A Study on Cross-Language Text Summarization Using Supervised Methods

Lei Yu
Graduate School of Advanced Science
Technology Education, The University of Tokushima
2-1 Minamijosanjima, Tokushima, 770-8506,
Japan
yulei@is.tokushima-u.ac.jp

Fuji Ren
[1]Graduate School of Advanced Science
Technology Education, The University of Tokushima
2-1 Minamijosanjima, Tokushima, 770-8506,
Japan
[2]School of Information Engineering, Beijing
University of Posts and Telecommunications
Beijing, 100876, China
ren@is.tokushima-u.ac.jp

Abstract:

In this work, we use Hidden Markov Models (HMM), Conditional Random Field (CRF), Gaussian Mixture Models (GMM) and Mathematical Methods of Statistics (MMS) for Chinese and Japanese text summarization. The purpose of this work is to study the applicability of mentioned three trainable models for cross-language text summarization. For model training, we use several training features such as sentence position, sentence centrality, number of Name Entity and so on. For model testing, Chinese on-line news and Japanese news are used as test data which are extracted from web pages. We evaluate each model by measuring the precision at the compression rate 10%, 20% and 30%. MMS is a baseline method. The results show that HMM, CRF and GMM have remarkable increases than MMS on both Chinese and Japanese text summarization by using the same training features. Especially, GMM model make a best performance in all tests.

Keywords:

Text Summarization; NLP; Machine Learning

1. Introduction

The process of text summarization can be seen as a data reduction to compress the content of document. Text summarization can give use a quick view of a text or document. Readers need not to read the whole original content of text, because an automatic can extract important sentences from text, and then generate a series of compressed machine-summaries. With the repaid development of internet, more and more online information is available, as a result information explosion has made us more difficult to browse the information we are interested in. So it is necessary to develop automatic summarization systems. In other words, the main task of text summarization is to distilling the most important information from an original text (or texts), or a source. There are two important tasks in text summarization, one is how to extract important sentences from original text, and the other one is how two order these selected sentences.

The early research use methods based on structural features is a general idea for extracting importance sentence from the source text. Edmundson and Luhn have proposed four sentence features to decide importance of sentences, including the high frequency words (keywords), pragmatic words (cue words), title and heading words, and structural indicators (sentence location). This basic method is still used as a paradigm by the other researches. And then several extended approaches such as lexical chains, structure of article, rules and statistics are proposed.

Based on several reports about text summarization, the general abstract architecture for summarization is viewed in three steps: analysis, transformation or called refinement, synthesis. The first step analysis builds an internal representation, and the second step transforms the internal representation. The last step uses natural language to render the summary representation. The compression of summaries is another key problem for text summarization, based on the research about compression rate of summaries, it is said that the compression rate at 5-30% are acceptable by the human. When the compression rate is out of 30%, some noisy information will increase. In this paper, for Chinese text and Japanese text, we use the compression rate: 10%, 20%, and 30%.

In the last decades, with the development of computer hardware, it is possible to use a large amount corpus to finish a machine learning experiment.

Furthermore, artificial intelligence (AI) has been applied to Natural Language Processing field. Some efficient machine learning models such as Hidden Markov Model (HMM), Conditional Random Fields (CRF), and Maximal Entropy model have been used for text classification, Pos Tagging, Name Entity Recognition (NM) and the other NLP fields. In the summarization field, Shen et al. use CRF model for the document summarization. Mohamed et al. have used FFNN, PNN, GMM models for Arabic and English text summarization. There is also research for specific domains summarization such as biomedical field (Reeve et al., 2006, 2007; Ling et al. 2007).

In this paper, we assume that there are two types of sentences in the documents: "summarization sentence" and "non-summarization sentence". Based on this assumption, summarization can be seen as a two-class classification problem. For the classification model, "summarization sentence" is given a value '1' and "non summarization sentence" is given a value '0'. And then several classification models are applied to this problem. For the model training, the sentence score is computed based on the manually corpus extracted from manually corpus. Summaries of original texts are generated from a set of score sentences at the compression rate of 10%, 20% and 30%.

