Effectiveness Evaluation for Paraphrase Method of Chinese Double Negation Based on Templates

Zhongjian Wang

Harbin University of Commerce, Harbin 150028, China Harbin Far East Institute of Technology, Harbin 150025, China

wangzhj@hrbcu.edu.cn

Abstract

By demand of expression, a double negation sentences are used in many situations. According to characteristics of double negation sentences, a paraphrase method based on templates is used. The method analysis sentence structure of double negation sentences, extract syntax frame and construct paraphrase templates. The syntax structure of double negative sentence and affirmation sentence are suitable for paraphrase. By comparing the syntax structure of double negative sentences with their affirmation sentences, paraphrase templates are abstracted. Matching of paraphrase templates with paraphrase sentences is completed by calculating their similarity. To evaluate performance of the method, small-scale experiments have been done and got the coverage rate of templates and precision of paraphrase 98.9% and 84.7% respectively. The experimental results indicate the availability of the proposed method, and a small corpus resource is established that it can be provided for further research of sentence paraphrase.

Keywords: Paraphrase, Double Negation, Syntax Structure, Chinese Sentence, Paraphrase Template

1. Introduction

With the development and maturity of the underlying technology (such as participle, speech marking, syntactic analysis) of natural language processing, the research for deeper natural language processing technology, such as paraphrasing, has got a widespread concern by many researchers. Paraphrasing is a very important research topic in research of NLP. It is able to use in a lot of fields such as machine translation [1-2], information retrieval [3], question answer [4], text reading assistance [5] and other fields. Its use can improve the processing performance of the system.

The research of paraphrase is an important research field of NLP. The research of NLP has got a widespread concern by many researchers, Such as NLP technology is used in the protection of privacy in E-communication [6], the research of word sense disambiguation [7] uses Information Gain to calculate the weight of different position's context to construct the feature vectors, and get an improved Bayesian model. Especially the application of sentences paraphrase can be improved the performance of those systems.

Paraphrase is an alternative expression of source sentence without changing its meaning. Application of paraphrase technology can effectively optimize the performance of various fields, it is very significant to research sentence paraphrase.

The research of Chinese sentence paraphrase is not abundant compared with the research of English and Japanese paraphrase. The research of English paraphrase and Japanese paraphrase has achieved great progress, and a lot of research results have been applied to actual processing systems, for instance, paper [8] use paraphrase technology to generate object language sentences in machine translation system. Paper [9] use paraphrase technology to extend queries and answer of the question answer system for improving system performance, increasing recall ratio and precision.

About research of Chinese sentence paraphrase, Zong *et al.*, proposed a method of spoken Chinese paraphrasing, extracting feature of input utterance by analyzing phrase structure and chunk dependency of an utterance, finally got paraphrasing of input based on the analysis result [10]. Zhao *et al.*, proposed a method that acquires context-specific lexical paraphrases automatically, the method obtains paraphrases of a word depend on the specific sentence that the word occurs in, extracts candidate paraphrase by web mining [11]. Li *et al.*, proposed a template representation and generalization method, it represents paraphrasing template with semantic code of words and extracts slot word by using a dependency parser, using search engine to generalize paraphrasing examples[12].

Although the paraphrase research of English and Japanese has achieved many results, the research of Chinese sentence paraphrase can not use the same technique that process English or Japanese. Chinese language is neither like English which characterizes by tense, person, declension, *etc.*, nor like Japanese which characterizes by case-auxiliary word and declension. Sentence structure and empty words play a key role in Chinese expression. Due to the lack of a strict conformation variety, the study on Chinese paraphrase is relatively difficult. The means of paraphrasing English and Japanese sentences cannot be directly adopted in Chinese paraphrase.

This paper use template-based method to study on the paraphrase of Chinese double negative sentence. At first, the paper introduces Chinese double negative sentences. Then paraphrase templates are extracted by analyzing sentence structure, comparing characteristic of original sentences with paraphrase object sentences. The matching of paraphrase template with original sentences is by calculating similarity that contains POS information of sentences. The generation of object sentences is completed by matched templates. The last, small scale corpus is used to evaluate experiment and the method performance and lack is given by discussing.

