Diverse Opinions Extraction from Web Reviews Based on Word Frequency Distribution for Each Evaluation Value

Ayaka Kodaira, Ryosuke Yamanishi, *Member, IAENG,* Yoko Nishihara, *Member, IAENG*

Abstract—This paper proposes a method for extracting various opinions about evaluation points from web reviews. In the proposed method, the evaluation points which frequency is different for each evaluation value such as stars. Opinions in reviews with extremely high or low evaluation values are effective to evaluate the review targets in most cases. However, the reviewers themselves are diverse. The evaluation points that many reviewers highly evaluated sometimes can be the cause of the low evaluation in a specific context for the same target. It is thus reasonable to say that comparing those conflicted opinions should help us to understand the features of the services and products. To provide such findings to the readers, focusing on not only the evaluation points which has high frequency but also various opinions in different context should be required. In this paper, we conducted the extraction experiment and had the discussion about the extracted opinions.

Index Terms—review analysis, diverse opinions, web intelligence

I. INTRODUCTION

REVIEW is a good resource to indirectly know the contents, i.e., the review target. As reading reviews, the user can know the opinions of those who used the products/facilities. Reading reviews can be more useful especially in case that the reader can not experience before, e.g., planning to travel.

Most of the web reviews prepare the evaluation value system such as five-stars with free descriptions for the evaluation of the review target. Although the five-stars evaluation system is easy to understand with the numbers, opinions with extremely high or low evaluation have much more attention than expected; the users might focus on only the frequent evaluation points. However, for some specific context, such overlooked evaluation points can be more important.

By paying attention to opinions that are contradictory evaluations and latent in many similar opinions, it can be possible to extract specific contexts and evaluation points which are hard to find among varied and many opinions. Diverse opinions can be extracted by extracting such opinions including those evaluation points. The extracted opinions should not depends on the polarity of words. Moreover, since the users can obtain evaluations that each reviewer felt important, it is possible to read the review while comparing the evaluation points with the users' feelings.

The final purpose of this research is to extract information not only the evaluation values but also the information necessary for the users to decide an item matching each needs. In this paper, we focus on the frequency of words for each evaluation value and the proposed method extracts various opinions. We have discussions about the opinions extracted by the proposed method.

II. RELATED WORK

As a study related to the analysis of the review, Albornoz *et al.* classified the opinions as positive, negative, and intermediate [1]. In their method, the strength of those polarity is also indicated . Their method extracts different evaluations depending on the field of the targeted document. However, it is difficult to deal with some sentences when a certain word is used in the meaning of popular and unpopular bipolar and emotional meaning is not included. In the study of polarity analysis, Turney's method evaluates sentences focusing on co-occurrence of words expressing positive/negative and words of adverb/adjective [2]. Wilson *et al.* propose the method to analyze the polarity of multiple words rather than one word [3]. Dave *et al.* showed the polarity evaluation with word correlation using N-gram [4].

Also, in the analysis of the review, Popescu *et al.* propose a method to evaluate the features of the product and its opinions and polarity. Their method excels in accurately extracting the object/attribute though, the target sentences are limited to those which are composed of the fixed some elements. [5]. Kim *et al.* propose a method to identify opinions including the reasons [6]. Their aim is to identify the reasons to justify the reputation, and it is different from the comparison of diverse opinions, which is the purpose of our research. As a research focusing on words, Riloff *et al.* extracted the nouns that have a strong influence on sentence subjectivity. [7]. Narayanan *et al.* focused on conditions [8]. They propose a method to judge whether the polarity of the condition sentence is affirmative, negative or neutral. The method can judge the sentences with the conditions including the user's situation, however, the input data is limited to the conditional sentences.

