

Method for Atypical Opinion Extraction from Answers in Open-ended Questions

Ayako Hiramatsu

Department of Information Systems Engineering
Osaka Sangyo University
3-1-1, Nakagaito, Daitou, Osaka 574-8530
Japan
ayako@ise.osaka-sandai.ac.jp

Hiroaki Oiso

Codetoys K. K.
2-6-8 Nishitenma, Kita-ku,
Osaka 530-0047
Japan
oiso@codetoys.com

Shingo Tamura

Graduate School of Information
Science and Technology, Osaka University
2-1, Yamada-oka, Suita 565-0871
Japan
tamura.shingo@ist.osaka-u.ac.jp

Norihisa Komoda

Graduate School of Information
Science and Technology, Osaka University
2-1, Yamada-oka, Suita 565-0871
Japan
komoda@ist.osaka-u.ac.jp

Abstract – This paper presents a method for atypical opinion extraction from answers to open-ended questions supplied by users of a mobile game content when they unsubscribe from the service. The proposed system excludes typical opinions and extracts only atypical opinions. To cope with incomplete syntax of texts due to the input by mobile phone, the system treats the opinions as the sets of keywords. The combination of words are established beforehand in a typical word database. Based on the ratio of typical word combinations in sentences of an opinion, the system classifies the opinion typical or atypical. When typical word combinations are sought in an opinion, the system considers the word order and the distance of difference between the positions of words to exclude unnecessary combinations. Furthermore, when an opinion includes meanings the system divides the opinion into phrases at each typical word combination.

With an experiment, the extraction accuracy of the proposed system was confirmed.

I. INTRODUCTION

The mobile game market has been expanding rapidly since the late 1990s. The mobile game market has flourished in Japan for two significant reasons. First, the combination of fixed monthly fees and variable packet charges is reasonable for consumers. Second, the commission charged by carriers for collecting the content fee is a low 10%, which encourages the participation of providers. However, the users actually have to pay a large sum as a packet charge and their share of monthly fee of total payments is extremely low. Besides, carriers only collect monthly fees as an allocation for the game provider. Game providers must, therefore, attract more users and prolong the subscription period per user with effective marketing.

Our target mobile game is a quiz game (from now referred to as the game) on the official menus of three domestic carriers since 2002. The purpose of this paper is the extraction of effective opinions that include ideas for extending subscription periods from questionnaires that subscribers answer when cancelling their account. The questionnaire consists of closed-ended and open-ended questions answered by cellular phones.

Using closed-ended questions, the questionnaire asked users to choose from a limited number of pre-selected answers and responses that will be analyzed statistically and rapidly [1]. The language of the opinions from closed-ended

questions, however, is completely provided by the consumer. To acquire unexpected original ideas, answers to open-ended questions must be analyzed.

Since open-ended questions, however, place no restrictions on descriptions, the answers include an enormous amount of text data for the game provider. It is time-consuming to read all of the texts one by one. Since a large number of answers are identical to choices included in the closed-ended questions or unnecessary opinions unconcerned with the game, there are few useful answers. Additionally, since answers are input through cellular phones, they often include many symbols dependent on various kinds of terminals and grammatical mistakes, making them hard to understand. Our research, therefore, aims to create a system that efficiently extracts unexpectedly unique ideas by culling useless opinions from the data of open-ended questions.

Recently, many text mining studies aim to identify efficient patterns from enormous text data. In text mining research for questionnaire analysis, the main keywords are extracted with the word concurrence graphs; search words are recognized as important by conditional probabilities that show relationships between words or main keywords, expressed by many users, they are extracted [2]. Thus the process of many text mining research methods includes summarization, extraction of themes, and categorization [3][4][5][6][7]. However, little text mining research has sought to extract unique data.

To extract unexpected opinions from open-ended questionnaire data, our research aims to comprehensively exclude common opinions by the following method. Open-ended questions place no restrictions on descriptions. Consumers supply answers by cellular phone, producing answers that are often ungrammatical because of troublesome input tools. Our approach grasps the meaning of opinions with word units. First, the answer data from open-ended questions are decomposed into word lists by morphological analysis. Combinations of words are set up in a typical word database beforehand. Based on the ratio of typical word combinations in the word list of an opinion, the system classifies the opinion typical or atypical. When typical word combinations are searched for in an opinion, the system excludes unnecessary combinations by considering the word order and the distance of difference between the positions of words. Furthermore, when an

opinion includes meanings the system divides the opinion into phrases at each typical word combination.