2. Double Negative Sentences

A double negative sentence, comparing with normal negative sentence, it has two forms of negation that is used in the same sentence. In double negative sentences, double negatives cancel one another and produce an affirmative sense. Mandarin Chinese also employs litotes in a like manner. For example, '我不是不喜欢。' (I do not dislike.) can express by '我喜欢。' (I certainly like.).

Some common construction are '不能不,不得不,不会不' (Pinyin: bu neng bu, bu de bu, bu hui bu, 'cannot but', 'cannot not', 'will do'), which is used to express a necessity more regretful and polite than that expressed by '必须' (bixu, 'must'). Compared with '我必须说。' (wo bi xu shuo, 'I must say'), '我不得不说。' (wo bu de bu shuo, 'I cannot, but I must say') tries to emphasize that the situation is out of the speaker's hands and that the speaker has no choice in the matter: 'Unfortunately, I have to say'. Similarly, '没有人不佩服他。' (mei you ren bu pei fu ta, 'There is no person who does not admire he') is a more emphatic way to express 'Everyone admire he'.

Double negatives sentences nearly always express a positive meaning even in colloquial speech but illogically, triple negatives can as well. For instance '他不相信没人没来。' (ta bu xiang xin mei ren mei lai, 'He does not believe no one will not come') should mean 'He believes some people will not come' but is more often

understood to mean 'He thinks everyone will come'. However, triple and multiple negatives are considered obscure and are typically avoided. Some times, a double negative sentence maybe expresses an ambiguity meaning, but this type sentence is suitable to research of paraphrase.

The relation of the two negative words or negatives meaning in double negative sentences are not parataxis, but are cancel out each other, reinforce each other or weaken each other. According to these relations, a double negative sentence is divided into three types: affirmative double negative sentence; emphatic double negative sentence; euphemistic double negative sentence [13].

A double negative sentence has three expression functions. It can express affirmation, emphasis and euphemism.

2.1. Classification of Double Negative Sentences

For the three kinds of double negative sentences, there are three structures of POS (part of speak). The three structures show as follows.

- Double negation consists of two negative adverbs of using together.
- Double negation consists of a negative adverb and a verb that has negative meaning.
- Double negation consists of a negative adverb (or a verb that has negative meaning) and tone of one's voice when asking a rhetorical question.

No.	Double Negation	Examples
1	不得不	这部电影使我感动得不得不流泪。
2	不能不	他不能不喜欢吃苹果。
3	不会不	他不会不同情我的。
4	不是不	他不是不想这样做。
5	不该不	你不该不去图书馆。

Table 1. Double Negative Sentences by Two '不'

Table 2. Double Negative Sentences b	y a	'不'	and a	Word	of I	Negative	Meaning

No.	Double Negation	Examples
1	不无	他的话不无道理。
2	无不	所有人无不为他的精彩表演喝彩。
3	未必不	我未必不是你的对手。
4	难道不	这难道不是他干的吗?
5	不是没	开会前不是没通知。

Commonly, a double negative sentence consists of two ' π ' or a ' π ' plus asked word. We classify those double negative sentences to three kinds. The first, a double negative sentence consists of two negative 'no', show in Table 1. The second, a double negative sentence consists of a negative word 'no' and a word with negative meaning, show in Table 2. The third, a double negative sentence consists of a ' π ' and a word with negative meaning, but the ' π ' and the word with negative meaning are separate, we call this kind of double negative sentence as separation double negative sentence, show in Table 3.

Following Table 1, 2, 3 are familiar double negation component and example sentences.

Table 2 shows double negative sentences that consist of a ' π ' and a word of negative meaning. Table 3 shows a double negative sentence consists of a ' π ' and a word with negative meaning, but the ' π ' and the word with negative meaning are separate.

No.	Double Negation	Examples
1	不是不	他不是一个不听话的孩子。
2	非不	这件棘手的事情非他不能解决。
3	没有不	没有一个人不佩服他。
4	不会不	这么用功学习不会考的不好吧。
5	非得不	要拿好成绩,非得刻苦学习不可。

