Appraisal Expression Recognition Based on Generalized Mutual Information

Huan Huang
Central China Normal University/National Engineering Research Center for E-Learning, Wuhan, China
Email: hbdx_hh@yahoo.com.cn

Qingtang Liu and Tao Huang
Central China Normal University/National Engineering Research Center for E-Learning, Wuhan, China
Email: {liuqtang, tmht}@mail.ccnu.edu.cn

***Abstract*—The polarity of a subjective sentence in customers' reviews depends on not only the semantic orientation of opinion words, but also their context, especially the modified features. An Appraisal Expression consists of an opinion word and the modified feature, and can accurately describe the holders' opinion about the feature. In this paper, we propose a novel method that uses generalized mutual information to automatically recognize the appraisal expressions from customers' reviews. Our method does not fill in any template and is domain independent. More important, our method can avoid the complex syntactic analysis while keeping the comparable accuracy, which greatly improves the efficiency of appraisal expression recognition. Our experimental results show that the F-measure of our method is up to 80.18% and 80.08% respectively, which is higher than the existing methods, and the efficiency is also comparable to the past methods.**

***Index Terms*—Appraisal Expression, Generalized Mutual Information, Sentiment Analysis**

I. Introduction

As the rapid expansion of e-commerce, the product reviews provided by users on the Internet is becoming more and more. These reviews express the customers' sentiment and opinions about the products or services. With the help of these reviews, potential customers can make a decision on whether to buy the products, and manufacturers can find the products' drawbacks to improve the design of products, and product sellers can design an appropriate marketing strategy. However, as the product reviews become more and more, it is difficult to read by potential customers. The large number of reviews also makes it hard for product manufacturers and sellers to keep track of customers' opinions on their products.

Sentiment Analysis, also known as Opinion Mining or Orientation Mining, means using the computer to process and analyze the subjective information on the Internet, and automatically recognizes the holders' sentiment and opinions about some persons, events, or products [1]. Sentiment Analysis can not only used to mine customers' opinions from product reviews, but also used to analyze web public sentiment from the posts of BBS, news comments, blogs and micro-blogs [2-3]. Sentiment Analysis derives from the identification of subjective words and their semantic orientation [4-5]. For example, how to automatically identify that the words *beautiful* and *ugly* are subjective and their polarities are positive, and negative respectively. Based on the identification of subjective words and their polarities, researchers gradually used word's polarity and machine learning methods to automatically predict the polarity of a subjective sentence or a short text, which is usually called sentiment classification [6-7]. Support Vector Machine (SVM) is the common algorithm used in sentiment classification, and the new development of SVM was introduced by LIU Taian [8]. However, a subjective sentence may contain more than one topic, and the holder may have different sentiment or opinion on different topic, thus it is unreasonable to simply decide the polarity of a sentence to be positive or negative. Moreover, if we can know the users' opinion on different features of a product, it may be more useful. According to these, Liu B. et al. proposed feature-based opinion mining [9-12]. These methods described an opinion as a four-tuple [Topic, Holder, Claim, Sentiment], in which the Holder believes a Claim about the Topic, and in many cases associates a Sentiment, such as good or bad, with the belief [9]. According to these four elements, the feature-based opinion mining needs some fundamental tasks to support, such as Holder Identification, Topic/Feature Mining and Opinion/Sentiment Word Extraction. Recently, researchers have proposed a lot of excellent methods for Feature Mining and Sentiment Word Extraction [13-18]. However, whether the polarity of a subjective sentence is positive, negative or neutral depends on not only the semantic orientation of opinion words, but also their context, especially the topics or features they modified. For example, the polarity of the subjective sentence, *the battery life is very long*, is positive, but the polarity of another subjective sentence, *the camera has long startup*, is negative. Although both of them have the same sentiment word *long*, they have different polarity. So except Topic/Feature Mining and Sentiment/Opinion Word Extraction, a more important

doi:10.4304/jcp.8.7.1715-1721

thing we need to do is to relate the Topics/Features with their appropriate Sentiment/Opinion Words. This fundamental task was called Appraisal Expression recognition in Bloom, Garg and Argamon's work [20]. Appraisal Expression is also called Private State by Wiebe and Wilson [19].

