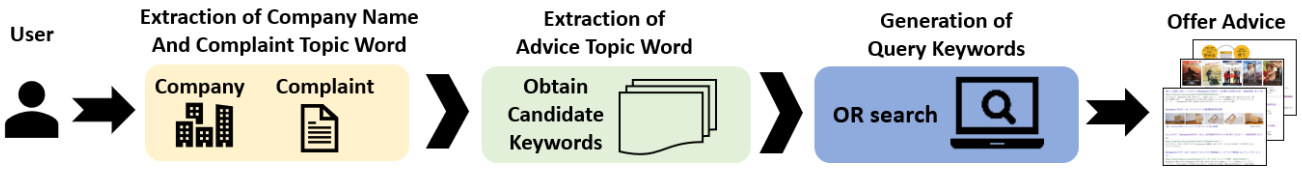


Figure 1 System Flow

recommendation could directly resolve users' dissatisfaction with certain products. In our work, we propose a method to resolve user complaints about services instead of products in an indirect way.

### 2.2 Topic Word Extraction

Sakai et al. [5] proposed a method to extract negative words as the expressions of dissatisfaction from blogs. They extracted nouns, adjectives to make a dissatisfaction expression dictionary. In our work, we only extract nouns because nouns can explain and represent the content of users' complaints.

Hashimoto et al. [6] proposed a method to extract important topics from newspaper and detect social problems based on document clustering. Ustumi et al. [4] proposed a method to extract technological solutions to social problems such as medical issues from the news. They extracted technological solution words by calculating the relevance of problems and technologies. They defined the relevance calculation as problem relevancy and technical relevancy. A higher value of relevancy indicated a higher possibility of being able to extract a technological solution word. In our work, we use this concept and extract the advice topic word by calculating the relevance of the company and complaint topic. However, we hypothesize that a lower relevancy indicates a higher probability that a word is an advice topic word.

Yoshida et al. [7] proposed a method to extract features terms from the customer reviews of e-commerce sites in order to recommend similar items to users. They used polarity analysis to calculate the degree of importance of feature words by counting the number of positive reviews, negative reviews, and positive ratings. In our work, we also use polarity analysis to evaluate advice topic words. However, we weight words according to the result of polarity analysis.

### 2.3 Query Generation

Song et al. [9] and Kajinami et al. [10] proposed a system to generate query keywords that can support a user's search intention. Kakimoto et al. [8] proposed a system to extract query keywords from the closed caption data of TV programs to recommend web pages related to tourism and events based on users' preferences. In our work, we extract query keywords from negative reviews to recommend web pages of advice with the aim of addressing a user's dissatisfaction with a certain service.

## 3. Proposed Method

### 3.1 Overview

In this paper, we propose a method to generate query keywords for users to search for advice according to their various complaints about a certain service. Figure 1 shows the system flow of our proposed method. First, we extract the company name and complaint topic words by calculating the importance of the nouns in the negative reviews. Second, we obtain candidate search keywords of these extracted words. Then, the system extracts the advice topic word by calculating the relevancy of the candidate keywords to the FKC dataset and score them using morphological and polarity analyses. Third, we create the query by combining the company name, complaint topic word, and advice topic word. These query keywords could respond to the needs and search intentions of different users in a comprehensive way. Finally, suitable web pages containing advice that could address a user's complaint are recommended from the results of the query.

### 3.2 Dataset

In this study, we analyze a dataset of complaints from the Fuman Kaitori Center, which is provided by Insight Tech Inc. from the National Institute of Informatics. In this paper, we refer to the Fuman Kaitori Center's dataset as the FKC dataset. The Fuman Kaitori Center is a website on which users can post their complaints about topics such as products, services, education, work, and relationships. Moreover, users get points when they post complaints that they can exchange for coupons for online shopping websites. This dataset contains about 5 million negative reviews that were posted from 18 March 2015 to 12 March 2017 by around 100,000 users. Each negative review contains the information shown in TABLE I. For example, the FKC dataset contains categories like "public transportation," "restaurants," and "industry." Each category contains several subcategories that describe the detailed information of the category. For in-

stance, the category "industry" contains subcategories such as "agriculture," "architecture," and "IT web services." In addition, each subcategory contains several companies that are recorded in the dataset. Moreover, because we focus on user service complaints, we only analyze the data fields for "company" and "text" in the method proposed in this paper.