II. SUPPORT SYSTEM FOR ANALYSIS OF UNSUBSCRIBING QUESTIONNAIRE DATA

A. Target mobile games

A target game is provided in the form of a quiz game provided by three mobile carriers. There are no differences in game rules, although there are minor differences (e.g.: display layout). A game consists of a maximum of 15 questions. All questions are answered by choosing a correct answer from 4 choices. If the first question is answered correctly, the player receives 10,000 points, which doubles every time the player answers correctly. Apparently players can amass a lot of score full points because it is easy to cheat through a mobile phone. Not only points are counted but also the time required answering questions, which is part of the final score calculated at the end of each month. The best score in a game each month is the consumer's score and is used in the ranking. In this game, four levels - 1st stage (lowest stage) through 4th stage (highest stage) - are prepared. The ranking is announced at the end of each month and only the top 25% players of each stage can move to the next stage. If consumers unsubscribe, all information is lost. So if he resubscribes, he starts from the first stage.

The fee is 180 Japanese yen (about \$ 1.6) per month. The packet charge is about 4 yen per quiz (= 1 page). Therefore, the packet charge amounts to about 40 yen. If a player plays 150 games a month, it will cost 6,000 yen per month. In application versions, the packet charge is 20 yen per 90 questions-quiz pack that is downloaded and 200 to 400 to download the application. To prolong subscriptions, prizes are provided to the leading scorers. In this game, only the top scorers at the highest stage receive prizes.

B. Unsubscribing questionnaire

The questionnaire given to consumers who unsubscribe from the game consists of closed-ended and open-ended questions. In the closed-ended questions, the questionnaire asks consumers to choose from a limited number of prelisted answers. In the open-ended questions, consumers freely write opinions.

1. Closed-ended questions

In questions about reasons for unsubscribing, the following 11 items are available from which users can choose any number.

(a) I stopped playing the game. (b) There are not enough incentives. (c) I could not win a prize. (d) Quizzes are not interesting. (e) I lost interest in the quizzes. (f) There is little number of "fastest finger first" quiz. (g) Communication conditions are bad. (h) The 180 yen fee is too high. (i) The packet charge is too high. (j) Support is bad. (k) Other.

2. Open-ended questions

The open-ended questions are the followings: (a) Improvement demands and (b) Other opinions. Consumers can freely write opinions for each question. However, because of an ambiguity in Japanese expressions, some consumers have misunderstood

question (a) as a question about asking if they were happy with the games. The writing rate of opinions for each question is around 10%.

C. Problems with open-ended questionnaire data

The answers from the consumers include such important remarks as dissatisfaction that cannot be captured in the closed-ended questions. The answers also include impressions or demands of which the provider is unaware. Unfortunately, useful opinions comprise only about 5% of all the opinions. Most answers reflect opinions already known by the provider or duplicate the meaning as of other answers.

To grasp unexpected opinions, our system tries to classify the opinions of open-ended questionnaire data into typical and atypical opinions. The definition of typical and atypical opinions follows:

1. Typical opinions:

- (a) Opinions having the same meaning as items of the closed-ended questions. (e.g.: The packet charge is too expensive.)
- (b) Frequent opinions that the provider has already heard. (e.g.: My knowledge increased.)
- (c) Irrelevant opinions. (e.g.: I had a baby!)

2. Atypical opinions: Any opinions not typical. (e.g.: Quizzes are for kids)

Although opinions are typical, because of the variety of expression, it is difficult to judge whether opinions offer the same meanings or new and useful information. Opinions input on cellular phones are also often not punctuated, ungrammatical, and abbreviated. For incomplete syntax of the texts due to the input by mobile phone, an approach based on natural language processing takes little effect.

