Chinese Volitive Words Mining*

Jian-feng Zhang, Yu Hong, Bin Ma, Jian-min Yao, Qiao-ming Zhu
School of Computer Science & Technology, Soochow University, Suzhou, China
tianxianer@gmail.com

Abstract

This paper proposes a grammar-based unsupervised method to automatically mine the Chinese volitive words, which are the important clues of intention and desiration in literal content, such as "can", "must", "rather than", etc. Besides, the paper introduces a scheme of manually tagging volitive words from large-scale Chinese blogs. And the tagged blogs are adopted as corpus to evaluate our unsupervised method in experiments. The results show a precision of 74.25% and a recall of 76.03%. Based on the above method, the paper constructs a statistical model to acquire all the volitive words with the trend of the mining, which improves the performance further.

Keywords

Opinion Mining, Volitive Words, Grammar-based, Statistical Model

1 INTRODUCTION

In the field of opinion mining, a challenge task is to acquire people's willingness from large-scale literal contents, such as "I would like to buy a cheaper camera". A possible solution is to identify the characteristics of these sentences which contain willingness. Through amounts of observations, the characteristics are the words in the sentence which can express the intention and desiration of the people. For example, in the above sentence, "would like to" suggests the potential intention of the speaker. If we can acquire these characteristic words, they can be seen as the reference to judge the people's willingness. The words similar to the "would like to" such as "can", "must", "rather than" and so on are called modal words in English, and in Chinese, we call them volitive words.

This paper focuses on exploring an unsupervised method of mining volitive words based on their grammatical features in sentences, especially that in Chinese. For example, in the sentence "I prefer to buy a cheaper camera", the grammatical feature is an ordered sequence of POS "n+v+v" (viz. "I / n + prefer / v + to / null + buy / v") where the first verb "prefer" is the candidate volitive word. And according to the three groups of experiments, we analyze the trend of the mining to constructs a statistical model, which is helpful to improve the efficiency of our method.

Besides, a frequency-based n-gram model is used to create the corpus of volitive words in large-scale blogs, which adopt the n-gram to acquire all possible volitive words that involve the specific Chinese volitive character, and constrict the scope of tagging by filtering low-frequent occurred n-grams in dataset.

* Supported by Natural Science Foundation of China (60970057, 60873105, 61003152).

The remainder of this paper is organized as follows: Section 2 reviews relevant work performed in the field. Section 3 describes the main method of mining volitive words. Section 4 introduces how to construct evaluation corpus. Section 5 proposes a statistical model based on the analysis of the experiment and results, assessing the strengths and weakness of our approach. And in Section 6, we draw our conclusion and state the future work.

2 RELATED WORK

Opinion mining and analysis have been studied by many researchers in recent years. Most researches focus on two main research directions: sentiment classification and feature-based opinion mining. Sentiment classification investigates the ways to classify texts into three classes: positive, negative, and neutral. Opinion mining emphasizes on mining the views of people but not normally their polarities.

Based on the definition of opinion mining initially proposed by Pang et al [1], much research on it has been carried out. Ding et al., [2] propose a holistic lexicon-based method to mine opinions which regards some polarity words as seeds and expands the polarity lexicon by involving their synonyms and antonyms. The expanded lexicon is helpful to obtain high recall of opinion mining. On the basis, Lee et al., [3] adopt conjunction rules and machine learning technologies to further develop the linguistic resource for sentiment classification. Esuli et al., [4] propose an subjectivity-based method of opinion mining, which well uses the bootstrapping algorithm to obtain subjective words and determines their orientation based on the probability of PMI (viz., Pointwise Mutual Information) . Conrad et al., [5] adopt the language model to identify the polarity. Additionally, in the field of application, Ghose et al., [6] adopt econometrics to realize opinion mining.

In Chinese, Yao et al., [7] summarize various key issues of Chinese opinion mining. On this basis, Song et al., [8] use some corpus tools to construct the subjective texts corpus. Li et al., [9] combine the semantics analysis with the sentiment recognition. And Liao et al., [10] utilizes the probability model to judge the sentiment of the blogs. We summarize the disadvantages of the previous methods, propose a grammar-based unsupervised method and construct a statistical model to mine Chinese volitive words.

3 MOTIVATION

The purpose of mining the volitive words is to achieve the depiction of volitive tendency mining and the content description. Especially, we want to integrate the volitive words and the event extraction to realize the dynamic evolution of the topic, then get a probabilistic model to measure topic variation, finally, fulfill the aim that to predict the ultimate form of the event or topic.

