Data Collection for Investigation of Reliable Reviews

Jun Kikuchi
Software Engineering Lab, University of Aizu
Tsuruga, Ikki-machi, Aizu-Wakamatsu City,
Fukushima, 965-8580 Japan
(+81)242-37-2603
s1190017@u-aizu.ac.jp

Vitaly Klyuev
Software Engineering Lab, University of Aizu
Tsuruga, Ikki-machi, Aizu-Wakamatsu City,
Fukushima, 965-8580 Japan
(+81)242-37-2603
vkluev@u-aizu.ac.jp

ABSTRACT
Nowadays the Internet is growing quickly. People use for many different purposes, mainly for online shopping. However, there is a large number of websites, which offer similar products and features. It makes a choice for customers difficult. As a result, they refer to online reviews. The main purpose of this research is to improve a reliability of a review. Increasing reliability is useful and helpful for customer's decision, and it can also filter out spam reviews. We are detecting a pattern by morphological analysis. Obtaining a reliable review pattern could be used to collect useful reviews automatically. Results of this study can be applied to any kinds of reviews and opinions.

Categories and Subject Descriptors
H.3.3[Information Storage and Retrieval]: Information Search and Retrieval – *information filtering*

General Terms
Algorithms, Experimentation, Languages, Theory

Keywords
pattern, text mining, opinion extraction

1. INTRODUCTION
Nowadays, many people access to the Internet for many different purposes. The main purpose is to obtain useful and helpful information about products, places, restaurant, and etc. There are a lot of websites and web applications which provide this information, and most people rely on them as important tool because it is easy to exchange opinion between individuals. In addition, many people search for products and buy them online, so reviews are important and necessary because they provide customers with an easy way to evaluate a product. Changing a style of purchasing makes review websites and applications to analyze user features more accurate. These review sites and applications, however, do not cover all kinds of products, services, or places which customers want to know in detail. Moreover, these review websites and applications utilize different evaluation approaches. Another problem is that we are not certain that these reviews are reliable and precisely.

*IWAIT'*15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.

To evaluate a product, we need a measurement and an evaluation model. For example, people could not rely on a certain review which is provided by a few users because reviews reflect a personal opinion. This research collected certain amount of opinions and reviews expressed by people who experienced or used products or services before to seek some features of useful reviews.

Our study goal is to create a valuable and reliable review evaluations system in different fields. We are planning to detect standardized patterns of these reviews. The first step is creating and developing a review corpus. We will find features of valuable review by part of speech (POS) tagging. We will also make these reviews up-to-date and current because people always need actual information to help in their decision making processes. One solution is social networking sites (SNS). The SNS provides customers with an opportunity to exchange information easily by any kinds of devices. The SNS are creating different kinds of information quickly in large quantities. Another merit of using this kind of data is that these opinions are very frank and clear, so product evaluation is more reliable. An application which collects and analyzes big data for gathering better reviews is necessary and helpful in the current circumstance.

2. RELATED WORK

2.1 Opinion Mining
Opinion mining applied to the review analysis deals with types of reviews and problems in this area. Reviews have many kinds of writing style and a large variety of words, so opinion mining concentrates on the automatic analysis of the reviews in natural language.

Generally, opinion mining organizes and obtains an entity, aspect, and opinion orientation. An entity is a target object such as a product that has been evaluated by someone. An aspect mentions a component or a sub-component of the entity. An opinion orientation expresses positive, negative or neutral attitude in an opinion sentence. These three components are essential to analyze any kind of reviews. This basic organization of reviews has been used to solve a lot of problems of the review analysis such as sentence subjectivity, sentiment classification, lexical expansion, and etc. [1]

Reference [2] used a support vector regression (SVR) based outlier detectors which designed by collecting reviews on the Internet. The SVR detector predicts a score of hotel ranking to

improve a reliability of hotel ranking. This research did morphological analysis to collect an opinion dictionary using part of speech (POS) pattern. This dictionary focused on patterns of adverb and intransitive adverb. Therefore, a feature value assigned each opinion word in this dictionary. As a result, the SVR detector witch trained by the opinion dictionary also has a feature of TF-IDF. These features help to analyze reviews on the Internet and to measure reliability.

3. APPROACH

3.1 Implementation

We analyzed several kinds of reviews on the Internet, and investigated critical and important information that a better review has. Accordingly we collected reviews from the Amazon.co.jp website. Amazon offers many kinds of products. In this study, our objection is to find out features of useful reviews, so we created a review corpus from gathering reviews. We gathered 169 reviews on the Sporting Goods from Amazon.co.jp. In addition to, it has implemented a rating system that asks customers for a feedback on a review. We also used this rating system for judging whether a review is useful or not.

Amazon provides developers with the APIs to obtain information for products, so we used it to obtain title, reviews, products' name, and reviews' rating to analyze patterns or template for good reviews. All reviews are only in Japanese. We analyzed opinion sentences with MeCab and TermExtract [3] in order to look at POS's features from collected reviews. Based on this POS tagging, we investigate a reliable and useful review.