Table 1 Details of the FKC dataset

| Data item | Content |
|---|---|
| post_id | complaint ID |
| user_id | Fuman Kaitori Center ID |
| category | complaint category |
| subcategory | detailed complaint category |
| company | company name |
| product | product name |
| text | negative review |

### 3.3 Extraction of Company Names and Complaint Topic Words

Our proposed method extracts company names from FKC dataset directly from the company field of each record. Next, we extract the complaint topic word by analyzing negative reviews. In this paper, we only use the negative reviews that are labeled with the company name. To extract complaint topic words, we first extract all companies' negative reviews for one subcategory. Then, we use morphological analysis to extract all nouns from the negative reviews. Next, we calculate the importance of each noun using the following equation.

$$\frac{tf}{|A|} \times \frac{tf}{\sum_{d \in D} tf_d} \tag{1}$$

Here, $tf$ is defined as the number of occurrences of a particular noun in the complaints for a certain company, $|A|$ is defined as the number of all nouns in the complaints for a certain company, and $\sum_{d \in D} tf_d$ is defined as the number of occurrences of certain noun for the complaints for all companies. After calculating the importance of each noun, we determine the threshold value for that company. Finally, we extract all nouns whose importance values are above the threshold and define them as that company's complaint topic words.

### 3.4 Extraction of Advice Topic Words

To extract the advice topic word, we first obtain candidate search keywords from a Google search of the company name and complaint topic word. Because the FKC dataset is full of negative reviews, we hypothesize that candidate keyword that are less relevant to the complaint data will make better advice topic words. To verify this hypothesis, we calculated the relevance of these candidate keywords for each company and each complaint topic word. In this study, we define the "company relevancy" as the relevance between a candidate keyword and a particular company. It is calculated using the following equation.

$$company\,relevancy = \frac{R_{cd}}{R_c} \tag{2}$$

Here, $R_{cd}$ is defined as the number of occurrences of certain candidate keyword in complaints for the company in the FKC dataset and $R_c$ is defined as the number of negative reviews of that company.

Next, we define the relevance between a candidate keyword and a complaint topic word as the "complaint topic relevancy" using the following equation.

$$complaint\,topic\,relevancy = \frac{R_{td}}{R_t} \tag{3}$$

Here, $R_{td}$ is defined as the number of occurrences of the candidate keyword with the complaint topic word in the negative reviews of the FKC dataset and $R_t$ is defined as the number of negative reviews with that complaint topic word.

However, not all candidate keywords with low relevance in the FKC dataset can be used as advice topic words. For example, some negative words as well as verbs and adjectives do not help users acquire advice. Hence, it is necessary to exclude these kinds of candidate keywords. To do this, we weight candidate keywords using morphological and polarity analyses, as shown in TABLE II.

Table 2 Weights for Candidate Keywords

| Result of Analysis | Weight |
|---|---|
| negative | 0.8 |
| verb | 0.7 |
| adjective | 0.7 |
| proper noun (place name) | 0.7 |
| proper noun (organization name) | 0.3 |
| common noun | 0.3 |
| verbal noun | 0.1 |

Finally, we calculate the final score of the candidate keywords by combining the arithmetic mean of the company and complaint topic relevancies with the weight as the following equation. Here, relevancies are calculated by adding the results of equation 2 and 3.

$$Score = \frac{relevancies}{2} \times Weight \tag{4}$$

After calculating the final score of each candidate keyword, we determine the threshold value for each company. Candidate keywords those scores are under the threshold value become the advice topic words.