The rest of this paper is organized as follows: section 2 describes how to select important features parameters for summary models training. In section 3, we present several supervised models for text summarization. Section 4 discusses the difference between Asia language texts and English texts, and then describes effect of word segmentation to automatic summarization by using different parsers. Section 5 introduces evaluation test. Section 6 concludes our work and presents our future work.

2. Text Features

As a reason, that the grammar and syntax of Chinese and Japanese are different from English. The structure of the English texts is more intact compared with the Chinese and Japanese texts. Therefore, how to choose important text features for the model training is a key problem. Mohamed et al., use ten text features for Arabic and English text summarization. These ten features including sentence position, positive keyword, negative keyword, sentence centrality, sentence resemblance to the title, sentence inclusion of Name Entity, sentence inclusion of numerical data, sentence relative length, Bushy path of the sentence and aggregated similarity for each sentence to generate summaries. Based on these features, with the combination to the characters of Chinese and Japanese languages, we use following 6 text features:

(1). The similarity with other sentences. This feature is the vocabulary overlap between this sentence and other sentences in the whole texts. The score of this feature is calculated using the following formula:

$$Score_{f_3}(s) = \left| \frac{proper\ noun\text{s in } s}{Lengh(s)} \right| \qquad (1)$$

Where $Score_{f1}$ is weight of feature (1), *KWS* is number of keywords in sentence *s*, *KWOS* is number of keywords in other sentences.

(2). Sentence position in document. This feature is based on the assumption that the first sentence of a paragraph is the most important. It is generally extracted as the summary sentence. The weight of sentence is given based on the position. We assume that maximum of position of a paragraph is 5. Therefore, if the sentence in the first position in a paragraph, it will be given a value 0.2, the second will be given a value 0.4 and so on. If the sentence position is out of 5, it will be given a value 0.

(3). Number of Name entity that the sentence includes. Generally, summary sentence contains more name entities. It means the sentence contains more name entities is an important one and it is most probably a summary sentence. In this paper, we consider the name entities is nouns and nouns phrases. The weight of this feature can be calculated use the following formula:

$$W_{f_3} = \frac{NES}{WS} \qquad (2)$$

Where W_{f3} is weight of feature (3), *NES* is number of name entities in sentence *s*, *WS* is number of words in sentence *s*.

(4). Sentence that contains numeric character. Important sentence usually contains several conclusively numeric characters. The weight of this feature can be calculated use the following formula:

$$W_{f_4} = \frac{NCS}{WS} \qquad (3)$$

Where W_{f4} is weight of feature (4), *NCS* is number of name entities in sentence *s*, *WS* is number of words in sentence *s*.

(5). Degree of similarity between sentence and title. Sentence with high degree of similarity to title is important sentence. And it can be calculated using following formula:

$$W_{f_5} = \frac{KWS \cap KWT}{KWS \cup KWT} \qquad (4)$$

Where W_{f5} is weight of feature (5), *KWS* is number of keywords in sentence *s*, *KWOS* is number of keywords in title.

(6). Links to the other sentences. Document can be seemed as a map made up of sentences. Each sentence is a node on a map. The score between every two sentences can be calculated based on the general Word Frequency

(WF) method. One node (sentences) linked with a number of branches means this sentence cover the other sentences in content and it is important sentence in the full document. Experiment of Mohamed et al., shows that links to the other nodes is a very effective feature for summarization task.

3. Summarization Model

In this section, we use four summarization models. One is a Statistical Mathematical model, which is used as a baseline approach. The other three models are HMM, CRF, and Gaussian Mixture model (GMM). A basic processing for a trainable text summarization is shown in Figure 1.

