DIVERSITY BASED TEXT SUMMARIZATION

Mohammed Salem Binwahlan[1], Naomie Salim[2], Ladda Suanmali[3]

[1,2] Faculty of Computer Science and Information Systems
University Teknologi Malaysia
81300 Skudai, Johor

[3]Faculty of Science and Technology,
Suan Dusit Rajabhat University
295 Rajasrima Rd, Dusit, Bangkok, Thailand 10300

E-mail: [1]moham2007med@yahoo.com, [2]naomie@utm.my, [3]ladda_sua@dusit.ac.th

Abstract: Diversity of selected sentences is an important factor in automatic text summarization to control redundancy in the summarized text. In this paper, we propose a method called maximal marginal importance (MMI) for text summarization based on the idea of the well-known diversity approach maximal marginal relevance (MMR) where an emphasis is on the diversity based binary tree is used to exploit the diversity among the document sentences, where the whole document is clustered into a number of clusters, and then each cluster is presented as one binary tree or more. In our method, the sentence is evaluated based on its importance and its relevance. Our experimental results shown that the proposed method outperforms the three benchmark methods used in this study.

Keywords: summarization, diversity, binary tree, similarity threshold.

1. INTRODUCTION

The automatic text summarization has gained high importance as an active research field in the recent years. The benefits of automatic text summarization system's availability increase the need for existence of such systems; the most important benefits of using a summary is its reduced reading time and providing quick guide to the interesting information.

Diversity, which refers to distinct ideas included in the document, became a very important factor in automatic text summarization to control the redundancy in the summarized text. Many approaches have been proposed for text summarization based on the diversity. For example, MMR (maximal marginal relevance) [1], maximizes marginal relevance in retrieval and summarization. The sentence with high maximal relevance means it is highly relevant to the query and less similar to the already selected sentences. Our modified version of MMR maximizes the marginal importance and minimizes the relevance. This approach treats sentence with high maximal importance as one that has high importance in the document and less relevance to already selected sentences.

MMR has been modified by many researchers [2, 3, 4, 5, 6, 7, 8]. Our modification for MMR formula is similar to [2] and [3] modifications where the importance of the sentence and the sentence relevance are added to the MMR formulation. Ribeiro and Matos [18] proved that the summary generated by MMR method is closed to the human summary, motivating us to choose MMR and modify it by including some documents features. The proposed approach uses a binary tree to exploit the diversity among the document sentences. Neto et al. [14] presented a procedure for creating approximate structure for document sentences in the form of a binary tree, in our study, we build a binary tree for each cluster of document sentences, where the document sentences are clustered using the K-means clustering algorithm into a number of clusters equal to the summary length. An objective of using the binary tree for diversity analysis is to optimize and minimize the text representation; this is achieved by selecting the most representative sentence of each sentences cluster. The redundant sentences are prevented from getting the chance to be candidate sentences for inclusion in the summary, serving as penalty for the most similar sentences. Our idea is similar to Zhu et al.'s idea [9] in terms of improving the diversity where he used absorbing Markov chain walks.

The rest of this paper is described as follows: section 2 presents the features used in this study, section 3 discusses the importance and relevance of the sentence, section 4 introduces the document-sentence tree building process, section 5 gives full description of the proposed method, section 6 discusses the experimental design, section 7 presents the experimental results and section 8 concludes our work and draws the future study plan.

2. SENTENCE FEATURES

The proposed method makes use of eight different surface level features; these features are identified after the preprocessing of the original document is done, like stemming using porter's stemmer[1] and removing stop words. The features are as follows.

a. Word sentence score (WSS): it is calculated using the summation of terms weights (TF-ISF, calculated using eq. 1, [11]) of those terms synthesizing the sentence and occur in at least in a number of sentences equal to half summary length(LS) divided by highest term weights

[1] http://www.tartarus.org/martin/PorterStemmer/

(TF-ISF) summation of a sentence in the document (HTFS) as shown in eq. 2, the idea of making the calculation of word sentence score under the condition of occurrence of its term in specific number of sentences is supported by two factors: excluding the unimportant terms and applying the mutual reinforcement principle [20]. MAN`A-LO`PEZ et al., [15] calculated the sentence score as proportion of the square of the query-word number of a cluster and the total number of words in that cluster.