In this research, we focus on recognizing the Appraisal Expressions from customers' reviews. The Appraisal Expression is defined as a tuple [Topic, Sentiment Phrase], in which the Topic refers to a feature of products, and the Sentiment Phrase consists of sentiment words and some decorators, such as negative or degree adverbs. We first analyzed the existing methods of appraisal expression recognition, and then proposed a novel method based on generalized mutual information. Our method of appraisal expression recognition performs in two steps: 1) use the techniques of association rule mining and some filtering strategies to identify the opinion features and opinion words, 2) and then calculate the generalized mutual information of each pair of opinion feature and opinion word in the same sentence, and choose the pairs which exceed the threshold as the proper appraisal expression. Our method is based on the observation that the opinion features and opinion words appear frequently in the reviews and the collocation of opinion features and opinion words is usually stable. Our experimental results with a large number of customer reviews of 2 types of products on Amazon show that the F-measure of our method is respectively up to 80.18% and 80.08%, and the efficiency is also comparable to the past methods.

II. Related Work

Although appraisal expression was firstly presented in Bloom, Garg and Argamon's work [20] until 2007, there has been some related work before. Until now, the methods of appraisal expression recognition fall into three categories: the methods based on distance, the methods based on hand-crafted templates and rules, and the methods based on syntactic path.

The distance-based methods used the distance between opinion feature and opinion word to identify the appraisal expressions. Kim and Hovy firstly identified the opinion holders using the techniques of Name Entity Identification, and then chose the adjective between the opinion holder and the given topic as the opinion word [9]. Hu and Liu firstly used the techniques of association rule mining and some filters to extract the opinion features, and then chose the nearby adjective as the feature's effective opinion [10]. The distance-based methods are simple to implement, however, the accuracy is not high because most of time opinion word is far away from its topic.

The second methods proposed to manually craft some templates or rules according to the relationship between the opinion feature and its opinion word, and then used the techniques of pattern matching to recognize the appraisal expressions in one-time. Kobayashi et al. described the evaluative expression as a similar triple [evaluated subject, focused attribute, value], and summarized 8 co-occurrence templates to represent the relationship among the evaluated subjects, focused attributes and value. At last, they used the co-occurrence templates to collect the evaluative expressions [21]. Bloom et al. applied the Stanford Parser to analyze the syntactic relationship between opinion features and their opinion words. They manually crafted 31 linkage specifications to recognize the appraisal expressions [20]. Similar to Bloom's work, Popescu et al. used MINIPAR Parser to analyze the dependency relationship of many subjective sentences, and manually summarized 10 grammatical templates to recognize the appraisal expressions [22]. The template/rule-based methods used the co-occurrence relationship or syntactic relationship between opinion features and their opinion words, which greatly improved the accuracy of appraisal expression recognition. However, it is difficult to manually summarize enough templates or patterns to cover all relationship between opinion features and their opinion words, so the recall rate is not very high.

In order to reduce the difficulty of manually crafting templates and rules, Zhao et al. proposed a method that used syntactic path to automatically recognize the appraisal expressions [23]. First, the syntactic paths are automatically collected to describe the relationship between the polarity words and their corresponding targets. Next, an exact syntactic path matching method and an edit distance based syntactic path matching strategy are used to recognize the appraisal expressions. This method can automatically learn syntactic path templates from corpus using statistical method, and avoid to manually crafting syntactic path templates. However, it depends on the techniques of syntactic analysis, which are not mature and the efficiency is low. Besides, this method used the HowNet to recognize the opinion words firstly, which does not contain enough opinion words.

Motivated by the above remarks, we attempt to use generalized mutual information for automatic recognition of appraisal expressions. Our method is different from the distance-based methods, because our method uses generalized mutual information to relate the opinion words with opinion features. Our method considers the co-occurrence pattern between opinion words and opinion features, so it is more accurate. Our method is also different from the methods based on rules or syntactic paths, because our method does not fill in any template and is domain independent. More important, our method can avoid the syntactic analysis but keep the comparable accuracy, which greatly improves the efficiency of appraisal expression recognition.

III. The Proposed Method Based on Generalized Mutual Information

Figure 1 gives the architectural overview of our method of appraisal expression recognition based on generalized mutual information. It mainly performs in two steps: firstly identify the opinion features and opinion words from the review database using POS Filter, Frequency Filter and Redundancy Pruning; secondly calculate the generalized mutual information of each pair of opinion feature and opinion word in the same sentence.