D. Outline of the support system

Since a large number of answers are the same as choices included in the closed-ended questions or unnecessary opinions unconcerned with the game, there are few useful answers. Neither can providers anticipate atypical answers and prepare such concepts. Our research, therefore, aims at a support system that efficiently extracts unexpected, unique ideas or atypical opinions by cutting typical opinions from the answer data of open-ended questions. Fig. 1 shows the outline of our support system.

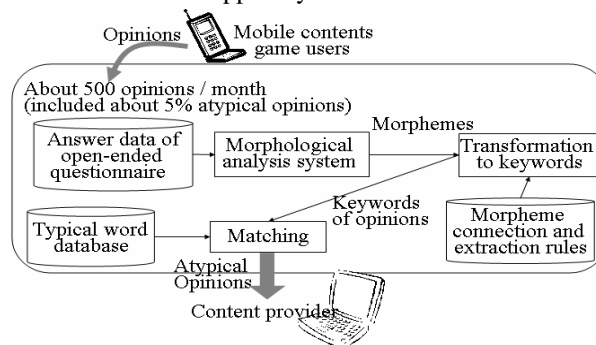


Fig.1. Outline of the support system

Because of the problems mentioned above, our approach grasps the meaning of opinions with word units. The answer data of open-ended questions is decomposed into word lists by morphological analysis with "ChaSen" [8], Japanese software.

From the word lists, we extract nouns, independent adjectives, and independent verbs as the minimum words for understanding a sentence. In Japanese, nouns are roughly divided into 14 kinds. There are also words that join keywords in context. Based on morpheme connection and extraction rules, morphemes are transformed to keywords. For example, "packet" and "fee" are transformed to keyword "packet fee" by the rules.

In the typical word database, the combinations of words are set up beforehand. Fig. 2 shows an example of an opinion; "the packet fee is high." Only "high" is not synonymous with "expensive." However, when "high" and "expensive" are with "packet fee," they have the same meaning.

A matching engine to extract atypical opinions compares keywords of opinions with the typical word database. For the extraction, we prepared three comparing methods explained in the following section.

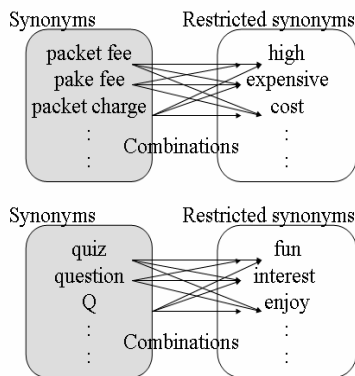


Fig.2. Combinations in typical word database

III. METHODS TO EXTRACT ATYPICAL OPINIONS

To extract atypical opinions, the proposed system needs to compare the keywords of each opinion with the typical word database. This section explains three methods. The first method is based on the ratio of typical word combinations in the sentences of an opinion, and classifies the opinion typical or atypical. In the second method, when typical word combinations are located in an opinion that excludes unnecessary combinations, the system considers the word order and the distance of difference between the positions of words. In the third method, when an opinion includes some ideas the system divides the opinion into phrases at each typical word combination.

A. The basic comparing method

At first, opinions that have neither keyword nor a noun keyword are excluded because these opinions make no sense; there is no atypical opinion without noun keywords. Other opinions are judged by comparing their keywords

with the combinations in the typical word database.

When an opinion includes a combination of words in the typical word database, the combination of words is called "typical element." If the ratio of typical elements to all keywords in an opinion is over a certain ratio, the opinion is classified typical. If an opinion satisfies the following formula (1), it is classified as typical.

$$\begin{aligned} & (\text{The number of keywords in typical elements}) \\ & + \alpha \times (\text{The number of typical elements}) \\ & \geq (\text{The number of all keywords}) \quad (1) \end{aligned}$$

Here, α shows how many other keywords are permitted to include in a typical opinion for each typical element.

Fig. 3 shows an example of classification with $\alpha = 2$. In Fig. 3, two word lists are enclosed by rectangles and typical elements are enclosed by ellipses. The order of keywords in these word list examples are based on the original Japanese texts. The upper word list includes one typical element and two other keywords, and this opinion with the upper word list is classified typical. On the other hand, the lower word list includes one typical element and 4 other keywords, and it is classified atypical.