Table 3. Separation Double Negative Sentences

2.2. Semantic Analysis of Double Negative Sentences

No.	Status	Analysis results of double negative sentences				
1	Original	我/rr 不得不/d 承认/v 你/rr 很/d 优秀/a。/wj				
	Paraphrased	我/rr 承认/v 你/rr 很/d 优秀/a。/wj				
		我/rr 应该/v 承认/v 你/rr 很/d 优秀/a。/wj				
		我/rr 必须/v 承认/v 你/rr 很/d 优秀/a。/wj				
	Structure compare	(rr/nr/r)+不得不/d+承认/v+(rr/nr/r)+(*)				
		(rr/nr/r)+&+v+(rr/nr/r)+(*)				
		(rr/nr/r)+ 应该/v+v+(rr/nr/r)+(*)				
		(rr/nr/r)+必须/v+v+(rr/nr/r)+(*)				
2	Original	我们/rr 不/d 能/v 否认/v 老一辈/n 的/uj 功劳/n。/wj				
	Paraphrased	我们/rr 承认/v 老一辈/n 的/uj 功劳/n。/wj				
		我们/rr 应该/v 承认/v 老一辈/n 的/uj 功劳/n。/wj				
		我们/rr 必须/v 承认/v 老一辈/n 的/uj 功劳/n。/wj				
	Structure compare	(rr/nr/r)+不能否认+(rr/nr/r/n)+(*)				
		(rr/nr/r)+承认+(rr/nr/r/n)+(*)				
		(rr/nr/r)+要承认+(rr/nr/r/n)+(*)				
		(rr/nr/r)+要接受+(rr/nr/r/n)+(*)				
		(rr/nr/r)+应该承认+(rr/nr/r/n)+(*)				
		(rr/nr/r)+必须承认+(rr/nr/r/n)+(*)				
3	Original	他/rr 这么/d 做/v,/w 我/rr 怎么/dy 能/v 不/d 感动/v 呢/y?/w				
	Paraphrased	他/rr 这么/d 做/v, /w 我/rr 非常/d 感动/v。/wj				
		他/rr 这么/d 做/v, /w 我/rr 很/d 感动/v。/wj				
		他/rr 这么/d 做/v, /w 我/rr 特别/d 感动/v。/wj				
	Structure compare	(*),(rr/nr/n)+非常/d 感动/v。				
		(*),(rr/nr/n)+很/d 感动/v。				
		(*),(rr/nr/n)+特别/d 感动/v。				
		(*),(rr/nr/n)+真的/a 很/d 感动/v。				

Table 4. Semantic Analysis and Extraction of Templates

We use the word segmentation and part-of-speech tagging tool to analyze those double negative sentences [14], and compare original sentences with paraphrase sentences to extract paraphrasing templates. The analysis results of double negative sentences are as following Table 4.

In the Table 4, '*' is words that does not effect sentence meaning and can be exchanged by other words. In addition rr, nr, w, v, d, dy and n are POS of words of corresponding position.

2.3. Paraphrase Templates of Double Negative Sentences

The extracted paraphrasing templates are shown in Table 5.

Table 5 shows double negative paraphrase templates that are generally used. Sometime a double negative sentence can be paraphrased to several affirmation sentences. For example, the paraphrase sentence of '我不得不去图书馆看书。' (I have to go to library to read book.) can be expressed by '我去图书馆看书。', '我必须去图书馆看书。', '我应该去图书馆看书。' and so on. The paraphrase templates can complete those paraphrasing process.