And as follows, it is the figure 1 to integrate the volitive words mining and event extraction. In this paper, we just research the volitive words mining.

4 THE GRAMMAR-BASED UNSUPERVISED METHOD

An obvious grammar characteristic of volitive words is their specific position in the ordered sequence of POS (part-of-speech), e.g., a sequence of POS is “n+v+v”, by terms of the numerous observations, where the second position of POS, viz the first verb is probably a volitive word. Based on the characteristic, we propose a grammar-based method which adopts the grammatical traits, especially the position information in ordered sequence of

POS, to mine volitive words. Take "I prefer to buy a cheaper camera" as an example. According to the result of POS tagging, "I / noun + prefer / verb + buying / verb", the first verb "prefer" is a volitive word.

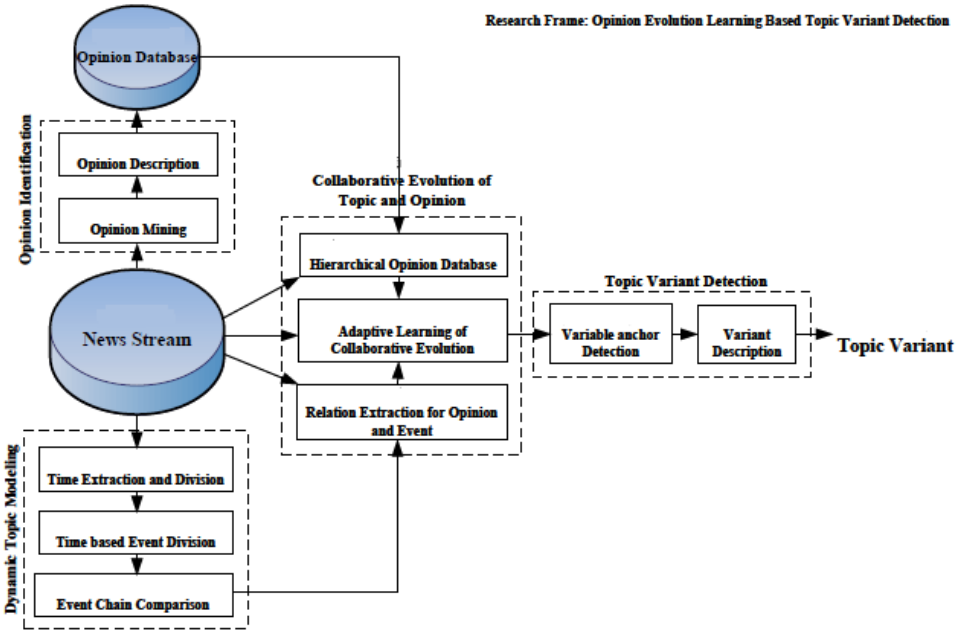

Fig.1. The integration of volitive words mining and event extraction

4.1. Local Mining

The basic treatment is to locally mine the volitive word which focuses on obtain the whole volitive words that involve the given volitive characters. In detail, including three steps as follows:

Several initial volitive words can be regarded as the input for expansion; the main purpose is to mine the volitive words as completely as possible. Take "*愿意*(will)" as an example, we segment the word into characters, i.e., *"愿", "意"*, then we can find out that other words which contains the characters like *"愿望", "有意"* are probably the new volitive words. So it can be concluded that at the initial condition of given few words and their characters, we can acquire the whole volitive words which contain the characters immensely.

Step 1: We firstly segment texts, viz. blogs from web by crawler, into sentence sets. Then use the POS tagging system of Chinese Academy of Sciences to realize word division and acquire the ordered sequence of POS in each sentence.

Step 2: Then we give several known volitive words (5 words in this paper) as seeds, and segment them into volitive characters. For example, given the seeds "*渴望*(desire)", "*必须*(must)", "*愿意*(will)", "*可以*(may)", "*应该*(should)", we segment them into the Chinese characters *"渴", "望", "必", "须", "愿", "意", "可", "以", "应", "该"*. After that we filter the sentences which don't include any seed volitive character. This step aims at improving the precision of volitive word mining by the strict restraint that the desirable sentences must involve volitive characters.

Step 3: For each sentence, we detect volitive words based on the ordered sequences of POS and corresponding positions in Table 1.

In the table, Two "/d" combined means combining the two "/d" into one word, e.g., "*不*/d *可能*/d", combine the two "/d" into "*不可能*".