As a comparison of the reviews under review, Hoque *et al.* proposed a mechanism to facilitate comparison of opinions by visualizing the polarity in the community conversation on the Web [9]. Kaschesky *et al.* automatically detect the topics and opinions in social media and visualize the communication network for identifying key arguments and positions [10]. In our proposed method, rather than presenting the polarity for each opinion, discovery of "the evaluation points in which various opinions exist" is set as the final goal.
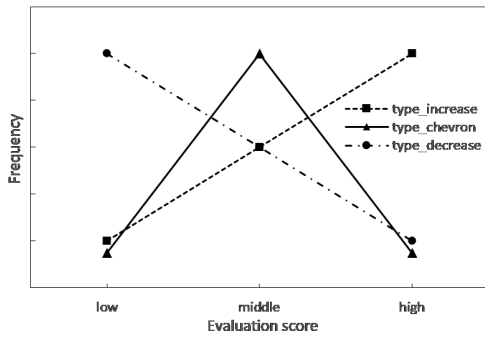
Fig. 1. Image of frequency distribution pattern. In the figure, *type_increase* shows "More frequently in high evaluation," *type_decrease* shows "More frequently in low evaluation," *type_chevron* shows "Medium evaluation is the most frequent, low frequency in high and low evaluations."

Fig. 2. Average of frequency ratio of words obtained in the evaluation experiment for each corpus.

III. ANALYSIS METHOD

In our method, we use the the frequency distribution pattern for analysis. Fig. 1 shows the frequency distribution pattern for each evaluation value. Using the distribution pattern, words with three types of frequency can be extracted: "more frequently in high evaluation," "more frequently in low evaluation," and "medium evaluation is the most frequent, low frequency in high and low evaluations." Therefore, it is considered that various opinions about the evaluation points can be extracted with the proposed method. It is expected that we can obtain the findings such as "for the same review point mentioned in highly evaluated reviews, what is the matter in the lowly evaluated review."

A. Division of review data and calculation of word frequency in each corpus

Let the corpus of the reviews for a certain review object h be c. In this paper, it is assumed that a discrete value of 1 to 5 is given to each value. Review sentences on the review target h are classified into three corpora according to the evaluation value. The corpora for evaluation values 1 (corpus low), values 2 to 4 (corpus middle), and value 5 (corpus high) are prepared.

The frequency of noun w in corpus c for the review target h is defined as $Freq_c^h(w)$. Since the number of reviews and nouns differs for each corpus, $Freq_c^h(w)$ is normalized by the total number of nouns in the corpus. The frequency of the noun w normalized for each corpus is defined as $StFreq_c^h(w)$ and calculated as the following equation (1);

$$StFreq_c^h(w) = \frac{Freq_c^h(w)}{N_c}, \quad (1)$$

where, $c \in \{low, middle, high\}$ and N_c represents the number of all nouns in each corpus.

The analysis target is the noun w that commonly appears in the three corpora. The proposed method detects the evaluation point w which is likely to have diversity in opinions by relational expressions using $StFreq_c^h(w)$.

B. Extraction of word set according to each frequency distribution pattern

In order to extract nouns satisfying *type_increase* in Fig. 1, the following conditional expressions (2) and (3) are used.

$$\frac{StFreq_{high}^h(w)}{StFreq_{middle}^h(w)} > 1.0, \quad (2)$$

$$\frac{StFreq_{middle}^h(w)}{StFreq_{low}^h(w)} > 1.0. \quad (3)$$

A noun w that satisfies both of the expressions (2) and (3) is extracted. The extracted noun should be more frequent with the increasing the evaluation values. The nouns obtained with these expressions are assumed as the word set set_{inc}^h.

For extracting nouns satisfying *type_increase* in Fig. 1, the following conditional expressions (4) and (5)are used;

$$\frac{StFreq_{low}^h(w)}{StFreq_{middle}^h(w)} > 1.0, \quad (4)$$

$$\frac{StFreq_{middle}^h(w)}{StFreq_{high}^h(w)} > 1.0. \quad (5)$$

By satisfying both the expressions (4) and (5), a word set set_{dec}^h, in which the frequency of noun decreases with the increasing evaluation values, is obtained.

Also, for *type_chevron* in Fig. 1, we use the following conditional expressions (6) and (7);

$$\frac{StFreq_{middle}^h(w)}{StFreq_{high}^h(w)} > 1.0, \quad (6)$$

$$\frac{StFreq_{middle}^h(w)}{StFreq_{low}^h(w)} > 1.0. \quad (7)$$

When both of the expressions (6) and (7) are satisfied, the word set set_{chev}^h is obtained. The obtained words in set_{chev}^h are the most frequent in medium evaluation and low frequent in high and low evaluations.