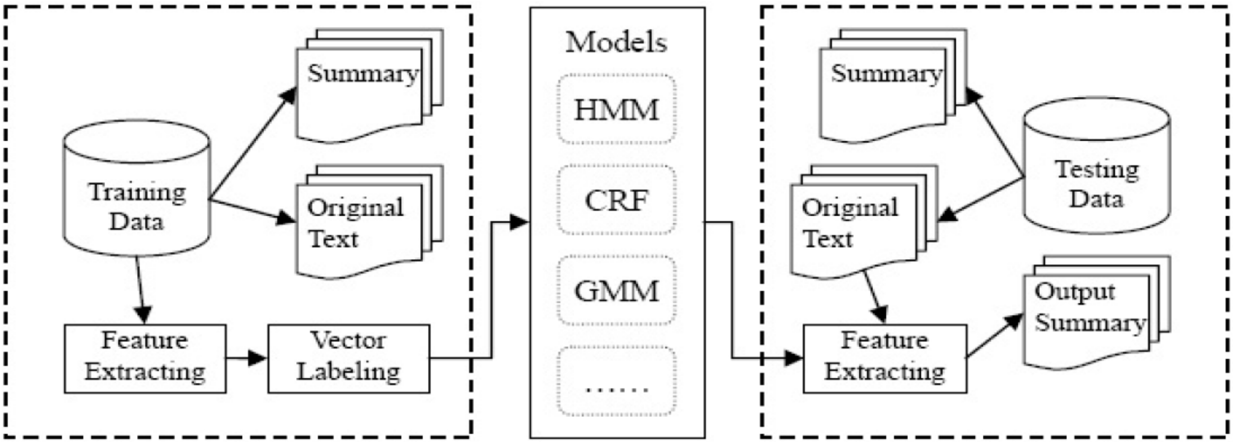

Figure 1
General process of corpus-based text summarization

3.1 Mathematical Methods of Statistics (MMS)

For a sentence *s*, a weighted score function is shown as follows:

$$Score(s) = \alpha_1 \cdot Score_{f_1}(s) + \alpha_2 \cdot Score_{f_2}(s) + \alpha_3 \cdot Score_{f_3}(s)$$
$$+\alpha_4 \cdot Score_{f_4}(s) + \alpha_5 \cdot Score_{f_5}(s) + \alpha_6 \cdot Score_{f_6}(s) \quad (5)$$

Where α_i is the weight of feature f_i, the value of i is [1, 2]. All the features are integrated to calculate the weight of each feature. For the weight calculation, the key problem is how to estimate weight value of α_i. This is an optimization problem. We use a general method, genetic algorithms (GAs), to estimate the weights of α_i. Genetic algorithms are categorized as global search heuristics. Usually, a genetic algorithm (GA) is used for a search technique to find approximate solutions to optimization. With the development of Internet, there are many optimization problems came out frequently in NLP field. The results make GAs quite useful around these classical problems of optimization. As a classical optimization problem, how to maximize/minimize an objective function f over a given space X of arbitrary dimension is the critical component. Therefore, GAs can be applied to estimate the weight of each feature mentioned in section 2.

For obtaining optimal weights of each parameter α_i, we use manually summarized Chinese documents and Japanese documents. A chromosome of GAs is defined by using the integrating all of these six feature scores in the form of (α_1, α_2, α_3, α_4, α_5, α_6). After training data is used for model training, we can get many precision about genome, and a number of genomes of GA are obtained. According to the average precision, we select most suitable genomes as the weights of these six parameters. We get two weights sets of α_I by using Chinese generations and Japanese generations.

For testing, a set of Chinese documents and Japanese documents were used. We apply formula mentioned in 3.1 to calculate weight of each sentence. After using the defined weights of α_I from GA execution. All sentences in documents are ranked in a descending order based their scores. A set of highest score sentences are extracted from all sentences in document to construct a set of summary sentences using three different compression rate, 10%, 20%, and 30%.

3.2 Supervised Learning Models

Supervised learning models such as Hidden Markov Models (HMM) and Conditional Random Fields (CRFs) have been used for text summarization in recent years (John et al., 2001; Pascale et al., 2003; Shen et al., 2007). A hidden Markov model (HMM) is widely used as a general statistical model in NLP field such as text classification and word segmentation. HMM can be considered as a dynamic Bayesian network. A state in HMM is not directly visible, but output dependent on the state is visible. In particular, it does not assume that the probability that *sentence$_i$* is in the summary is independent of whether *sentence$_{i-1}$* is in the summary. Furthermore in our task, a joint distribution is proposed for the text features set, unlike the independence-of-features assumption used by naive Bayesian methods.