### 3.5 Generation of web search queries

To respond to the different demands of users, each company name and complaint topic word are matched with several advice topic words. To search for suitable websites, We use an OR-based search method to acquire advice websites. Our proposed system generates the query based on one company name, one complaint topic word, and one advice topic word.

### 3.6 Recommendation of Advice

Our proposed system recommends suitable web pages containing advice from the results of the query which is based on users' complaints. Figure 2 shows the user interface of our proposed system. First, the system generates several queries by analyzing user's negative review. Next, When user clicks on one of the offered query, the query will turn into the web page of the results of search. Finally, user can choose and browse the web page based on their needs. The system shows the advice that could address user's dissatisfaction expressed in the negative review.


Figure 2 User Interface

## 4. Experiment and Evaluation

### 4.1 Dataset

In this study, we conducted an experiment to extract the complaint and advice topic words in order to verify the feasibility of proposed method. For this experiment, we analyzed the subcategory of "IT web services" of the FKC dataset, which is under the category "industry." We extracted 1,000 negative reviews for each of the six companies. The final dataset consisted of 6,000 negative reviews in total. And we analyzed three of the companies for experiment.

### 4.2 Experiment

First, we extracted the complaint topic words and determined different threshold values for each of the three companies. For company A, we extracted 186 complaint topic words above the threshold value of 0.00080. For company B, we extracted 144 complaint topic words above the threshold value of 0.00076. For company C, we extracted 86 complaint topic words above the threshold value 0.00080. Examples of the complaint topic words for each company are shown in TABLE III. These examples show that each complaint topic word implies the object of different users' dissatisfactions.

Table 3 Example Complaint Topic Words

| Company | Complaint topic word |
|---|---|
| A | prime, delivery, movie, cardboard box, delivery fee, order, reservation, book, print, return |
| B | stamp, code, block, coin backup, telephone call, camera, setting, post, commercial |
| C | news, question, premium, answer, auction, article, title, ID, weather forecast, transaction |

Next,we extracted advice topic words from the candidate keywords that had a score less than 0.0043, 0.0020, and 0.0033 for companies A, B, and C, respectively. We extracted advice topic words that match with 20, 10, and 10 complaint topic words for companies A, B, and C, respectively. Some examples of these words are shown in TABLE IV. As TABLE IV shows, even if the candidate keywords are under the threshold value, not all of them can be used as advice topic words. However, the proposed method is sufficient for ranking candidate keywords. Most of the extracted results can be used as advice topic words. In the future, we plan to develop a method to determine a more accurate threshold value to better exclude noise in the results.

Table 4 Example Advice Topic Words

| Company name | Complaint topic words | Advice topic words |
|---|---|---|
| A | point | charge, present, how to save up, how to use, credit card |
| B | setting | security, group friend, privacy, initialization, recommendation |
| C | premium | privilege, merit cancellation of agreement, magazine |

### 4.3 Evaluation

We also conducted a questionnaire-based survey to evaluate the usability and effectiveness of the proposed method.The questionnaire-based survey contained following 2 questions. For Q1, we extracted all nouns from the negative review for respondents to choose from. For Q2, we provided 10 candidate queries for each negative review to choose from.
Q1:Please choose one word which you think could represent the dissatisfaction of the following reviews.
Q2:please choose the query that you think the contents returned by a search using this query keyword could address

the complaints found in the negative review.(multiple choices are allowed)

### 4. 4 Result and Discussion

We collected the answers of 10 respondents, and the results are shown in Table 5 and Table 6. We defined those nouns and queries were chosen by over 5 answers as true positive.

Table 5 Result of Q1

| | p@1 |
|---|---|
| Average of p@k | 0.60 |

The result of Q1 shows that if we search by using those nouns are with the highest value of importance for one time only, 60% of appropriate complaint topic word can be extracted from the negative review. It not only shows that the proposed method is effective, but also explains the method to rank nouns by calculating the importance is correct.