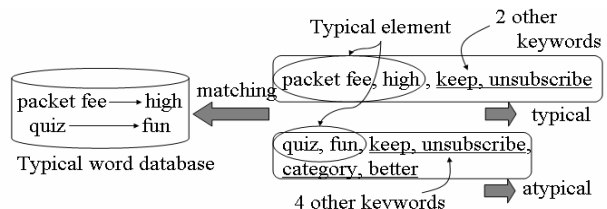


Fig.3. Classification based on first method

B. Problems with basic comparing method

The above mentioned basic comparing method counts combinations of words that appeared in any positions of word lists as typical elements. Fig. 4 shows an example of this problem. The combination of "quiz" and "fun" is extracted as a typical element. However, the opinion is not "quizzes are fun", but rather "history quizzes would be more fun". Misrecognizing creates unnecessary combinations of typical elements, and atypical opinions are treated as typical opinions.

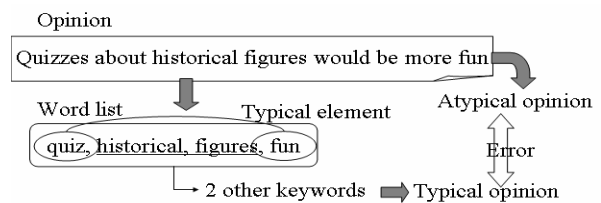


Fig.4. Classification problem

Another classification problem occurs in long sentences that include more than one meaning in an opinion. If the ratio of keywords about atypical elements is small compared to all the keywords of the opinion, it is classified typical, even though it includes atypical meanings. As a long sentence example, there is an opinion whose word list

has 15 keywords. When parameter α is 2 and the word list includes 5 typical elements, including 10 other keywords defines the opinion as typical. However, with less than 10 keywords, an atypical opinion can be expressed.

C. Method based on the keyword distance (The second method)

In the basic method, typical elements are extracted with all matched keywords in any word order, and unnecessary combinations are also extracted as typical elements. Especially when an opinion includes such frequently used keywords as "fun" and "happy," misrecognized combinations in the sentences often appear. Normally, if sentences are grammatical, relations between keywords can be found by syntax analysis to identify "what is fun" or "what is happy." In our open-ended questionnaire data, however, it is difficult to identify relations between keywords in sentences by syntax analysis. Therefore, we considered the order of keywords in an opinion, and the position differences of keywords are treated as keyword distance. If the keyword distance is short (that is, 2 keywords appearing near), the keywords are regarded as one combination.

In this method, if keyword combinations in the typical word database appear in the word lists of opinions within a certain keyword distance, we regard them as typical elements. Based on this typical element recognition, formula (1) is applied to classify the opinions. Fig. 5 shows a classification example where the keyword distance is within 2 and the parameter is 2.

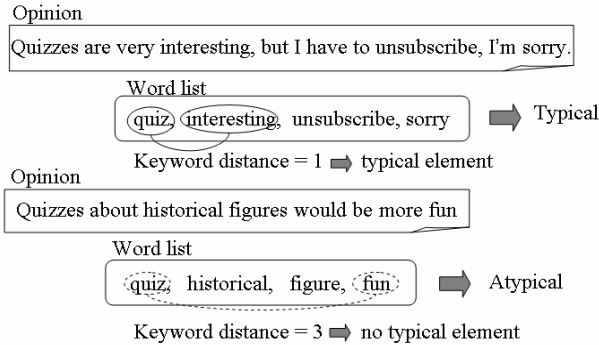


Fig.5. Method based on keyword distance

Originally, the upper opinion in Fig. 5 was typical and the lower opinion was atypical. The word lists of extracted keywords are shown under each opinion. In the upper opinion, the keyword distance between "quiz" and "interesting" is 1, and the combination of these two keywords is regarded as one typical element. By formula (1), since 2 other keywords are permitted to include in each typical element, the upper opinion is classified typical. On the other hand, in the lower opinion, although "quiz" and "fun" are matched with the combination in the typical word database, the keyword distance is 3. In this case, the opinion is not considered to be "quizzes are fun" because the combination of "quiz" and "fun" is not regarded as one typical element. The lower opinion, therefore, has no typical element and is classified as an atypical opinion.