The opinions including nouns in in the word sets set_{inc}^h, set_{dec}^h and set_{chev}^h are extracted from each corpus. The nouns in those word set should be the major point of review in either of the three corpus and minor in other corpora. So we expect that the extracted opinions should have various aspects for each corpus.

IV. EXPERIMENT AND CONSIDERATION

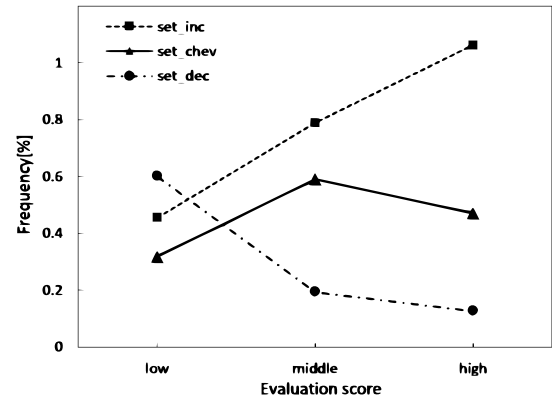In the experiment, the review data for analysis is part of Rakuten Travel's review data included in Rakuten Public

TABLE I

EXAMPLES OF WORDS EXTRACTED FROM THE REVIEW OF A HOTEL. FOR EACH SET, THE TOP TEN WORDS WITH HIGH VALUES OF $StFreq_c^h(w)$ ARE SHOWN IN THE TABLE. TRANSLATED ENGLISH IS SHOWN IN BRACKETS AFTER THE ORIGINAL JAPANESE WORD; SOME OF THE WORDS ARE NOT ASSUMED AS NOUNS IN ENGLISH THOUGH.

set_{inc}^h	set_{dec}^h	set_{chev}^h
部屋 (room)	フロント (front desk)	の (one)
利用 (use)	時 (time)	便利 (useful)
宿泊 (staying)	人 (people)	立地 (location)
こと (thing)	サービス (service)	駅 (station)
今回 (this time)	よう (sense)	品川駅 (Shinagawa Station)
大変 (tough)	さ (-ish)	前 (front)
東京 (Tokyo)	ため (for)	階 (floor)
泊 (staying night)	スタッフ (staff)	非常 (extraordinary)
出張 (business trip)	もの (thing)	1 (one)
チェックイン (check-in)	翌朝 (next morning)	一 (one)

Data in Japanese, which Rakuten Co., Ltd. provides for research purposes in cooperation with the National Institute of Informatics. The analysis target is 69,803 sentences obtained from the reviews for the 20 hotels randomly selected. The information used in these review data is "posted text," "overall evaluation value," "facility number."

A. Extraction of word set

Table I shows an example of the word set obtained by the proposed method. The words shown in the table are the top ten with high values of $StFreq_c^h(w)$ in the word set.

set_{inc}^h includes words common to all facility such as "room" and "accommodation," while set_{dec}^h is related to communication like "front" and "service." In set_{chev}^h, we can see some characteristic words concerning the contents outside the accommodation facilities such as "station" and "location."

Fig. 2 shows a graph of the proportion of the average value at each accommodations for $StFreq_c^h(w)$ of each word included in the obtained word set set_{inc}^h, set_{dec}^h and set_{chev}^h. The maximum value of $StFreq_c^h(w)$ was 6.153 and the minimum value was 0.003. The calculation method was the same as the equation (1). In set_{inc}, $StFreq_c(w)$ decreases by approximately 0.3 % as the evaluation decreases.

In set_{dec}, the average value of $StFreq_c(w)$ in the corpus low is significantly different from the corpora middle and high by 0.5 % or more. The words included in set_{dec} are rarely appeared excepting for corpus low. Thus it means that those words are characteristically appear in reviews with low rating. The distribution tendency is mostly similar to the image shown in the Fig. 1. We suggest that we could design the mathematical modeling as we expected.

B. Extraction of opinions

Opinions including words in word set were extracted from each corpus of each accommodation facility. As the data for the consideration, 30 sentences were randomly taken out from each obtained sentence set.

Table II shows examples of the opinions extracted by the proposed method. Opinions were randomly selected from sentences used for consideration with regardless of facilities.