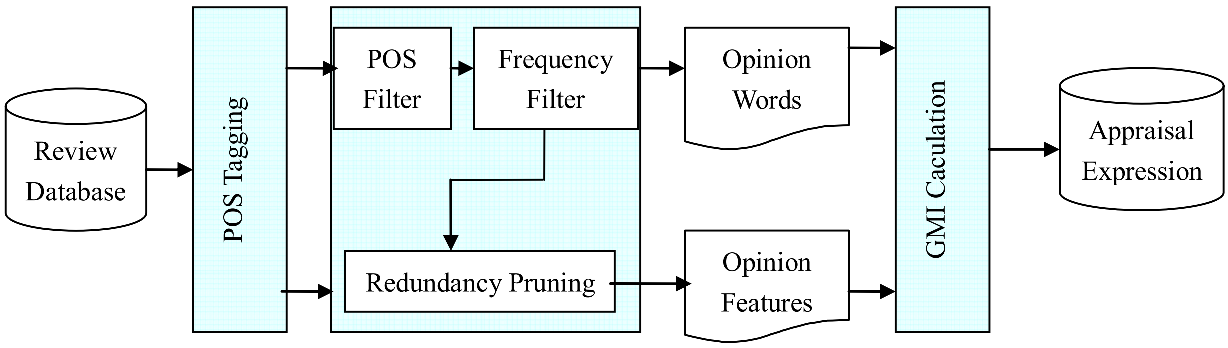


Figure 1. Architecture of the proposed method

A. Opinion features and opinion words identification

The existing researches have suggested that the opinion features are usually nouns or nouns phrase in review sentence, and the opinion words are usually adjectives or adverbs. Moreover, the opinion features and opinion words frequently appear in one kind of product reviews. According to these, we firstly split each review into sentences and produce the part-of-speech tag for each word. After that, we choose the nouns and nouns phrases as the candidates of opinion features, and choose the adjectives or adverbs as the candidates of opinion words. Finally, we use frequency filtering strategy and redundancy pruning techniques to remove meaningless and redundant opinion features and opinion words.

Currently there have been a lot of perfect methods and tools for text segmentation and part-of-speech (POS) tagging. We use the Natural Language Toolkit (NLTK) [25] to split each review into sentences and produce the part-of-speech tag for each word. NLTK is a leading platform for building Python programs to work with human language data. It provides easy-to-use interfaces to over 50 corpora and lexical resources such as WordNet, along with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning. The following shows a sentence with POS tags.

[('it', 'PRP'), ('is', 'VBZ'), ('amazing', JJ), ('that', 'IN'), ('the', 'DT'), ('battery', 'NN'), ('lasts', VBZ), ('so', 'RB'), ('long', 'JJ'), ('when', 'WRB'), ('the', 'DT'), ('phone', 'NN'), ('is', 'VBZ'), ('so', 'RB'), ('small', 'JJ'), ('and', 'CC'), ('light', 'JJ'), ('.', '.')]

As some opinion features are nouns phrases, thus we further use shallow analysis to identify the nouns groups, such as *battery life*, *color screen*, *voice dialing* and so on. After the POS tagging and simple chunk parsing, we extract the nouns and nouns phrases as the candidate set of opinion features, and extract the adjectives and adverbs as the candidate set of opinion words. The following shows the POS tags of candidate opinion features and opinion words.

OPINION FEATURES: NN, NNS, NNP, NNPS, NP

OPINION WORDS: JJ, JJR, JJS, RB, RBR, RBS

The opinion features and opinion words discovered by POS Filter contain a lot of meaningless and redundant items, which will reduce the efficiency and accuracy of appraisal expression recognition later. Thus we use two strategies to remove the meaningless and redundant items. The first strategy is frequency filtering, and the second strategy is redundancy pruning. The reason of using frequency filtering is the observation that the words converge when customers comment on the features of the same kind of products. Thus using frequency filtering can find the items which are more likely to be opinion features and opinion words. Although frequency filtering can remove some infrequent opinion features and opinion words, it is not important for the results because most of people do not care of them. Redundancy pruning is mainly used to remove the redundant opinion features that contain single words. To describe the meaning of redundant opinion features, Liu et al. proposed a concept *p-support* (pure support) [10]. If the *p-support* value of an opinion feature is lower than the minimum *p-support*, and the feature is a subset of another feature phrase, it is pruned.