In the real mining process, the overlapping words will be kept only once. All the mined words consist of a word set.

Table 1. ORDERED SEQUENCES OF POS AND POSITION

	ordered sequence of POS	position
grammar 1	/n+/v+/v	The first "/v"
grammar 2	/d+/v+/v	The first "/v"
grammar 3	/d+/v+/p	The "/v"
grammar 4	/d+/v+/r	The "/v"
grammar 5	/d+/v+/rr	The "/v"
grammar 6	/rr+/d	The "/d"
grammar 7	/rr+/v+/v	The first "/v"
grammar 8	/rr+/v+/p	The "/v"
grammar 9	/d+/d	Two "/d" combined
grammar 10	/d+/v	Two "/d" combined

4.2. Global Mining

Even though the words which contain the given volitive characters (mentioned in section 3.1) can be mined completely, numerous other volitive words that don't involve the characters cannot be obtained by the local mining for only one time. To solve the problem, we propose a global mining method which iteratively runs the local mining procedure under given different groups of volitive characters as input.

But the key issue for the global mining is how many times should be used to iteratively run the local mining? Actually, we find a hyperbola trend of the number of unduplicated volitive words obtained by the iterative running in a small-scale label corpus (extracted from Synonymy Thesaurus developed by HIT) [2]. As shown by the series of gray histograms above the horizontal axis in the Figure 1, which illustrates the occupied proportions of the locally mined volitive words decrease by degrees along with the iterative times *n*. On the basis, we use a unilateral hyperbola-based function to roughly estimate the iterative times *n* as the equation (1).

$$(n+\alpha)\cdot(y+4\cdot\alpha-160)=4\cdot(\alpha+1)\cdot(\alpha+2) \quad (1)$$

where *n* denotes the iterative times, *y* is the remainder of volitive words that have not been acquired in the labeled corpus (230 words totally at the beginning). And *a* is a parameter, its optimal value, i.e., 900, is acquired after the numerous training.

From the function, it can be inferred that when *n* is 45, the value of *y* approaches illimitably to 0. It could be concluded that at least 45 iterative times are needed to acquire all the labeled volitive words. In other words, we have to choose 5 unduplicated words (viz., 10 different volitive characters) each time as the initial seeds, and run the local mining for 45 times.

In fact, we globally mine the volitive words from 35,000 blog texts. Thus, the corpus of HIT is only a small portion of total Chinese volitive words. So the hyperbola trend is only a partial phenomenon. But because the corpus of HIT is an approximatively random portion, we can roughly think that the iterative times for whole volitive words mining also obey the hyperbola trend. It just likes we divide the Chinese volitive words into two parts, one is the previously known volitive words (viz., corpus of HIT), the other is the unknown ones, then

we use the laws on the former to process the later, as shown in the Figure 2, where the series of blank histograms under the horizontal axis denote the similar trend with that achieved on the corpus of HIT.

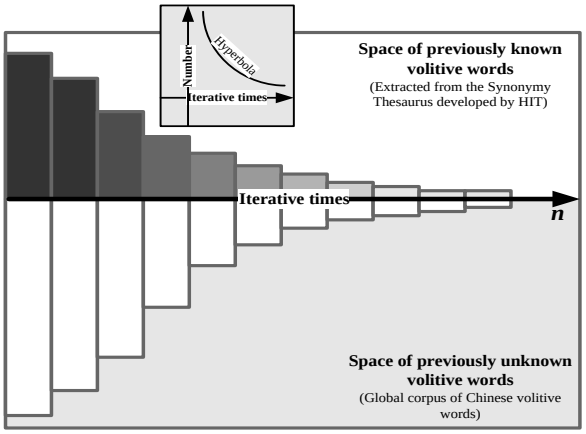

Fig. 2. The hyperbola trend of the iteratively mining

4.3. Grammar ranking in global mining

But it is conceivable that the global mining will result in low precision because of the cumulative errors from each iterative local mining. In fact, the procedure of local mining focuses on acquiring the approximatively whole volitive words that involve the initially given volitive characters. So it is inevitable that the local mining will involve many noises. In other words, the local mining dotes on the recall but ignores the precision. Thus the persistently iterative running of local mining will result in unaccepted poor precision of global mining. To solve the problem, we use a grammar-based ranking of candidate volitive words to filter the noises.