CRFs is a widely used in a probabilistic framework. It has a better performance in many labeling task than Hidden Markov Models (HMMs) and Markov Random Fields (MRFs). If CRFs is given input sample x, it can directly model $P(Y|x)$. In CRF model, the joint density $P(X,Y)$ is modeled based on inputs X and labels Y. In addition, CRF model has a better capability on processing long-range interactions. This simulating is better than the other model which has a disadvantage on so-called label bias.

Gaussian Mixture models (GMM) is a statistical learning model, which can work well in text clustering. It is reported that "The components of GMM can be used to represent some general output dependent features". The following formula can calculate the probability of observing X for a given class modelλ_c.

$$p(X \mid \lambda c) = \sum_{k=1}^{K} w_{c,k} N(X; \vec{\mu}, \sum{}_{c,k}) \tag{6}$$

Based on this approach, a trainable text summarizer can be built to do a two-class classification model.

In a task of two-categories based text summarization, the classification model is given a class-dependent reference models (λ_1, λ_2). And then, an input feature X which is made up of $x_1, x_2, x_3, \dots, x_n$ is classified as one of two categories by using the following formula:

$$\arg\max_{\iota} p(\lambda\iota \mid X) = \arg\max_{\iota} \frac{p(X \mid \lambda\iota)}{p(X)} p(\lambda\iota) \tag{7}$$

Where the value range of l is [1, 2].

4. Effects of POS tagging to Summarization

Word segmentation is first step of many natural language processing (NLP). Unlike English, some Asia language document is written without space. Therefore, word segmentation is very important for the other applications such as text summarization, text classification, machine translation. Results of word segmentation directly affect the precision of subsequent application of NLP. As a reason that Chinese text is written without natural delimiters, word segmentation is a basic step for Chinese natural language processing. As a fact, NLP techniques in Chinese are quite different from English, and the following applications of Chinese NLP are affected by the results of word segmentation. It seems that word segmentation is a very simple problem, but in Chinese NLP field, research have paid attention and studied this problem for almost 20 years. As a basic key task for Chinese NLP, many methods and many practical methods have been developed.

In all of proposed methods, character based tagging method quickly rose as a remarkable one with state-of-the-art performance. "character-based" approach has a higher performance than word-based approach. In all of recent proposed methods, character-based tagging method gets a remarkable performance then the others. Character-based tagging method can has higher precision than the word-based tagging method in processing the unknown word, so that the character-based tagging method gains an overall higher score than the word-based tagging method.

The character has some cost problems, because the character-based method has to process the sentences treating them as sequences of characters. Meanwhile, the character-based method loses some information about the words that might be important for processing the known words. We build a CRF-based Chinese word segmentation system. We improve CRF-based tagging method of Chinese word segmentation. There two cut-in points in our methods. First one we focus on how to construct a effective feature template. Second one is we use a more effective tag set to train model. For Japanese word segmentation, we use the Japanese morphological analysis System "ChaSen". "ChaSen" is widely used in Japanese NLP field, which is based on HMM model. All of test data is manually segmented to construct a "golden set" for investigating negative effects of word segmentation to text Summarization. This means precision of word segmentation result is 100%. We will use these two sets to compare the precision of text summarization in section 5.

5. Experiment

5.1 Test Data

For testing, we select 200 Chinese articles from web pages and 120 Japanese articles from NTCIR summarization corpus. The average number of sentences included in articles is 24.2. Among these articles, 150 Chinese articles and 90 Japanese articles are manually summarized as training data by using compression rate 30%. The rest of Chinese and Japanese articles are used as test data to evaluate performances of 4 methods mentioned in section 3. All of articles are segmented by human and segmentation tools to construct two set of test data. We called manually set "golden set". We use an intrinsic evaluation to compare the performance of each summarizer. The scores of each summarizer are calculated by using the following formula:

$$Precision = \frac{S \cap T}{S} \tag{8}$$

$$Precision = \frac{S \cap T}{T} \quad (9)$$

Where S is the set of summary sentences generated by the trainable summarizer; T is the set of summary sentences manually generated by human. Because the candidate summaries and the reference summaries are compared at the same compression ratio, it means $S = T$, respectively, *precision* = *recall*.

5.2 Test Performance for each Feature

For investigating the effect of each feature on summarization task, we use the Mathematical Methods of Statistics (MMS) and the formula mentioned in section 3.1 to test the performance of each feature.