Key words	Paraphrase key words	Paraphrase templates		
不得不	@,必须,应该	If (rr/nr/r)+不得不+(rr/nr/r)+(*),		
		then (rr/nr/r)+X+(rr/nr/r)+(*) and X={@, 必须, 应该}		
不能不	@,要,一定,必	If (rr/nr/r)+不能不+(rr/nr/r)+(*),		
	须,	then $(rr/nr/r)+X+(rr/nr/r)+(*)$ and X={能}		
不会不	@,会,确实,应该	If (rr/nr/r)+不会不+(rr/nr/r)+(*),		
		then $(rr/nr/r)+X+(rr/nr/r)+(*)$ and $X=\{ \Delta_{\overline{x}} \}$		
不可不	肯定,可能,一定,	If (rr/nr/r)+ 不可不+(rr/nr/r)+(*),		
	会	then $(rr/nr/r)+X+(rr/nr/r)+(*)$ and $X=\{\overrightarrow{II}\}$		
不更不	更 该 旦	If (rr/nr/r)+ 不要不+(rr/nr/r)+(*),		
小女小	安, 以, 定	then $(rr/nr/r)+X+(rr/nr/r)+(*)$ and $X=\{ \Xi \}$		
不该不	该	If (rr/nr/r)+不该不+(rr/nr/r)+(*),		
		then $(rr/nr/r)+X+(rr/nr/r)+(*)$ and $X=\{ig\}$		
不是不	@,是	If (rr/nr/r)+不是不+(rr/nr/r)+(*),		
		then (rr/nr/r)+X+(rr/nr/r)+(*) and X={是}		
		If (rr/nr/r)+ 不得不承认+(rr/nr/r)+(*).		
不得不承认	承认,必须承认	then (rr/nr/r)+X+(rr/nr/r)+(*) and X={必须承认}		
		If (*)+v+(rr/nr/n)+不于+v+v		
不无	有	$(-)^{+v+(1/10/10+v+v+y)}$		
		If (*)+ v+(rr/nr/n)+无不+v+v.		
无不	都	(+) + (+)		
		If (*)+ v+(rr/nr/n)+未必不+v+v。		
未必不	@,也	then $(*)+v+(rr/nr/n)+X+v_{\circ}$ and $X=\{@, \pm\}$		
おどろ	0	If (*)+v+(rr/nr/n)+难道不+v+y。		
难迫个	@	then $(*)+v+(rr/nr/n)+X+v_{\circ}$ and $X=\{@\}$		
		If (*)+v+(rr/nr/n)+不是没+v+y。		
不是没	@,有,还是有	then (*)+v+(rr/nr/n)+X+v。and X={@,有,还是有}		

Table 5	Paraphrase	Templates 1
---------	------------	--------------------

In Table 5, '@' is deletion. The paraphrase processing is deletion the double negative words.

Table 6 shows the double negative paraphrase templates that negative words are separate. For example:

这件棘手的事情非他不能解决。

Paraphrase: '这件棘手的事情他能解决。', '这件棘手的事情只有他能解决。'.

We called this kind of negative sentences is separate double negative sentences.

Key words	Paraphrase key words	Paraphrase templates
不是不	是@	If (*), v+(rr/nr/n)+不是不+n,
		then (*), v+(rr/nr/n)+X+v。 and X={是}
いた て	都@	If (*), v+(rr/nr/n)+没有不+n,
没有个		
		then (*), $v+(rr/nr/n)+X+v_{\circ}$ and X={都}
	@很	If (*), v+(rr/nr/n)+不会不+n,
个会个		
		then (*), $v+(rr/nr/n)+X+v_{\circ}$ and $X=\{\&, \Leftrightarrow\}$
		If (*), y+(rr/nr/n)+非 不+n,
非不	只有	
		then (*), v+(rr/nr/n)+X+v。 and X={只有}
		If (*) v+(rr/nr/n)+非 不可+n.
非不可	@@	$(), v + (11/11/11) + \exists r r + v : 1 + 11 $
		then (*), $y+(rr/nr/n)+X+y$, and $X=\{@, 必须\}$
	@必须	

Table 6. Paraphrase Templates 2

3. Paraphrasing Process

Paraphrasing process consists of templates matching, similarity calculation of paraphrase sentences with paraphrase templates and generation of paraphrase sentences.

Templates Matching

The key words of original sentences, the double negative part, are compared with paraphrase template to select paraphrase template. When there is not matched template that is same key word of original sentence with template in the corpus, the system finds a matched template by calculating similarity of original sentence with template.

About text similarity method, there are some researches. Paper [15] proposes a similarity model to measure the similarity between sentences. The method supposes the sentence similarity depends on the morphological similarity and word order similarity, and the former plays more important role than the latter. To support the retrieval of semi-structure date, paper [16] proposed a fast and effective method that can have ontological similarity flexibility same as XClust, but does not have big velocity delay. Paper [17] proposed a sentence similarity computation method that integrates the syntax feature, semantic feature and word feature of the sentences. And the method endows the three features with different weights.