B. Recognize the appraisal expressions using generalized mutual information

The same opinion word may have the different semantic orientation when modifies the different features or be in different context. After identify the opinion features and opinion words, we should proceed to relate the opinion words with its opinion features.

According to analyzing the review corpus, we found that an opinion word is usually not appropriate to modify all opinion features, and the modified opinion features is usually stable. It means that opinion features and opinion words are interdependent. Mutual information statistic was popularly used to measure the interdependence of two signals in a message. If we look the opinion feature and opinion words as two random variables, then we can use their mutual information as a measure of their interdependence. Although mutual information statistic can reveal the interdependence between opinion features and opinion words, it can not make use of the contextual information. Thus Magerman and Marcus proposed an improved method named generalized mutual information to measure the interdependence of two adjacent n-grams [26]. Generalized mutual information is actually the weighted average of mutual information with different contextual window size. For example, for a window size w=4, give the context $\chi_1\chi_2\chi_3\chi_4$, the generalized mutual information of χ_2 and χ_3 is:

$$GMI(\chi_2,\chi_3)=k_1MI(\chi_2,\chi_3)+k_2MI(\chi_2,\chi_3\chi_4)+k_3MI(\chi_1\chi_2,\chi_3)+k_4MI(\chi_1\chi_2,\chi_3\chi_4)\,. \quad (1)$$

The purpose of generalized mutual information is to identify the constituent boundary in a sentence, so Magerman and Marcus used the standard deviation of the values of the bigram mutual information vector of an n-gram as the weighting function. However, it is not suitable for our method, because the opinion features and opinion words are usually not adjacent.

We suppose that 1) if a pair of opinion feature and opinion word is a correct appraisal expression, its mutual information should be larger than the mutual information with some context; 2) and the mutual information between opinion feature and opinion word is more important than the mutual information with some context. Thus we use the ratio of the value of mutual information against the sum of all mutual information as the weighting function.

In summary, if a customer review is represented as $D = w_1w_2w_3\ldots w_i\ldots w_{j-1}w_jw_{j+1}\ldots$, the term w_i is an opinion feature, and the term w_j is an opinion word, $i\neq j$, for a window of size n, the generalized mutual information of w_i and w_j is defined to be:

$$GMI(w_i,w_j)=\sum\frac{MI(X_i,Y_j)}{\sum MI(X_i,Y_j)}MI(X_i,Y_j)\,. \quad (2)$$

In the formula (2), X_i is a consecutive substring which includes the term w_i but without X_j, and X_j is a consecutive substring which includes the term w_j but without X_i, $MI(X_i,Y_j)$ is the mutual information of X_i and X_j.

According to the formula (2), we can calculate the generalized mutual information of each pair of opinion feature and opinion word in one review, and choose the ones that exceed the threshold as the correct appraisal expressions.

IV. EXPERIMENT

A. Dataset and Evaluation Method

To the best of our knowledge, there is not a public benchmark data for appraisal expression recognition until now, so we manually construct a small-scale test corpus in the experiment. The test corpus consists of 500 sentences that come from the reviews of MP3/4 Player and 500 sentences that come from the reviews of Portable Digital Camera in Amazon, and each sentence contains more than 10 words. We gave these 1000 sentences to two annotators respectively, and asked them to manually find out the appraisal expressions. The percentage of agreement is used to measure the inter-rater reliability. The results are up to 87% and 89%, which suggests the consistency and validity of the test corpus. We used the coincident results as the benchmark in the experiment. The following TABLE I shows the statistical information of the test corpus.

TABLE I.
THE STATISTICAL INFORMATION OF THE TEST CORPUS

Statistical information	MP3/4 Player	Portable Digital Camera
Number of sentiment sentences	500	500
Number of opinion features	38	52
Number of opinion words	74	96
Total of appraisal expressions	759	893
Number of distinct appraisal expressions	92	107

According to the statistical information in TABLE I, we can find that 1) the average frequencies of opinion features and opinion words are 759/38=17.34, 759/74=10.26 in the reviews of MP3/4 Player, and they are 893/52=17.17, 893/96=9.30 in the reviews of Portable Digital Camera. These results suggest that opinion features and opinion words are frequently appeared in customer reviews, and the frequency filtering strategy is appropriate for identification of opinion features and opinion words; 2) the average frequency of each appraisal expression in reviews of MP3/4 Player is 759/92=8.25, and it is 893/107=8.35 in reviews of Portable Digital Camera. These results are consistent with our initial assumption that opinion features and opinion words are interdependent, and mutual information statistic is suitable for appraisal expression recognition.