Table 6 Result of Q2

| | p@1 | p@3 |
|---|---|---|
| Average of p@k | 0.70 | 0.50 |
| Average of r@k | 0.32 | 0.68 |
| F-measure | 0.44 | 0.58 |

The result of Q2 shows that if we search by using the query which is with the lowest score for one time only, 70% of appropriate queries can be offered to address the complaints expresses in the negative reviews. What's more, it shows that users can acquire nearly 70% of true positive by using query keywords those are with the lowest 3 scores instead of 10 queries from search engine. This result does not only demonstrate that the query keywords generated by the proposed method by scoring candidate keyword is effective, but also implies by using proposed method can help users to release their burden when searching for advice . Moreover, it is believed that the proposed method can respond to the demands of different users.

## 5. Conclusion

In this paper, we proposed a method to generate query keywords that will enable users to collect advice about a certain service. First, we analyzed the FKC dataset and extracted particular company names from the data. Second, we extracted complaint topic word that can express the dissatisfaction of users from the negative reviews for each company. Third, we extracted advice topic word that can help users to collect advice to address their complaints. Finally, we generated a query based on the company name, complaint topic words, and advice topic words. The advice on web pages returned by this query could address users' dissatisfaction with a service and respond to different user demands in a comprehensive way. In addition, we evaluated the effectiveness and usability of the proposed method through a questionnaire survey, and the results shows thought the generated query keywords would be useful for collecting advice.

In the future, we plan to analyze other categories of the FKC dataset. Moreover, we plan to create a website that can automatically recommend advice web pages when users post their complaints on the website.

## Acknowledgment

In this paper, we used FKC Data Set provided for research purposes by National Institute of Informatics in cooperation with Insight Tech Inc.

### References

[1] K. Mitsuzawa, M. Tauchi, M. Domoulin, M, Nakashima and T. Mizumoto. " FKC Corpus: a Japanese Corpus from New Opinion Survey Service," In proceedings of the Novel Incentives for Collecting Data and Annotation from People: types, implementation, tasking requirements, workflow and results, Portoro, Slovenia pp.11-18, May. 2016

[2] T. Hasegawa and D. Kitayama. "The Visualization of Dissatisfaction Groups using Dissatisfaction Dataset," *DEIM Forum 2017*, P7-1.

[3] T. Hayashi, Y. Wang, Y. Kawai, and K. Sumiya. "An E-Commerce Recommender System using Complaint Data and Review Data," *Proc. of ACM IUI2018 Workshop on Web Intelligence and Interaction (WII 2018).*

[4] K. Utsumi, K. Inui, T. Hashimoto, H. Kimura, and M. Ishikawa. "Extraction of Critical Knowledge concerning Social Problems and their Technological Solutions," *Socio Technology Research Journal*, vol. 6, pp. 187–198, Mar. 2009.

[5] T. Sakai and K. Fujimura. "Discovering Latent Solutions from Expressions of Dissatisfaction in Blogs," *Information Processing Society of Japan*, vol. 52, no. 12, pp. 3806–3816, Dec. 2011.

[6] T. Hashimoto and H. Kimura and K. Inui and K. Utsumi and M. Ishikawa. "Topic Extraction and Social Problem Detection based on Document Clustering," *Socio Technology Research Journal*, vol. 5, pp. 216-226, Mar. 2008.

[7] T. Yoshida and D. Kitayama. "An Evaluation of Feature Term Extraction Method based on Polarity Analysis from Customer Reviews," IEICE-DE2016-5 vol. 116, no. 105, pp. 19–24, Jun. 2016.

[8] H. Kakimoto, T. Hayashi, Y. Wang, Y. Kawai, and K. Sumiya. "Query Keyword Extraction from Video Caption Data based on Spatio-Temporal Features ," *Lecture Notes in Engineering and Computer Science: Proceedings of the International MultiConference of Engineers and Computer Scientists 2018*, pp. 405–408.

[9] X. Song and M. Takaku. "Study on Navigation support System based on User ' s Search Intents," *ARG WI2*, no. 9, 2016.