The grammar-based ranking procedure includes three components as follows:

- Acquiring the precision of the grammar

In this component, the key issue is to acquire the precision of each grammar-based determination, To solve that, we iteratively run our local mining for three times and verify the probability of correct volitive words in the whole results determined by a specific grammar. Thus we can roughly achieve the distributions of precision among different grammars (viz., the grammars in Table 1).

- Acquiring the professional score of the grammar

In this component, we focus on giving a score to describe the professional abilities of grammars in determining volitive words. That is, given a specific local mining, the grammar, which achieves the most quantity of correct volitive words, has the most professional mining ability. Based on this assumption, we calculate the scores for all of the grammars based on their occupied rates in the correctly determined volitive words. Here, the results of the three local mining (mentioned in the first component) are also used.

- Ranking the candidate volitive words

In this component, for a grammar, we calculate its ranking score as follows:

$$Score_{rnk} = P \cdot Score_{pro} \tag{2}$$

where $core_{rnk}$ denotes the ranking score of a specific grammar; P denotes the prior-precision of the grammar; $Score_{pro}$ denotes the professional score of the grammar. Then, we run our global mining procedure, and for all of its output candidate words, we rank them based on the $Score_{rnk}$ of their corresponding grammars.

On the basis, we filter the words which refer to the grammars having lower $Score_{rnk}$ than a threshold. In this paper, we simply make the threshold equal to the average value of the $Score_{rnk}$ in the training procedure (viz., the scores achieved in the three iterative local mining as mention in the first component).

5 CORPUS ESTABLISHMENT

Until now, there is not proprietary corpus to evaluate the performance of Chinese volitive word mining. In this paper, we establish a corpus by using a grammar-frequency based labeling method, and use it to evaluate the performance of our local mining method. Especially, we use a long-tail detection method to evaluate the global mining method.

- Grammar-frequency based corpus establishment

For a specific local mining, the corresponding grammar-frequency based corpus establishment includes seven steps as follows:

Step 1: Given the initial volitive characters of the local mining, which are the input of the mining procedure as mentioned in section 3.1, we extract all sentences which involve the characters from the source corpus, viz., the 35,000 blog texts, as the basic corpus.

Step 2: Given the basic corpus, we acquire their grammar-based descriptions. In detail, every sentence in the corpus will be segmented into bi-grams, tri-grams, quad-grams. Thus we can obtain three kind of gram-datasets which are respectively named as Gset_1, Gset_2 and Gset_3.

Step 3: Given any gram-datasets, we calculate the frequencies of the grams and filter out the sentences which involve the low-frequency grams. In other words, only the high-frequency grams are regarded as the probable volitive words. The main reason is if a grammar has a high frequency, it is probably a word then a volitive word.

Step 4: The remaining grams are manually labeled by six volunteers in department of Chinese language who additionally perform the cross-checking. We regard the labeled grams and all sentences in the basic corpus (viz., the corpus mentioned in the step 2) as a test sub-corpus.

We perform the *Step 3* and *Step 4* for each gram-dataset, viz., Gset_1, Gset_2 and Gset_3, to generate their test sub-corpus. At last, all of the sub-corpus are combined to generate our final test corpus.

- Long-tail detection based evaluation method

A key difficulty in our experiment is to evaluate the performance of the global mining. In fact, the quantity of grams that need to be labeled for a local mining is few. But, it is difficult to manually label all Chinese volitive words in all 35,000 blog texts for evaluating the global mining. Its possible evidence is that at least 45 times are needed to iteratively run the local mining to simultaneously acquire enough volitive words.

To solve the problem, we propose a long-tail detection based evaluation method. The method focuses on verifying the quantities of volitive words correctly determined by the later-period local mining. In detail, it regards the times of local mining, which is calculated to be 45 by the hyperbola-based function in section 3.2, as the center of a detection window, and verify the quantity of the volitive words correctly mined by each-time local mining in the window. If the quantities are always few and nearly the same, we can determine that the

global mining simultaneously converges and obtains the most of Chinese volitive words. In our experiments, the radius of the window is appointed to be 5 times.

6 EXPERIMENTS

6.1. Experiment Design

In the experiments, the gram-frequency based test corpus is in use which is generated from the 35, 000 blog texts of the websites Sina, Sohu, Netease, Tencent et al. And the precision, recall and F-score are used to evaluate our mining method.