Evaluation of the experimental results was performed using standard Information Retrieval (IR) metrics Precision, Recall and F-score that are defined in formula (3), (4) and (5), respectively. In addition, we used the total time to evaluate the efficiency of the three methods.

$$\mathrm{Pr}ecision=\frac{TP}{TP+FP}\,. \quad (3)$$

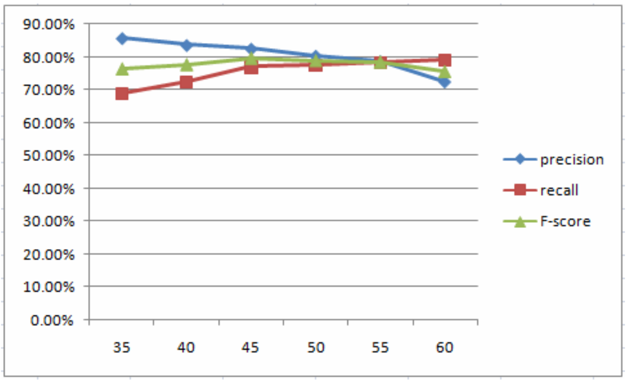
$$\mathrm{Re}call=\frac{TP}{TP+FN}\,. \quad (4)$$

$$F-score=\frac{2\times \mathrm{Pr}ecision\times \mathrm{Re}call}{\mathrm{Pr}ecision+\mathrm{Re}call}\,. \quad (5)$$

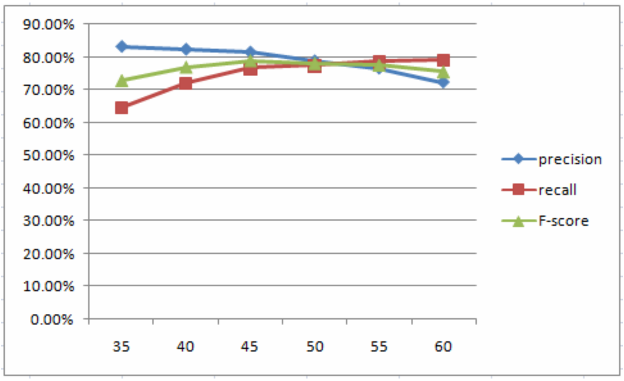
In above formulas, TP is the number of appraisal expressions that the system recognized correctly, FP is the number of appraisal expressions that recognized falsely by the system, and FN is the number of appraisal expressions which the system fails to recognize.

B. Experimental Method and Results

To evaluate the performance of the proposed method, in experiment, we implemented three methods of appraisal expression recognition: the nearest neighbor method, the method based on syntactic path and the method based on generalized mutual information. The nearest neighbor method supposes that an appraisal expression consists of an opinion feature and its nearby adjective [10]. The second method first automatically collects the syntactic paths to describe the relationship between the opinion feature and its corresponding opinion word, and then uses syntactic paths matching to recognize the appraisal expressions [23]. The proposed method also first collects the opinion features and opinion words, and then uses the generalized mutual information to predict their relationship.



(a) Experimental Result on reviews of MP3/4 Player



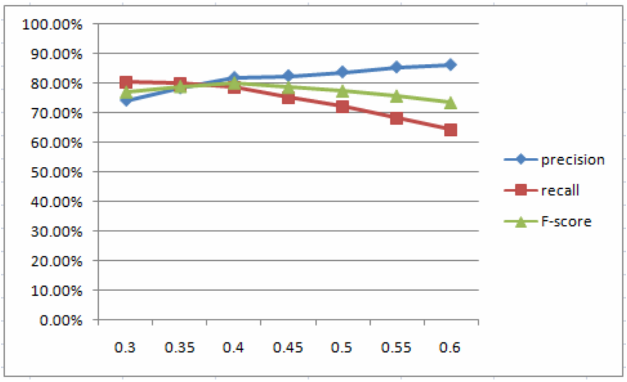
(b) Experimental Result on reviews of Portable Digital Camera