3.2 Data Collection

We divided all gathered reviews into two groups. One includes opinion sentence, and another one consists of non-opinion sentences. We defined a difference between a review and an opinion. An opinion should include a target and subjective expression. An opinion target must be related a product or this aspects, and subjective expression are personal feelings, views or beliefs. In this study, opinion sentences are defined by two main types: regular opinion and comparative opinion. A regular opinion commonly expresses an opinion on any aspects of products. On the other hand, a comparative opinion expresses a relation of similarities or differences between two or more products or aspects. All opinion sentences commonly include positive, negative, or neutral sentences, emotion, attitude, etc. Table 1 summarizes collected Sporting Goods reviews. It shows two data. One is a number of reviews in Sporting Goods category. Second data shows numbers after calculating an average of positive feedback numbers and an average of all feedbacks number. A rate of reviews must be evaluated by more than 5 customers to obtain more accurate template. In this case, about 24 customers evaluated a review, and about 22 customers think a review is useful on average. Table 2 shows summarizing data on two types of sentences: opinion sentences or non-opinion sentences.

Table 1. Statistics of Sporting Goods Review Data from Amazon

Collection	Number of Reviews	Number of Positive Feedbacks / Number of All Feedbacks
Sporting Goods	169	22.54/24.51

Table 2. Statistics of Sporting Goods Review for each Review Group

Collection	Number of Sentences	Number of Words
All Sentences	839	27193
Opinion Sentences	532	17299
Non-Opinion Sentences	307	9894

3.3 Collected Data Analysis

All collected reviews are rated as a useful review by Amazon customers, so it is critical to look at any features in these collected reviews. Opinion sentences are more informative. They are focused on the product. Therefore, we analyzed only collected opinion sentences. As a first step of the analysis, we did POS tagging to all opinion sentences with MeCab and TermExtract. Important information of a review is opinion targets, subjectivity and emotions. A noun could indicate an opinion target or an aspect, and an adjective also could indicate emotions and subjectivity. Because of these reasons, we concentrated on two POSs, which are noun and adjective. In this study, these two word lists are created from collected opinion sentences. Table 3 shows the result after gathering on two POSs words.

Using these created POSs lists, we automatically collected sentences which include two POSs as a keyword from all collected reviews. This second step of the analysis identifies how these two POSs relate to important opinion sentences. Moreover, we investigated what kinds of words are necessary for a review corpus. Table 4 summarizes the result of the analysis. It shows a number of collecting sentences as an opinion sentence (OS) by each word lists and a number of correct opinion sentences. Correct opinion sentences include a manually collection of opinion sentences, and have at least one word in POSs lists. Accuracy is calculated by dividing a number of correct opinion sentences by a number of all collected sentences.

Table 3. Statistics of Nouns and Adjectives

Collection	Number of Words
Nouns	277
Adjectives	33

Table 4. Statistics after each POS collection

Collection	# OS	# Correct OS	Accuracy
All Noun	751	437	58.20
Top 20 Noun	563	315	55.95
Top 10 Noun	486	267	54.94
All Adj	469	284	60.55
Top 20 Adj	400	263	65.75
Top 10 Adj	365	242	66.30
Top 10 Noun AND Top 10 Adj	229	141	61.57
Top 10 Noun OR Top 10 Adj	622	367	59.00

4. DISCUSSION

Collected data on noun and adjectives show us two points. One point is that noun keywords have too many varieties and less duplication to obtain better collection of opinion sentences. In this case, a noun list has many words, so we obtain automatically correct opinion sentences with a high rate. Each noun collection, however, is not better accuracy than adjective collections. Adjective keywords could not gather much number of opinion sentences, but the collection of top 10 adjective keywords result on a highest accuracy. In this search area, adjective always treat as a special keyword because an opinion sentence mainly expresses based on adjective words. Therefore, an adjective word is an important issue to achieve a better pattern of reviews, so we need a deeper analysis about adjectives. Moreover, combination of nouns and adjectives shows an interesting and better result. Collected opinion sentences that use top 10 noun keywords or top 10 adjective keywords have a good number of correct opinion sentences. It means that we need to inspect opinion sentences in details in order to know a relation between noun keywords and adjective keywords.

5. CONCLUSION

We collected a Japanese corpus of reviews from Amazon.co.jp. We conducted a preliminary analysis of every reviews manually and automatically to find the common features or patterns. We concentrated on analysis of opinion sentences. An analysis of opinion sentences in the way of inspecting adjectives and nouns shows relation between two POSs and opinion sentences. This data is necessary information for creating a better review corpus. However, using two POS words picks many non-opinion sentences up, and it missed a lots opinion sentences as well.

Therefore, this analysis needs deeper research to seek a feature or pattern. To increase the accuracy in obtaining correct opinion sentences, we need to analyze details of nouns, adjectives, and relations of two types of words. Moreover, it is necessary to study that these keywords could achieve effectively opinion sentences in different categories.

6. REFERENCES

[1] Liu, B. and Zhang, L. 2012. A survey of opinion mining and sentiment analysis. *Mining Text Data*, 415-463.

[2] Hsieh, H. Klyuev, V. Zhao, Q. and Wu, S. 2014. SVR-based outlier detection and its application to hotel ranking. *In Proc. of the 2014 IEEE 6th International Conference on Awareness Science and Technology (iCAST)*, 1-6.

[3] Miyashita, M. and Klyuev, V. 2014.TermExtract: Accuracy of Compound Noun Detection in Japanese. *Future Information Technology*, 189-194.