At first, we would like to consider the opinions including words in set_{inc} From the corpus high, overall evaluation such as "the room is clean" and "satisfied" is obtained. Also, there is a characteristics in description style: the sentence describes the multiple items in one sentence. However, the details on travel and correspondence depend on the facility. The opinions obtained from corpus low tended to state the detailed reasons of evaluation regarding the evaluation points. Referring to a review with a low evaluation about the evaluation points evaluated in the highly rated review, it was suggested that the detailed opinions on the evaluation points can be obtained.

For set_{dec}, it contains a lot of unpopular opinions. Also we found that features of description way is different between corpus low and others. Many sentences included in the corpus low had no subject to be evaluated. That is, the sentence showed only the evaluation part "too bad" and "worst" or just a part of conversation. On the other hand, opinions indicating "what was wrong" were extracted from corpus middle and high. This pattern makes it possible to judge on the review object not only from unpopular opinions and requests but also from various perspectives such as reasons and favorable opinions for the same object.

In the case of using set_{chev}, a characteristic evaluation points appeared. The extracted points are different for each facility and not in opinions extracted by using set_{inc}. The extracted opinions contained the details of the rooms and facilities such as "bath amenities" and "elevator." And, it was suggested that the description of the opinions reflected the characteristics of the hotel and the individual's prejudice.

Opinions in this case were not limited to popular and unpopular, then it was expected that the detailed opinions for various contents were stated. Although such information might be suitable for individual needs, it is necessary to narrow down the information according to the user's context.

V. CONCLUSION

This paper proposed a method for extracting diverse opinions by focusing on the words with different frequency for each evaluation value on the web review. Also, we examined the contents of opinions obtained by the proposed method. The analysis result showed that it was successful to extract the words of the target distribution. Moreover, it sugested the possibility to extract opinions complying with the evaluation point the individual sense.

On the other hand, it is necessary to narrow down the detailed classification of nouns. There is also a problem that meaning of sentences can not be grasped due to bias of output number of words for each evaluation value and division way of sentences. In our future, to solve the above problem, the analytical method would be improved for incorporating characteristic of sentences such as conditional statements showing the users' situation.

REFERENCES

[1] L. P. Jorge Carrillo de Albornoz and P. Gervs, "A hybrid approach to emotional sentence polarity and intensity classification," in *Proc. of the Fourteenth Conference on Computational Natural Language Learning*, 2010, pp. 153–161.

[2] P. D. Turney, "Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews," in *Proc. of the 40th Annual Meeting on Association for Computational Linguistics(ACL)*, 2002, pp. 417–424.

[3] J. W. Theresa Wilson and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in *Proc. of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, 2005, pp. 347–354.

[4] S. L. Kushal Dave and D. M. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," in *Proc. of the 12th international conference on World Wide Web*, 2003, pp. 519–528.

[5] A.-M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in *Proc. of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, 2005, pp. 339–346.

[6] S.-M. Kim and E. Hovy, "Automatic identification of pro and con reasons in online reviews," in *Proc. of the COLING/ACL on Main conference poster sessions*, 2006, pp. 483–490.

[7] J. W. Ellen Riloff and T. Wilson, "Learning subjective nouns using extraction pattern bootstrapping," in *Proc. of the seventh conference on Natural language learning at HLT-NAACL 2003 - Volume 4*, 2003, pp. 25–32.

[8] B. L. Ramanathan Narayanan and A. Choudhary, "Sentiment analysis of conditional sentences," in *Proc. of the 2009 Conference on Empirical Methods in Natural Language Processing*, 2009, pp. 180–189.

[9] E. Hoque, S. Joty, L. Marquez, and G. Carenini, "Cqavis: Visual text analytics for community question answering," in *Proc. of the 22nd International Conference on Intelligent User Interfaces*, 2017, pp. 161–172.

[10] P. S. Michael Kaschesky and G. Bouchard, "Extracting product features and opinions from reviews," in *Proc. of the 12th Annual International Conference on Digital Government Research*, 2011, pp. 317–326.

2078-0958

TABLE II
EXAMPLES OF OPINIONS OBTAINED WITH WORD SETS. THE SENTENCES IN THE TABLE SHOWS THE ORIGINAL SENTENCES EXTRACTED FROM THE
DATABASE. THE BOLD WORD IS THE KEYWORD TO EXTRACT THE SENTENCE, WHICH IS THE WORD IN THE WORD SET OBTAINED BY THE PROPOSED
METHOD. THE TRANSLATED ENGLISH SENTENCE IS SHOWN AFTER THE ORIGINAL JAPANESE SENTENCE IN THE BRACKETS.