Figure 2. Experimental Results Using Different Syntactic Paths

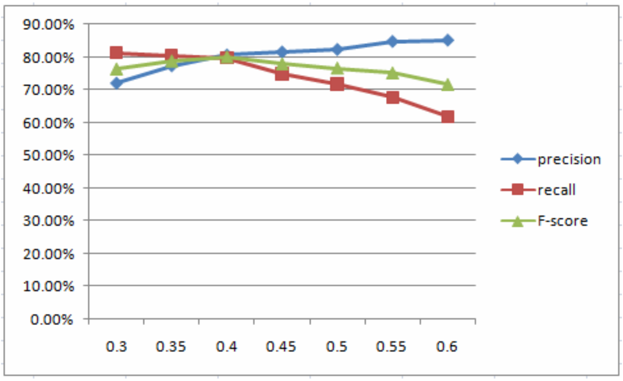
All of the three methods need to firstly identify the candidates of opinion features and opinion words, but they are different when relate the opinion word with its corresponding opinion feature. The first method relates the opinion word with its nearest opinion feature, and the second method uses the syntactic path to relate the opinion word with its opinion feature, and our method relates the opinion word with its opinion feature according to their generalized mutual information. Accordingly, the implementation of our experiment can be divided into the following five steps.

- Use the NLTK to parse each review to split the review into sentences and to produce the POS tags, and then further use chunking analysis to identify the nouns phrases.
- Lemmatize each word according its POS tag and count its frequency, and then group the words by POS tags and sort them by their frequency. After that, we choose the top 120 adjectives and adverbs as the candidate set of opinion words. For the candidate set of opinion features, we choose the nouns and nouns phrases which are in the top 100 and their *p-support* must exceed 0.5.
- Implement the nearest neighbor method. For each review in review database, identify the opinion features according to the candidate set of opinion features which produced in above step, and then choose the nearest adjective as its opinion word. The statistic information of experimental results is shown in the third row of the TABLE II.
- Implement the second method based on syntactic paths. We used the Stanford Parser [26] to parse the grammatical structure of each review sentences, and then extracted and generalized the syntactic paths between the identified opinion feature and the identified opinion word. We respectively chose the top 35, 40, 45, 50, 55, and 60 frequent syntactic paths, and used the exact path matching method to recognize the appraisal expressions. The experimental results shown in figure 2 suggest that it is best when choose the top 45 frequent syntactic paths to recognize the appraisal expressions. The fourth row of TABLE II shows the experimental results when we chose the top 45 frequent syntactic paths.
- Implement the proposed method based on generalized mutual information. For each review in review database, we first identified the opinion features and opinion words, and build the inverted index of the reviews. We only used the opinion features and opinion words as the index terms. After that, we calculated the generalized mutual information of each pair of opinion feature and opinion word. When calculate the generalized mutual information, we only considered the context of opinion words, which means the contextual window size of opinion feature was set to 0, and the context window size of opinion word was set to 3. Additionally, we respectively observed the experimental results when the threshold of generalized mutual information was set to 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, and 0.60. According to the figure 3, we can see that it is best when the threshold is set to 0.4 in our experiment. The last row of TABLE II shows the experimental results at the moment.

The TABLE II shows the comparison between our proposed method and two baselines. According to the results, we can find that: 1) the precision, recall and F-score of our method based on generalized mutual information are all higher than the nearest neighbor



(a) Experimental Result on reviews of MP3/4 Player



(b) Experimental Result on reviews of Portable Digital Camera

Figure 3. Experimental Results Using Different Threshold of Generalized Mutual Information