Table 1
Performances of each feature at different compression rate (Chinese)

Feature	CR 10%	CR 20%	CR 30%
feature(1)	0.423	0.433	0.421
feature(2)	0.409	0.413	0.420
feature(3)	0.385	0.391	0.382
feature(4)	0.332	0.337	0.331
feature(5)	0.412	0.419	0.421
feature(6)	0.486	0.491	0.484

Table 2
Performances of each feature at different compression rate (Japanese)

Feature	CR 10%	CR 20%	CR 30%
feature(1)	0.452	0.437	0.442
feature(2)	0.432	0.429	0.423
feature(3)	0.402	0.383	0.392
feature(4)	0.358	0.361	0.353
feature(5)	0.456	0.469	0.461
feature(6)	0.521	0.529	0.518

The results are showed in table 1 and table 2. Performance of Japanese summarization is higher than Chinese. And performance of feature (6) is highest in both Chinese and Japanese. It seems that this feature have a stronger applicability in Chinese and Japanese text summarization.

5.3 Results of Each model

For training the models, we use Chinese and Japanese manually summarized articles as the training data. The rest of articles are used for testing performance of each model. All the features used as training parameters. In addition, the "golden" segmentation result of Chinese and Japanese articles are applied to the test data. The results using "golden set" are compared to the results using segmentation tools, which are our Chinese segmentation tool and Japanese segmentation tool "ChaSen".

We use MMS as a baseline method. All features are used for testing. Weight of sentence is calculated by using all features and summary sentences are extracted at the compression rate of 10%, 20% and 30%. We also use two segmentation results in the test. Table 3 and Table 4 show the results of each model at different compression rate by using segmentation tools. From this table, the performance of HMM is higher than MMS. The performance of GMM model has a remarkable improvement than the other models. The highest precision of this model is 0.603 at the compression rate of 20% by using Chinese segmentation tool. In the results of Japanese text summarization, the highest performance is also obtained by GMM model.

Table 3
Performances of each model (Chinese)

Models	CR 10%	CR 20%	CR 30%
MMS	0.501	0.505	0.503
HMM	0.504	0.510	0.507
CRF	0.532	0.527	0.530
GMM	0.590	0.603	0.601

Table 4
Performances of each model (Japanese)

Models	CR 10%	CR 20%	CR 30%
MMS	0.524	0.530	0.528
HMM	0.531	0.540	0.537
CRF	0.545	0.539	0.541
GMM	0.611	0.623	0.618

5.4 Discussion

It is clear from Table 1 and Table 2 that the most important text feature for summarization is feature 6 (links to the other sentences), as a result that it gets the highest precision in all features. Feature 1 (similarity with other sentences) also get good performance since it conveys the vocabulary overlap between this sentence and other sentences from the document. The performance of feature 6 makes it reasonable that the sentence which has a maximum number of branches convey the most important part of documents. Usually, the document title conveys the main topic of this document. Therefore, feature 5 (Degree of similarity between sentence and title) which is the vocabulary overlap between this sentence and the title of document gets good performance. Feature 4 (Sentence that contains numeric character) gets the lowest results in the test, after analyzing content of articles, we find that there are not a lot of numerical characters in sports articles and the other religious news. The summarization results also show that these text features have a steady applicability in Chinese and Japanese text summarization.

Table 3 and Table 4 show that supervised methods improve the performance of text summarization. As GMM is a probability mixture model and it is a probability distribution which is a convex combination of other probability distributions. GMM has a better capability of simulating arbitrary densities than the other supervised models. As a result, this model gets good performance in all tests. The results compared with MMS show that GMM makes a further improvement in the tests of Chinese and Japanese summarization tasks. It means that GMM is a more powerful method in exploiting dependent features than the other models. Both HMM and CRF improve the performance of summarization as compared to MMS. As a reason that HMM cannot exploit the rich features better, the model does not get a better performance than CRF model and GMM model. Furthermore, the evaluation results prove the possibility of using supervised methods in cross-language text summarization.