In the pretreatment, we use the POS tagging system developed by Chinese Academy of Sciences to acquire the ordered sequence of POS. Additionally; the Synonymy Thesaurus developed by HIT (Harbin Institute of Technology) is adopted to collect initially known volitive words whose total number is 230 according to the synonym and antonym rule proposed by Ding (2008). In the global mining, only 5 random known volitive words and their characters are adopted as the seeds to run the first local mining. And in the process, the obtained words by a local mining are used as the new seeds to run the next local mining iteratively.

The experiments give three main results: first, we give the performances of the local mining method which runs on three different beginnings (viz., given different 5 known volitive words as seeds); second, we give the precision of global mining by iteratively running local mining for 45 times (viz., iterative times n equaling to 45 as mentioned in section 3.2) under randomly given 5 known volitive words; third, we give the recall of later-period local mining in the window of the long-tail detection whose radius is appointed to be 5.

6.2. Performances of local mining

In this test, we do three groups of experiments totally. Each group begins at 5 different volitive words and iteratively run the local mining for three times. We simply name it as 3×3 test (viz., 3 round iterative running for each of the three groups). In the second and third group, we choose different 5 seeds. Finally, we can acquire three word sets mined by the grammar-based unsupervised method.

The precision, recall, F-score and the number of volitive words acquired by the first group of local mining are shown in the Table 2 and Table 3.

Table 2 THE RESULTS OF FIRST-GOURP LOCAL MINING

	Round 1	*Round 2*	*Round 3*	*total*
Precision(%)	85.6	80.2	77.6	76.0
Recall(%)	73.4	75.5	71.1	72.2
F(%)	79.0	77.8	74.2	74.1

Table 3 NUMBER OF VOLITIVE WORDS ACQUIRED (FIRST-G)

	Round1	Round2	Round3	total
Manual	70	68	72	151
Grammar-based	60	64	66	159

According to the results in the tables above, we can find that the precision is similar among the three rounds of iterative running, and the recall and F-score are in the same way. It illustrates that each iterative running doesn't result in much loss of performance. Thus the global mining, which needs iteratively run local mining for 45 times, will not go downhill

quickly. Besides, all precisions are high, but recall are relative low. It seems that the local mining are adept in accurately identifying volitive words from noisy language resources, but relatively weak in acquiring the words completely. Thus the global mining will not reduce the precision when it focuses on improving the recall by iteratively running local mining.

To verify whether the performances above is an isolated phenomenon, we additionally run another two groups of local mining when given different random volitive words as initial inputs. Their performances are shown in Table 4~7 respectively.

Table 4 THE RESULTS OF SECOND-GROUP LOCAL MINING

	Round 1	*Round 2*	*Round 3*	*total*
Precision(%)	82.3	79.2	75.3	78.0
Recall(%)	72.7	84.2	88.9	80.9
F(%)	77.2	81.6	81.5	79.4

Table 5 NUMBER OF VOLITIVE WORDS ACQUIRED (SECOND-G)

	Round1	*Round2*	*Round3*	*total*
Manual	77	95	72	136
Grammar-based	68	101	85	141

Table 6 THE RESULTS OF THIRD-GROUP LOCAL MINING

	Round 1	*Round 2*	*Round 3*	*total*
Precision(%)	77.6	76.8	73.1	68.7
Recall(%)	85.7	88.7	91.9	84.6
F(%)	81.4	82.3	81.4	75.8

Table 7 NUMBER OF VOLITIVE WORDS ACQUIRED (THIRD-G)

	Round1	*Round2*	*Round3*	*total*
Manual	77	71	78	91
Grammar-based	94	82	62	112

By the Table 4~5, we in surprise find that the recall in the second group of test continuously increase along with iterative local mining, and they even exceed the precision after the second-round mining. That is because more volitive characters are used as the initial seeds to perform the local mining, although they still extracted from only 5 different volitive words as the same as that in the first group of test. It should be mentioned that we cannot avoid the repetitive characters because of the random rule of selecting initial seeds, such as the random seeds "*愿意*" (viz., "*be willing to*" in English) and "*愿望*" (viz., "*wish*" in English) having the same character "*愿*". And the same phenomenon also occurs in the third group of test, and we find its initial volitive characters are even more than that in second group. So it can be concluded that when more different volitive characters (>8) are used as initial seeds, the local mining can achieve higher recall. Although the increasing trend of recall is different from the reducing one in the first group of test, it is helpful for global mining. So whatever initial seeds are selected, they wouldn't seriously hurt the performance of global mining.