Word set	Corpus	Obtained opinions
set_{inc}	high	駅からも近く、ホテル自体もとても綺麗でした (It was close to the station and the hotel itself was also very **beautiful**) お部屋はカジュアルツインの予約だったのにデラックスツインを用意いただいてラッキーでした (I was lucky to stay in a deluxe twin although the reservation was for a casual **twin**) 結果は大満足です (The result is very **satisfying**)
	middle	お部屋は狭かったのですが、ベッドはよかったので疲れもとれました (The room was small, but the bed was good and I feel refreshed) 部屋の広さ・立地など価格に見合った満足感だと思います (The largeness of **room** and the location are reasonable matching the price) GW の５月３日直前予約でお部屋が空いておられたのは、とってもラッキーでした (It was so lucky that I could book the room just before May 3rd in a long holidays)
	low	朝食は、値段の割に品数が少ないと思いました (I thought the number of items in the breakfast is small compared to the **price**) 素泊まりプランで利用をさせていただきました (We used that with accommodation **plan** without meal) 朝食のウリであるワッフルも品切れしたまま補充なし (The waffle which is a selling point of **breakfast** is out of stock and it is not replenished)
set_{dec}	high	ホテルの問題ではないですが、意外に駅から遠かったです (It is not a **problem** of the hotel, but it was unexpectedly far from the station) 朝食はブッフェ式で品数はそんなに多いわけではないが、味・内容ともに断満足でした (Breakfast is buffet style and the number of items is not so much, but both taste and **content** were satisfying) 多少トイレが古いタイプの感じですが，十分に使えます (The toilet is somewhat old type **feeling**, but it can be used enough)
	middle	ちょっとさすがに部屋のお風呂に入るのが気持ち悪いなと思い、大浴場へ (I went to the big bath because I felt the bath in the room was yuck) 上の階の方 (子供？) が飛び跳ねているのか寝れませんでした (I could not sleep because the **people** (child?) on the upper floor was jumping) 今回初めての和室でしたが、風呂場の換気扇の音の大きさに驚きました (It was the first Japanese-style room this time, but I was surprised by the volume of the **sound** of the ventilator in the bathroom)
	low	今まで相当数のシティホテルに宿泊しましたが、残念ながら不満でした (I have stayed in number of city hotels until **now**, but unfortunately I couldn't be satisfied with this) その間もホテルからは何のアナウンスも無く警報も止まりません (In the meantime, there is no announcement from the **hotel** and the alarm can not stop) 他の客は案内有りなのになぜですかね？ 楽しい気分を害された感じで最悪 (Why did they have guidance for **other** guests? It was the worst because I felt my pleasant feeling was harmed)
set_{chav}	high	車は少し離れのパーキングに駐車となりますが、ホテルに荷物を預けてパーキングに行くことが出来ます (The car will be parked at a short distance from the hotel, but you can ask the hotel to keep your **luggage** to go to the parking) 娘と何度か利用させて頂いていますが、とてもお気に入りのホテルです (I **use** it several times with my daughter, it is a very favorite hotel for us) お風呂もエレベーターも宿泊者しか使えないようにしてありました (Both the bath and **elevator** were available for only guests)
	middle	チェックアウト前に 荷物を送ろうと荷物を詰め そのブースに行きました (I tried to send my **luggage** before checkout and packed **luggage**, and went to the booth) ビジネスでしたが、繁華街も近く（ごはんマップなども参考にしました）良かったです (It was business, but It was good because nearby the downtown (used Meal Map and others as reference). シャワーとトイレが一体型のため、子供は使いづらかったようです (**Shower** and bath are in one unit, so it seemed to be hard for children to use)
	low	良かったのは立地だけ、自社の駐車場に車を止めるのに 900 円も徴収するし (The good point is only location, they took 900 JPY to park at their parking) 部屋はビジネスホテルのそれ (The **room** is that of a business hotel) 全体的にスタッフの印象、すごく悪かったです (**Overall**, the impression for the staff was really bad)