Table 5
Performances of each model using "golden set" (Chinese)

Models	CR 10%	CR 20%	CR 30%
MMS	0.513	0.518	0.515
HMM	0.519	0.527	0.523
CRF	0.546	0.550	0.542
GMM	0.605	0.621	0.613

Table 6
Performances of each model using "golden set" (Japanese)

Models	CR 10%	CR 20%	CR 30%
MMS	0.529	0.539	0.528
HMM	0.538	0.550	0.543
CRF	0.551	0.558	0.554
GMM	0.619	0.630	0.625

For estimating negative effects of word segmentation to text summarization, we use a manually word segmentation result to compare the automatic segmentation tools in Chinese and Japanese. We build a Chinese segmentation tools based on CRF model. In the Japanese word segmentation we use the tool "ChaSen". The comparison results are showed in table 5 and table 6. From the results of Table we can clearly find that word segmentation brings negative effect to summarization results. In the test of Chinese text summarization, the average precision of GMM model by using segmentation tool is 0.598 and the average precision by using "golden set" is 0.613. The precision of GMM model in Japanese case are 0.617 and 0.625. "golden set" bring 0.015 and 0.008 higher precision in Chinese and Japanese text summarization task. In the current of Asia language processing field, the precision of Japanese word segmentation is higher than Chinese. The segmentation results by using "ChaSen" are closer to "golden set". As a result, Chinese get a higher growth rate than Japanese. In addition, "ChaSen" can recognize Name Entity (NM) better than our Chinese segmentation tool, so it also obtains the test results of text summarization.

6. Conclusion

In this paper, we use Mathematical Methods of Statistics (MMS), Hidden Markov Model (HMM), Conditional Random Fields (CRFs), and Gaussian Mixture Model (GMM) for text summarization task. Among these four methods, HMM, CRFs, and GMM are trainable method. For training the models, we use four text feature parameters: similarity with other sentences, sentence position in document, number of name entity that sentence includes, sentence that contains numeric character, degree of similarity between sentence and title, and links to the other sentences. 150 Chinese articles and 90 Japanese articles are used as training data, and the rest of articles are used as test data. For investigating whether the results of word segmentation bring effects to summarization, we use two set of word segmentation in the experiment. The results shows supervised methods based models improve the performance of summarization. Moreover, these supervised models are available in Cross-Language text summarization. The analyzing results also shows that word segmentation bring negative effect to the summarization task.

In our future work, we will use more data for the model training to investigate whether increasing data improve the performance of summarization results. More text features will be extracted and used for trainable summarizer.

Acknowledgment

This research has been partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (B), 19300029, and Challenging Exploratory Research, 21650030.

References

[1] J. Morris and G. Hirst, "Lexical Cohesion Computed by Thesaural Relations as an Indicator of the Structure of Text", Computational Linguistics, vol. 17, pp. 21-43 (1991)

[2] R. Barzilay and M. Elhadad, "Using Lexical Chains for Text Summarization", in Proceedings of the Intelligent Scalable Text Summarization Workshop (ISTS'97), ACL, pp. 10-18 (1997)

[3] FU Jian-lian, CHEN Qun-xiu, "Research on Automatic Summarization Based on Rules and Statistics for Chinese Texts", JOURNAL OF

CHINESE INFORMATION PROCESSING, Vol.20, No.52006 (2006)

[4] H. P. Luhn, "The automatic creation of literature abstracts", IBM Journal of Research and Development", 2-2, pp.159-165 (1958)

[5] H. P. Edmundson, "New Methods in Automatic Extraction [A], In Advances in Automatic Text Summarization[C]", pp.23-42 (1998)

[6] Hahn, U., Mani,I., "The challeges of automatic summarization", IEEE-computer 33 (11), pp.29-36 (2000)

[7] Hovy, E., Lin, C.Y., "Automatic text summarization in SUMMARIST", IN the proceceeding of the ACL'97/EACL'97 Workshop on INtelligent Scalable Text summarization, Madrid, Spain, pp.18-24 (1997)

[8] Mani, I., Maybury, M.T.(EDs.), "Advances in Automated Text Summarization", The MIT Press, Cambridge, MA (1999)

[9] Fuji Ren, Shigang Li, Kenji Kita, "Automatic abstracting important sentences of web articles", IEEE International Conference on Systems, Man, and Cybernetics, vol.3, pp.1705-1710 (2001)