Similarity calculation is used to match part of speech of sentence and key word of double negative part. The key word that is the double negative part in a sentence is called key item. The other parts which except key item in a sentence is called restriction item. The match of key item carries out by comparing original sentence with template. And the match of restriction item is completed by calculating similarity. Then the match of original sentence with template carries out by two steps, key item and restriction item. After match of key item, the similarity is calculated by formula (1).

$$RistrictionItemSimilarity[\%] = \frac{\alpha \cdot NRIW(T, S) + \beta \cdot NRIP(T, S)}{NRIW(T) + NRIP(T)} \times 100$$
(1)

Where *NRTW* is the number of same words in restriction item that compare original paraphrase sentence with the paraphrase template, and *NRTP* is the number of same POS in restriction item that compare original paraphrase sentence with the paraphrase template. T and S denote template and original paraphrase sentence respectively. *RistrictionItemSimilarity* is

the similarity of the being paraphrased sentence to match with the template. In addition, α and β are weighted parameters. The weighted parameter α and β are confirmed by using Greedy method through preliminary paraphrase experiment for 100 sentences, $\alpha = 0.7$ and $\beta = 0.3$.

4. Evaluation Experiments

A small scale corpus is used in evaluation the paraphrase performance of presented method. The experiment system is developed by JAVA program language. The experiment system is showed as Figure 1.

Par	raphrase System of Double Negative Sentence V1.0	
	Result of Paraphrase	
	我去图书馆看书。 我必须去图书馆看书	输入文件
	我应该去图书馆看书。	改写
		退出
	Input Sentences	
	我不得不去图书馆看书。	清零

Figure 1. The Paraphrase System

4.1. Experimental Data

To evaluate performance of the paraphrase system, the small scale corpus there are 398 sentences are used. The experiment is based on 19 kinds of templates; each kind of template is divided into several more idiographic templates, there are total of 85 templates. The paraphrase is not one-to-one, but one-to-many. That is, one sentence will be paraphrased to several sentences.

4.2. Experimental Results

The paraphrase results are 1050 sentences. The right or wrong judgment of paraphrase results is carry out by manual. The benchmark of judgment can be classified into correct paraphrasing and erroneous paraphrasing. The correct paraphrasing: the paraphrasing sentence is correct, to express the same meaning or less wrong and the meaning of the expression remain basically unchanged; erroneous paraphrasing: lack of information, chaotic word order and changing the meaning of the expression. The experimental results are showed as Table 7.

Original	Paraphrase	Non-	Right	Wrong
sentences	sentences	paraphrase	paraphrase	paraphrase
398	1050	12	879	159

Recall ratio and precision of paraphrase are calculated by formula (2) (3) respectively.

International Journal of Multimedia and Ubiquitous Engineering Vol. 8, No. 4, July, 2013

$$Recall[\%] = \frac{number \ of \ paraphrased \ sentence}{total \ number \ of \ original \ sentence} \times 100$$
(2)

$$precision[\%] = \frac{number \ of \ right \ paraphrase}{total \ number \ of \ original \ sentence} \times 100$$
(3)

The evaluation results are recall 98.9%, precision 84.7%.

5. Discussion

Because in this paper paraphrase object is only double negative sentences. The structure of double negative sentence is correspondingly regular, the research content are restricted in fixed scope. The extracted templates are also used in limitative object, so that a high precision is achieved.

To get the main reason of wrong paraphrased and non-paraphrase, we review the experimental results. The reason of non-paraphrase is not enough templates and experimental corpus contain non-double negative sentence. The main reason of wrong paraphrase is the problem of similarity calculation. The error template is selected at a lower similarity because there is not template that matches with original sentence.

6. Conclusion

In this paper we have discussed the paraphrase problem of double negative sentences. We use the method based on templates. At first, we have classified the double negative sentences to three kinds, and then analyzed the structure of each kind of double negative sentences. We have extracted the paraphrase templates by comparing the syntax structure of double negative sentences with their affirmation sentences.

We have also constructed a method of similarity calculation. The similarity calculation are consists of part of speech of sentence, key word of double negative part, key item and restriction item. The match of key item carries out by comparing original sentence with template. And the match of restriction item is completed by calculating similarity.

To evaluate the performance of proposed method, we have carried out a small scale experiment. The experimental results indicated the method based on template is effective to double negative sentences. The method has achieved a result of precision 87.5% and recall ratio 96.3%. Due to the templates are not enough and the similarity calculation method is not perfect, some sentences can not paraphrased and make some mistakes in paraphrase.

For the future work will be research the similarity calculation method, extract more paraphrase templates and carry out more large scale experiment.

Acknowledgements

This work was financially supported by Open Fund of Smart Education and Information Engineering (Harbin Normal University), the Scientific Research Foundation of Education Bureau of Heilongjiang Province: (12511127) and Natural Science Foundation of Heilongjiang province (F201243).