6.3. Performance of global mining

As mentioned above, the global mining can rely on the n-times iterative local mining to obtain high recall of all Chinese volitive words. But if the precision of each local mining is low, the global mining will involve a large quantity of noises even if it has high recall. So it additionally requests the high precision of each iterative local mining. But we find the

precision actually continuously decreases along with the iterative running. As shown in Fig. 1, the precision of the three-group local mining tests all decrease. To solve the problem, we adopt the grammar-based ranking in global mining as mentioned in section 4.3. Here, we firstly give the precision distribution of the grammars on determining Chinese volitive words, and then we give the performance of global mining that involves the grammar-based ranking treatment.

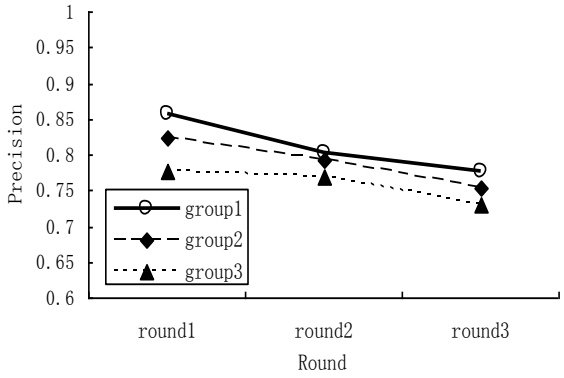

Fig. 3. The tendency of precision

- Precision distribution of grammar-based ranking

The ranking method needs to calculate two factors for each grammar: precision and professional score as mentioned in section 3.3. According to the results of the three-group local mining test, the average values of the factors are shown in Table 8, where the occupied rate of correctly determining volitive words is used as professional score, in other words, the grammar which can correctly determines more volitive words will be more professional. Based on the data in Table 8, we can calculate the ranking scores for the grammars, viz., the $Score_{rnk}$ in Equation (2), and the average ranking score, which is used as the threshold to filter out noises, equals to 0.08.

- Performance of global mining

Given the times of iterative local mining (viz., n=45) and the threshold (=0.08), we run our global mining and filter out the candidate volitive words whose grammar-based determination has the lower $Score_{rnk}$ than the threshold. The performance of global mining is shown in Table 9, where the symbol "Number_CD" denotes the number of volitive words being correctly determined by global mining, "Number_OP" is the number output by the mining procedure. Besides the total precision, we also give the precision of the grammar whose $Score_{rnk}$ are higher than the threshold.

The results show that the global mining achieves 0.67 precision which is mainly negatively influenced by the low precision of the grammar "/d+/d". But it seems unreasonable to filter out the volitive words mined by the grammar, otherwise global mining will loss a large part of correct words, as shown that the grammar "/d+/d" recall 335 correct volitive words which is nearly 15% of total correct results. Although we can ignore the grammar to achieve high precision and maintain high recall by improve the times of iterative local mining, the efficiency of global mining will be reduced. Besides most grammars have the similar professional score, viz., the occupied rate, with that in the three-group local mining. So we can conclude that the grammars are relatively robust.

Table 8 LOCAL-MINING TEST RESULTS

grammar	sequence of POS	amount	Occupied rates	Precision
grammar 1	/n+/v+/v	52	20.6%	0.85
grammar 2	/d+/v+/v	43	17.1%	0.79
grammar 3	/d+/v+/p	18	7.1%	0.33
grammar 4	/d+/v+/r	26	10.3%	0.69
grammar 5	/d+/v+/rr	3	1.2%	0.61
grammar 6	/rr+/d	24	9.5%	0.55
grammar 7	/rr+/d	21	8.3%	0.56
grammar 8	/rr+/v+/p	16	6.3%	0.45
grammar 9	/d+/d	38	15.1%	0.65
grammar 10	/d+/v	11	4.4%	0.30
total number		252		

Table 9 THE VOLITIVE-WORD NUMBER AND PRECISION ACQUIRED BY EACH POS SEQUENCE IN GLOBAL MINING

Rank	ordered sequence of POS	amount	precision
1	/d+/d	780	0.43
2	/n+/v+/v	643	0.80
3	/d+/v+/v	507	0.78
4	/d+/v+/r	301	0.68
5	/rr+/d	234	
6	/rr+/v+/v	202	
7	/d+/v+/p	183	
8	/rr+/v+/p	61	
9	/d+/v	32	
10	/d+/v+/rr	5	
Number_CD		2188	0.67
Number_OP		2948	