D. Introducing the delimiter (The third method)

When consisting of sentences with many keywords, the opinion includes many typical elements and few atypical elements. Since the provider needs atypical ideas, even if they are only a very small part of an opinion, the few atypical elements in an opinion must not be overlooked. To classify opinions without missing atypical elements, sentences in opinions should be divided at gaps of meanings. However, since target text data input through cellular phone often contains peculiar pictographs instead of punctuation marks or sentences are not delimited beforehand, it is difficult to delimit sentences to meaningful phrases.

In this method, the points where typical elements appear are regarded as the delimiters of the break points of meanings. Delimiters divide sentences in opinions into phrases. Formula (1) is applied to each phrase, which it judges as typical or atypical. The number of typical elements in a phrase, then, is always 1. If an opinion includes at least one atypical phrase, the opinion is classified atypical.

The following procedure inserts the delimiters.

1. The target sentence S consists of n keywords (w).

$$S = [w_1, w_2, \dots, w_n]$$
2. The set of word combinations in the typical word database is R , and the word combinations are

$$R = [(w_{a1}, w_{b1}), (w_{a2}, w_{b2}), \dots, (w_{am}, w_{bm})]$$
3. When a word combination (w_{ax}, w_{bx}) in R appears at $(w_i, w_j) (j < i + d) (d$: the keywords distance parameter) in S , the combination (w_i, w_j) is regarded as one typical element.
4. The number of appearances of typical elements is N . If $N \leq 2$, go to Step5. If $N > 2$, execute the following. If (w_i, w_j) is the $N - 1$ th typical element in S , and (w_k, w_l) is the N th typical element in S , then the delimiter is inserted between w_{k-1} and w_k .

$$s_{N-1} = [w_i, w_{i+1}, \dots, w_{k-1}]$$
 and

$$s_N = [w_k, w_{k+1}, \dots, w_n]$$
 are the subsets of S and are regarded as phrases.
5. When $N = 1$, $s_1 = S$. When $N = 2$, because of the appearance of the second typical element (w_i, w_j) , $s_1 = [w_0, w_1, \dots, w_{i-1}]$ and $s_2 = [w_i, w_{i+1}, \dots, w_n]$.

Fig. 6 shows an example of classification with this method. This example includes 4 typical elements and is divided into 4 phrases. Since the 4th phrase does not satisfy formula (1), the opinion is classified atypical.

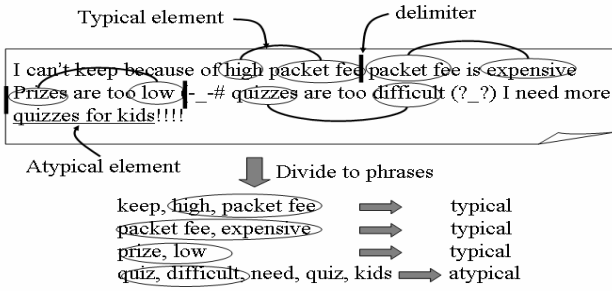


Fig.6. Method introducing the delimiters

IV. EXPERIMENTS

A. Application experiment

To compare the three proposed methods, we collected and used the questionnaire data of users who unsubscribed from a certain carrier for 7 months. The questionnaire data included 3263 opinions. The content provider spent more than 3 hours classifying this data into typical or atypical opinions. The results were 270 atypical opinions. Based on this data, we registered about 8000 kinds of word combinations to the typical word database.

Based on each proposed method, we classified the questionnaire data and calculated the recall and the precision ratios, defined as follows: Recall is the ratio of atypical opinions correctly extracted by the proposed method against the atypical opinions originally identified by the provider. Precision is the ratio of atypical opinions correctly extracted by the proposed method against the atypical opinions extracted by the proposed method. In other words, Recall shows the degree of the extraction leakage, and Precision shows the degree of misextraction.

TABLE I shows the results, when the allowing-other keywords parameter α is 2 and the keyword distance parameter d is 2.

TABLE I.
Results of application experiment

Method	Class		Recall (%)	Precision (%)
	Typical	Atypical		
1	3046	217	65.6	81.2
2	2957	306	94.1	83.0
3	2947	316	100	85.4

As shown by the results, the third method is best in both Recall and Precision. Many opinions in the questionnaire data are short sentences. Since the second method considers the distance of keywords and the leakages of extraction are reduced, Recall of the second method is better than Recall of the first method. Furthermore, the extraction leakages in the second method are caused by too many long sentences.