TABLE II.
COMPARISON BETWEEN OUR APPRAISAL EXPRESSION RECOGNITION METHOD AND TWO BASELINES

Methods	MP3/4 Player				Portable Digital Camera			
	P(%)	R(%)	F(%)	T(s)	P(%)	R(%)	F(%)	T(s)
Nearest-based	62.53	66.74	64.57	118.14	61.37	67.26	64.18	164.27
Syntactic Path-based	82.47	76.92	79.60	314.85	81.49	76.48	78.91	371.28
GMI-based(our)	81.86	78.66	**80.23**	**237.98**	80.63	79.53	**80.08**	**281.46**

method, but the efficiency of our method is lower than the nearest neighbor method. We think the reason is that the constraint of generalized mutual information is stricter than the nearest position, so it is more accurate. In addition, according to observing the recognized appraisal expressions, we found that our method can identify the appraisal expressions in which opinion word is far away from opinion features. However, our method needs to acquire the statistical information of the co-occurrence of opinion features and opinion words, thus it is less efficient than the nearest neighbor method. 2) The F-score, recall and efficiency of our method based on generalized mutual information is higher than the method based on syntactic path, but the precision of our method is slightly lower than the method based on syntactic path. We think the reason is that the syntactic path used the grammatical relationship, the constraint of which is stricter than the generalized mutual information, so the precision of the method based on syntactic path is higher than our method. However, the recall of our method is higher than the method based on syntactic path, as a result, the F-score of our method is slightly higher than the method based on syntactic path. Besides, as our method can avoid the complex syntactic analysis, so the efficiency of our method is higher than the method based on syntactic path.

In summary, according to the comparison between our method and the two baselines, we can clearly see that the proposed method is much more effective for the recognition of appraisal expressions.

V. CONCLUSION

In this paper, we analyzed the problems of the existing methods of appraisal expression recognition, and proposed a novel method based on generalized mutual information. Our experimental results indicated that the proposed method is very promising in automatically recognizing the appraisal expressions from the customers' reviews. It does not fill in any template and is domain independent. More important, it can avoid the complex syntactic analysis while keeping the comparable accuracy, which greatly improves the efficiency of appraisal expression recognition.

However, the precision of the proposed method is somewhat dependent on the identification of opinion features and opinion words. In addition, we did not consider the verbs, nouns in the process of opinion words extraction. In our future work, we will further improve and refine our method, and deal with the problems above.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (No. 61272205), and supported by the Project Fund of National Scientific and Technical Support Plan during the 11th Five-year Plan (No. 2009BAH51B03), and supported by the Social Science Fund Project, Ministry of Education of China (No. 09YJC870009).

REFERENCES

[1] ZHAO YY, QIN B, and LIU T, "Sentiment Analysis," *Journal of Software,* vol. 21, 2010, pp.1834-1848.

[2] Mingjun Xin, Hanxiang Wu and Zhihua Liu. A Quick Emergency Response Model for Micro-blog Public Opinion Crisis Based on Text Sentiment Intensity, *Journal of Software,* vol. 7, no. 6, 2012, pp. 1413-1420.

[3] Feng Zhao, Qianqiao Hu. Orientation Mining-Driven Approach to Analyze Web Public Sentiment, *Journal of Software,* vol. 6, no. 8, 2011, pp. 1417-1428.

[4] Hatzivassiloglou V, McKeown KR, "Predicting the semantic orientation of adjectives," in *Proceedings of eighth Conference of the European Chapter of the Association for Computational Linguistics (EACL-1997)*, Morristown: ACL, 1997, pp.174–181.

[5] Wiebe J, "Learning Subjective Adjectives from Corpora," in *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence (AAAI-2000)*, 2000, pp.735-740.

[6] Peter D. Turney, "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews," in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, Philadelphia: ACL, 2002, pp.417-424.

[7] Bo Pang, Lillian Lee and Shivakumar Vaithyanathan, "Thumbs up? Sentiment Classification using Machine Learning Techniques," in *Proceedings of ACL-2002 Conference on Empirical Methods in Natural Language Processing*, Morristown, NJ: Association for Computational Linguistics, 2002, pp.79- 86.

[8] LIU Taian, WANG Yunjia, WANG Yinlei and LIU Wentong. The new Development in Support Vector Machine Algorithm Theory and Its Application, *Journal of Software,* vol. 6, no. 2, 2011, pp. 273-280.

[9] Kim S M, Hovy E, "Determining the sentiment of opinions," in *Proceedings of the 20th International Conference on Computational Linguistics (COLING-2004)*. Switzerland, 2004, pp. 1367-1373.

[10] Minqing Hu, Bing Liu, "Mining and Summarizing customer reviews," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD-2004)*, Seattle, Washington, USA, Aug 22-25, 2004.