Additionally, we run the global mining without using grammar-based ranking, and compare its performance with that using the ranking. The comparison of results is shown in Table 10. It can be found that the global mining without ranking acquires more 2678 words, but its precision is very poor. Compared to the quantity of noises involved by the low-$Score_{rnk}$ grammars, that of volitive words correctly determined is so few. Thus if we attempt to solve the long-tail issue of global mining by increase the times of iteratively running local mining, it is inevitable that the low-$Score_{rnk}$ grammars will always provide noises but not correct volitive words. Such as, we iteratively run local mining for 60 times by which only 2379 correct words can be acquired, viz., additional 76 correct volitive words are recalled, but the total number of words output has reach 6959, viz., additional 1257 noises are recalled.

Table 10 TEST RESULTS (RANKING VS WITHOUT RANKING)

	Using ranking	Without using ranking
Precision	0.67	0.41
Number_CD	2188	2303
Number_OP	2948	5626

6.4. Long-tail detection

As mentioned above, it is difficult to labeled all Chinese volitive words from the 35.000 blog texts (In fact, the texts normally don't involve all Chinese volitive words). So we cannot give the comprehensive recall of global mining. Thus we propose the long-tail detection evaluation method, as mentioned in section 4, which focus on verify the quantity

trend of volitive words correctly mined by the iterative running in the later period of global mining. The Fig. 4 shows the long-tail detection results when given the window which uses the 45th iterative local mining as central and 5 neighbors as radius. It should be mentioned that the maximum quantity of correct volitive words mined by local mining is known to be 95, as shown in the Table 5, so the maximum value of y-axis in Fig.4 is simultaneously appointed to be 100.

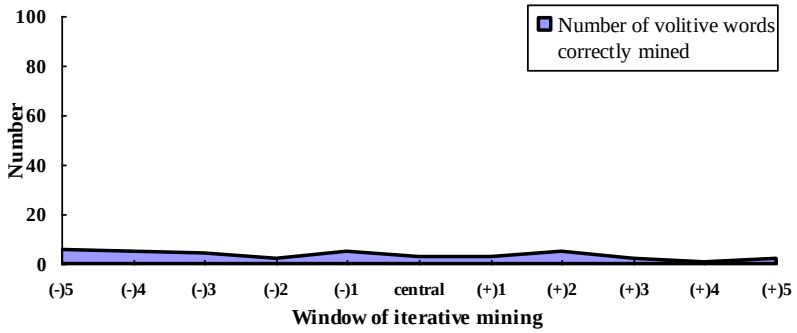

Fig. 4. Long-tail correctly mined in the detection window

By the results shown in the Fig. 4, we can find that all quantities of volitive words returned by the iterative local mining in the window (40~50 times) are very few, whose maximum value is 6, and most of them simultaneously equal to 2. It illustrates that the global mining have approached to the long-tail, in other words, most volitive words that it can mined have be returned. Additionally, as mentioned in section 5.2, we have illustrated that the selection of initial seeds will not affect the performance of global mining seriously, in other words, the recall of different global mining will be similar whatever the beginning is. So it can be concluded that the global mining simultaneously acquire the most of current Chinese volitive words.

In fact, it is difficult to give the quantity of the long tail and the real end of iterative mining. So it will not be known how many novel volitive words are still hiding in the long-tail. Thus, to acquire higher-recall of the words, we have to find a way to efficiently identify the volitive words in the long-tail. A method is to run multiple global mining under given different initial seeds, and at the same time detect their real-time unduplicated volitive words in the early period of iterative running. Thus we don't need expand the times of iterative local mining for each global mining, and acquire more correct volitive words when the precision of mining is still high. This will be our new attempt in future work.

7 FUTURE WORK

In the future, we will make use of our statistical model to further mine volitive words in larger-scale corpus, then utilize the mined volitive words to establish an automatic classified hierarchy based on the strength of the words. Each volitive word has different strength to express the desiration. For example, "We could…", "We should…", "We must…" have different meanings to convey the wish of the speaker. So we want to establish the structure as follows:

Normal
Level 1
Sentiment Degree: very low (score->1)
Meaning: possible
Database: 能(can), 能够(can), 会(could)...
Example (Sentence):
能(can): We can prevent the earthquake.
...
Level 2
Sentiment Degree: low (score->2)
Meaning: willing
Database: 愿意(willing), 想(want)...
Example (Sentence):
愿意(willing): We are willing to prevent the earthquake.
...
Level 3
Sentiment Degree: middle (score->3)
Meaning: evaluation
Database: 值得(worth)...
Example (Sentence):
值得(worth): We are worth preventing the earthquake.
...
Level 4
Sentiment Degree: high (score->4)
Meaning: necessary
Database: 应该(should)...
Example (Sentence):
应该(should): We should prevent the earthquake.
...
Level 5
Sentiment Degree: very high (score->5)
Meaning: must
Database: 必须(must)...
Example (Sentence):
必须(must): We must prevent the earthquake.
...