References

- D. Das and N. A. Smith, "Paraphrase Identification as Probabilistic Quasi-Synchronous Recognition", Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP, Suntec, Singapore, (2009) August 2-7.
- [2] K. Ohtake and K. Yamamoto, "Applicability analysis of corpus-derived paraphrases toward example-based paraphrasing", Proceedings of the 17th Pacific Asia Conference: Language, Information and computation, Sentosa, Singapore, (2003) October 1-3.
- [3] Fabio Rinaldi, James Dowdall and Diego Moll, "Exploiting Paraphrases in a Question Answering System", Proceedings of Workshop in Paraphrasing at ACL2003, Sapporo, Japan, (**2003**) July 7-12.
- [4] T. Tetsuro, N. Kozo, K. Shinya and I. Kentaro, "Applying Structural Matching and Paraphrasing", NTCIR Workshop 3 Question Answering Challenge (QAC), (2002).
- [5] K. Inui, A. Fujita, T. Takahashi and R. Iida, "Text Simplification for Reading Assistance: A Project Note", Proceedings of the second international workshop on Paraphrasing, Sapporo, Japan, (2003) July 11-16.
- [6] H. Kataoka, N. Watanabe, K. Mizutani and H. Yoshiura, "DCNL: Disclosure Control of Natural Language Information to Enable Secure and Enjoyable E-Communications", International Journal of u- and e- Service, Science and Technology, vol. 3, no. 1, (2010).
- [7] Y. Zheng-tao, D. Bin1, H. Bo, H. Lu1 and G. Jian-yi, "Word Sense Disambiguation Based on Bayes Model and Information Gain", International Journal of Advanced Science and Technology, vol. 3, (2009).
- [8] Y. Lepage Etienne Denoual, "Automatic generation of paraphrases to be used as translation references in objective evaluation measures of machine translation", Third International Workshop on Paraphrasing, (IWP2005), Jeju Island, Korea, (2005) October 14-16.
- [9] S. Dumais, M. Banko, E. Brill, J. Lin and A. Ng, "Web question answering: Is more always better?", SIGIR 2002: Proceedings of the 25th Annual International ACMSIGIR Conference on Research and Development in Information Retrieval, (2002), pp. 291-298.
- [10] C. Zong, Y. Zhang, K. Yamamoto, M. Sakamoto and S. Shirai, "Approach to Spoken Chinese Paraphrasing Based on Feature Extraction", 6th Natural Language Processing Pacific Rim Symposium, Tokyo, Japan, (2001) November, pp. 551-556.
- [11] S. Zhao, T. Liu, X. Yuan, S. Li and Y. Zhang, "Automatic Acquisition of Context-Specific Lexical Paraphrases", Proceedings of IJCAI, Hyderabad, India, (2007) January, pp. 1789-1794.
- [12] W. Li, T. Liu, Y. Zhang, S. Li and W. He, "Automated Generalization of Phrasal Paraphrases from the Web", The 3rd International Workshop on Paraphrasing, Jeju Island, South Korea, (2005) October, pp. 49-57.
- [13] Double negative http://en.wikipedia.org/wiki/Double_negative.
- [14] ICTCLAS (Institute of Computing Technology, Chinese Lexical Analysis System): http://www.ictclas.org/index.html.
- [15] L. Xueqiang, R. Feiliang, H. Zhidan and Y. tianshun, "Sentence Similarity Model and the Most Similar Sentence Search Algorithm (in Chinese)", Journal of Northeastern University(Natural Science), vol. 24, no. 6, (2003) June.
- [16] H. Moon, K. Kim, G. Park and C. Wooyoo, "Effective Similarity Discovery from Semi-structured Documents", International Journal of Multimedia and Ubiquitous Engineering, vol. 1, no. 3, (2006).
- [17] W. Quane and X. Hai-Ling, "Method for Sentence Similarity Computation by Integrating Multi-Features (in Chinese)", Computer system and application, vol. 9, no. 11, (2010) November.

Authors



Zhongjian Wang, Ph.D., professor, his main research interests include natural language process, Chinese sentence paraphrase, Chinese word segmentation and Information retrieval, *etc*.

International Journal of Multimedia and Ubiquitous Engineering Vol. 8, No. 4, July, 2013