The third method introduces delimiters to cope with these long sentences.

B. Evaluation experiment

To evaluate the proposed method, we examined the third method because it produced best results in the above experiment. For this experiment, we used the questionnaire data of consumers who unsubscribed from other carriers. This data contained 1764 opinions. The provider found 175 atypical opinions from 1764 opinions. Here, the typical word database is the same as in the application experiment. Namely, the typical word database in this experiment is not registered in accordance with the target experiment data. The extraction results are shown in TABLE II.

TABLE II.
Result of evaluation experiment (The third method)

Data	Class		Recall (%)	Precision (%)
	Typical	Atypical		
1764	1480	284	87.4	53.9

The results are less satisfactory than the application experiment. The opinions with short sentences having 3 or 4 keywords cause the low Recall scores. Some atypical opinions can be expressed in only one keyword. In such cases, the proposed method drops atypical opinions. If the allowing-other keywords parameter α is 1, this problem may be improved. However, stricter judgment of typical elements reduces the number of typical elements, and then the method extracts a huge number of atypical opinions; a difficult tradeoff problem.

V. CONCLUSIONS

This paper described a support system for atypical opinion extraction from answers in open-ended questions collected from consumers of mobile games when they unsubscribe. In this support system, we proposed three methods of extraction of atypical opinions: a basic method based on the ratio of typical word combinations in sentences of an opinion that classifies the opinion typical or atypical; a second method that considers word order and the distance of difference between the positions of words to exclude unnecessary combinations when typical word combinations are searched in an opinion; a third method that divides the opinion into phrases at each typical word combination. The experiments proved that the third method is the best. However, differences of carriers also affect the accuracy of extraction.

VI. REFERENCES

- [1] H. Oiso and N. Komoda: "Access Analysis of Monthly-charged Mobile Content Provision," in Proc. of Future Business Technology Conf. (FUBUTEC'2004), pp.76-80 (2004).
- [2] Kenji Yamanishi and Hang Li: "Mining Open Answers

- in Questionnaire Data,” *IEEE Intelligent Systems*, vol. 17, no. 5, pp. 58-63(2002).
- [3] N. C. Jr. Romano, C. Bauer, H. Chen, and J. F. Jr. Nunamaker: “The mindmine comment analysis tool for collaborative attitude solicitation, analysis, sense-making and visualization,” in *Proc. of the 33rd Hawaii Int. Conf. on System Sciences* (in CD-ROM)(2000).
- [4] H. Chen, O. Titkova, R. Orwig, and J. F. Jr. Nunamaker: “Information visualization for collaborative computing,” *IEEE Computer*, vol. 31, no. 8, pp. 75-82(1998).
- [5] P. Frasconi, M. Gori, and G. Soda: “Data categorization using decision trellises,” *IEEE Tran. on Knowledge and Data Engineering*, vol. 11, no. 5, pp. 697-712(1999).
- [6] Hiroshi Shibata, Takafumi Nozaki, Ayako Hiramatsu, and Norihisa Komoda; “A Support System for Requirement Extraction from BBS using Hierarchical Fish-Eye User Interface,” *Proc. of 2001 IEEE International Conference on Systems, Man and Cybernetics* (in CD-ROM)(2001).
- [7] Ayako Hiramatsu, Tatsuya Nakae, Hiroshi Shibata, and Norihisa Komoda: “An Information Extraction Support System from BBS Using Topic Mapping Based on Structure of Conversation with a 3-dimensional Space,” *TOWARDS THE KNOWLEDGE SOCIETY* (Proc. of The Second IFIP Conference on I3E 2002), pp.273-287 (2002).
- [8] Yuji Matsumoto, Akira Kitauchi, Tatsuo Yamashita, Yoshitaka Hirano, Hiroshi Matsuda, Kazuma Takaoka, and Masayuki Asahara; “Japanese Morphological Analysis System ChaSen version 2.2.1,” <http://chasen.aist-nara.ac.jp/chasen/doc/chasen-2.2.1.pdf> (2000).