[11] Yao TF, Nie QY, Li JC, Li LL, Lou DC, Chen K, and Fu Y, "An opinion mining system for Chinese Automobile reviews," in *Proceedings of the Frontiers of Chinese Information Processing*, Beijing, 2006, pp.260−281.

[12] Liu HY, Zhao YY, Qin B, and Liu T, "Target extraction and sentiment classification," *Journal of Chinese Information Processing*, vol. 24, No. 1, 2010, pp. 84−88.

[13] Minqing Hu, Bing Liu, "Mining Opinion Features in Customer Reviews," in *Proceedings of the 19th national conference on Artificial intelligence (AAAI-2004)*, 2004, pp.755-760.

[14] Zhou J, Lin C, and Li BC, "Research on the Identification of Opinion Topic Expressed in Web Comments," *Journal of the China Society for Science and Technical Information*, vol. 29, No. 5, October 2010, pp. 858-863.

[15] Riloff E, Wiebe J, "Learning Extraction Patterns for Subjective Expressions" in *Proceedings of the 2003 conference on Empirical methods in natural language processing (EMNLP-2003)*, 2003, pp. 105–112.

[16] Turney P, Littman ML, "Measuring Praise and Criticism: Inference of Semantic Orientation from Association," *ACM Transactions on Information Systems (TOIS)*, 2003, 21(4):315–346.

[17] Kim SM, Hovy E, "Automatic Detection of Opinion Bearing Words and Sentences," in *Proceedings of the Second International Joint Conference on Natural Language Processing (IJCNLP-2005)*, 2005, pp. 61–66.

[18] Zhu YL, Min J, Zhou YQ, Huang XJ and Wu LD, "Semantic Orientation Computing Based on HowNet," *Journal of Chinese Information Processing*, vol. 20, No.1, 2006, pp. 14-20.

[19] Wilson T, Wiebe J, "Annotating attributions and private states," in *Proceedings of the ACL Workshop on Frontiers in Corpus Annotation II: Pie in the Sky*, Morristown: Association for Computational Linguistics, 2005, pp. 53−60.

[20] Bloom K, Garg N and Argamon S, "Extracting appraisal expressions," in *Proceedings of Human Language Technologies/North American Association of Computational Linguists (HLT-NAACL 2007)*, Morristown: Association for Computational Linguistics, 2007, pp. 308−315.

[21] Kobayashi N, Inui K, Matsumoto Y, Tateishi K and Fukushima T, "Collecting evaluative expressions for opinion extraction," in *Proceedings of the 1st International Joint Conference on Natural Language Processing (IJCNLP-04)*, New York: Springer-Verlag, 2004, pp. 584−589.

[22] Popescu AM and Etzioni O, "Extracting Product Features and Opinions from Reviews," in Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLTEMNLP-2005), pp. 2005.339–346.

[23] Zhao YY, Qin B, Che WX, and Liu T, "Appraisal Expression Recognition Based on Syntactic Path," *Journal of Software*, vol. 22, No.5, 2011, pp.887-898.

[24] David M. Magerman and Mitchell P. Marcus, "Parsing a natural language using mutual information statistics," in *Proceedings of the eighth National conference on Artificial intelligenc (AAAI-1990)*, 1990, pp. 984-989.

[25] Steven Bird, Ewan Klein, and Edward Loper, *Natural Language Processing with Python*, 1st ed., vol. 1. O'Reilly Media, Inc., Sebastopol, 2009.

[26] Dan Klein and Christopher D. Manning, "Accurate Unlexicalized Parsing" in *Proceedings of the 41st Meeting of the Association for Computational Linguistics*, 2003, pp. 423-430.

Huan Huang was born in July 1986. Currently he is working toward the PhD degree in National Engineering Research Center for E-Learning at Center China Normal University. His research interests include affective computing, ITS, and semantic web.

Qingtang Liu was born in February 1969, and received the Doctor of Engineering degree in communication and information system from Huazhong University of Science & Technology in 2005. Currently he is the professor and the director of the Department of Information Technology at the Central China Normal University, and his research interests include web mining, e-learning and digital right management.

Tao Huang was born in October 1969, and received the PhD degree in educational information science and technology from Central China Normal University in 2007. Currently he is the professor of National Engineering Research Center for E-Learning at Central China Normal University, and his research interests include information retrieval, semantic web and e-learning.