Fig. 5. The hierarchy based on the strength of the volitive words

Then we want to mine the contents of the volitive words, combined with the semantic relations to form the semantic matching. The contents of the volitive words and semantic relation are depicted as the following figure 6:

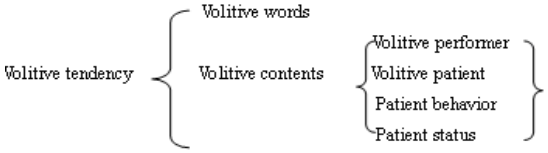

Fig. 6. The contents of the volitive words and semantic relation

For example, “Citizens are willing to see city hall took out valid earthquake prevention and control measures.” The depiction of the sentence is:

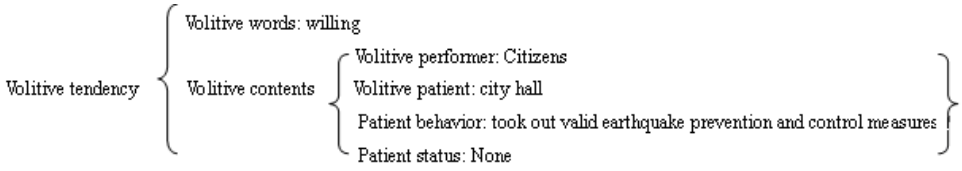

Fig.7. The depiction of an example

8 CONCLUSION

In this paper, we propose a new concept called volitive words, introduce the method of constructing the evaluation corpus for comparison and how to mine volitive words. And by means of experiments, we prove that our approach has a good performance, but the results also show that with the running of experiments, the precision of the approach becomes lower, the recall gets higher, and the F value is relatively stable. However, the mined words are more, the proportion of volitive words are lower. So we can conclude that if the mining is going on and on, all the volitive words can be mined completely.

9 REFERENCE

Pang B. and L. Lee. Opinion Mining and Sentiment Analysis. Now Publishers Inc, July 2008.

Ding. X, B. Liu, and P. S. Yu. A holistic lexicon-based approach to opinion mining. Proceedings of the Conference on Web Search and Web Data Mining. 2008.

Lee, D., Jeong, O.-R., & Lee, S.-G. Opinion mining of customer feedback data on the Web. Proceedings of the 2nd international conference on Ubiquitous information management and communication. 2008.

Esuli Andrea and Fabrizio Sebastiani. Determining term subjectivity and term orientation for opinion mining. In Proceedings of EACL-06, 11th Conference of the European Chapter of the Association for Computational Linguistics, Trento, IT. Forthcoming. 2006.

Conrad Jack G. and Frank Schilder. Opinion mining in legal blogs. In Proceedings of the 11th International Conference on Artificial Intelligence and Law (ICAIL07), pages 231-236, Palo Alto, CA. ACM Press. 2007.

Ghose A., P. G. Ipeirotis, and A. Sundararajan. Opinion mining using econometrics: A case study on reputation systems. Proceedings of the Association for Computational Linguistics. 2007.

Tianfang Yao, Xiwen Cheng, Feiyu Xu, Hans Uszkoreit, Rui Wang. The overview of the text opinion mining[J]. Journal of Chinese Information Processing, 2008, 22(5): 71-80.

Hongyan Song, Jun liu, Tianfang Yao, Quansheng Liu, Gaohui Huang. The labeled corpus construction of Chinese subjectivity texts[J]. Journal of Chinese Information Processing, 2009, 23(2): 123-128.

Dun Li, Baojun Qiao, Yuanda Cao, Yueliang Wan. The words tendency recognition based on the semantics analysis[J]. Pattern Recognition and Artificial Intelligence, 2008, 21(4): 482-487.

Xiangwen Liao, Donglin Cao, Binxing Fang, Hongbo Xu, Xueqi Cheng. The searching of the blog orientation based on the probabilistic inferential model[J]. Journal of Computer Research and Development, 2009, 46(9): 1530-1536.