[10] Hirao, T., Okumura, M., Yasuda, N., Isozaki, H., "Supervised automatic evalutation for summarization with voted regression model", Information Processing & Management 43 (6), pp. 1521-1535 (2007)

[11] Nomoto, T., "Discriminative sentence compression with conditional random fidels", Information Processing & Management, 43(6), pp.1571-1587 (2007)

[12] Reeve, L., Han, H., Brooks, A., "the use of domain-specific concepts in biomedical text summarization", Information Processing & Management, 43(6), pp.1765-1776 (2007)

[13] Mohamed Abdel Fattah, Fuji Ren, "Automatic Text Summarization", International Journal of Computer Science, Vol.3, No.1, pp.25-28 (2008)

[14] M.A. Fattah, Fuji Ren, "GA, MR, FFNN, PNN and GMM based models for automatic text summarization", Computer Speech and Language 23(1), pp.126-144 (2009)

[15] Lei Yu, Mengge Liu, Fuji Ren, Shingo Kuroiwa, "A Chinese Automatic Text Summarization system for mobile devices", The Pacific Asia Conference on Language, Information and Computation (PACLIC-2006), pp.426-429 (2006)

[16] Yuji Matsumoto and Masayuki Asahara, IPADIC User Manual version 2.2.4. Nara Institute of Science and Technology (2001)

[17] Yuji Matsumoto, Akira Kitauchi, Tatsuo Yamashita, Yoshitaka Hirano, Hiroshi Matsuda, Kazuma Takaoka, and Masayuki Asahara, Morphological Analysis System ChaSen version 2.2.8 Manual. Nara Institute of Science and Technology (2001)

[18] Huang Degen, Sun Xiao "An Integrative Approach to Chinese Named Entity Recognition", In proceedings of Sixth International Conference on Advanced Language Processing and Web Information Technology, pp.171-176 (2007)

[19] Pascale Fung, Grace Ngai, Chi-Shun Cheung, "Combining optimal clustering and Hidden Markov models for extractive summarization", Proceedings of the ACL 2003 workshop on Multilingual summarization and question answering, Vol.12, pp.21-28 (2003)

[20] Brants, Thorston Brants, "TnT-A Statistical Parts-of-Speech Tagger", Proceedings of Sixth Applied Natural Language Processing Conference ANLP-2000 (Seattle, WA), pp.224-231 (2000)

[21] Charniak,E., Hendrickson, C., Jacobson, N., and Perkowitz, M., "Equations for part of speech tagging", In Proceedings of the Conference of the American Association for Artificial Intelligence(AAAI-93), pp.784-789 (1993)

[22] Xiao SUN, Fuji Ren and Degen Huang, "Dual-chain Unequal-state CRF for Chinese New Word Detection and POS Tagging", IEEE NLP-KE 2008, pp.60-66, Beijing, Oct. (2008)

[23] Ani Nenkova, Rebecca Passonneau, Kathleen McKeown, "The pyramid method: incorporating human content selection variation in summarization evaluation", ACM Transactions on Speech and Language Processing, 4(2). (2007)

[24] Charles Sutton and Andrew McCallum and Khashayar Rohanimanesh, "Dynamic Conditional Random Fields: Factorized Probabilistic Models for Labeling and Segmenting Sequence Data", Journal of Machine Learning Research, Vol.8, pp.693-723 (2007)

[25] Dou Shen, Jian-Tao Sun, Hua Li, Qiang Yang, Zheng Chen, "Document Summarization using Conditional Random Fields", Proceeding of International Joint Conferences on Artificial Intelligence, pp.2862-2867 (2007)

[26] Hua-Ping ZHANG, Qun LIU, Xue-Qi CHENG, Hao Zhang, Hong-Kui Yu, "Chinese Lexical Analysis Using Hierarchical Hidden Markov Model", Second SIGHAN workshop affiliated with 41th ACL, Sapporo Japan, pp.63-70 (2003)

[27] John M. Conroy, Dianne P. O'leary, "Text summarization via hidden Markov models", In the Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval, pp.406-407 (2001)

[28] NTCIR, http://research.nii.ac.jp/ntcir/