Term frequency-inverse sentence frequency (TF-ISF) [11], term frequency is very important feature; its first use dates back to fifties [12] and still used.

$$W_{ij} = tf_{ij} \times isf = tf(t_{ij}, s_i)\left[1 - \frac{\log(sf(t_{ij}) + 1)}{\log(n + 1)}\right] \tag{1}$$

Where W_{ij} is the term weight (TF-ISF) of the term t_{ij} in the sentence s_i.

$$\text{WSS}(S_i) = 0.1 + \frac{\sum_{t_j \in S_i} W_{ij}}{HTFS} \quad |\, no.of\ sentences\ containing\ t_j >= \frac{1}{2}LS \tag{2}$$

Where 0.1 is minimum score the sentence gets in the case its terms are not important.

b. Key word feature: the top 10 words whose high TF-ISF (eq. 1) score are chosen as key words [23, 24]. Based on this feature, any sentence in the document is scored by the number of key words it contains; where the sentence receives 0.1 score for each key word.

c. N-friends feature: the n-friends feature measures the relevance degree between each pair of sentences by the number of sentences both are similar to. The friends of any sentence are selected based on the similarity degree and similarity threshold [19].

$$N-friends(s_i, s_j) = \frac{s_i(friends) \cap s_j(friends)}{|s_i(friends) \cup s_j(friends)|} \,|\, i \neq j \tag{3}$$

d. N-grams feature: this feature determines the relevance degree between each pair of sentences based on the number of n-grams they share. The skipped bigrams [25] used for this feature.

$$N-grams(s_i, s_j) = \frac{s_i(n-grams) \cap s_j(n-grams)}{|s_i(n-grams) \cup s_j(n-grams)|} \,|\, i \neq j \tag{4}$$

e. The similarity to first sentence (sim_fsd): This feature is to score the sentence based on its similarity to the first sentence in the document, where in news article, the first sentence in the article is very important sentence [21]. The similarity is calculated using eq. 11.

f. Sentence centrality (SC): the sentence has broad coverage of the sentence set (document) will get high score. Sentence centrality widely used as a feature [19, 16]. We calculate the sentence centrality based on three factors: the similarity, shared friends and shared n-grams between the sentence in hand and all other the document sentences, normalized by n-1, n is the number of sentences in the document.

$$SC(S_i)=\frac{\sum_{j=1}^{n-1} sim(S_i,d(S_j))+\sum_{j=1}^{n-1} n\text{-}friends(S_i,d(S_j))+\sum_{j=1}^{n-1} n\text{-}grams(S_i,d(S_j))}{n-1} \mid i\neq j \; and \; sim(S_i,d(S_j))>=\theta \qquad (5)$$

Where $d(S_j)$ is a document sentence except S_i, n is the number of sentences in the document. θ is the similarity threshold which is determined empirically, in an experiment was run to determine the best similarity threshold value, we have found that the similarity threshold can take two values, 0.03 and 0.16.

The following features are for those sentences containing n-grams [17] (consecutive terms) of title where n=1 in the case of the title contains only one term, n=2 otherwise:

g. Title-help sentence (THS): the sentence containing n-gram terms of title.

$$THS(s_i)=\frac{s_i(n-grams)\bigcap T(n-grams)}{|s_i(n-grams)\bigcup T(n-grams)|} \qquad (6)$$

h. Title-help sentence relevance sentence (THSRS): the sentence containing n-gram terms of any title-help sentence.

$$THSRS(s_j)=\frac{s_j(n-grams)\bigcap THS(s_i(n-grams))}{|s_j(n-grams)\bigcup THS(s_i(n-grams))|} \qquad (7)$$

The sentence score based on THS and THSRS is calculated as average of those two features:

$$SS\_NG=\frac{THS(s_i)+THSRS(s_i)}{2} \qquad (8)$$

3. THE SENTENCE IMPORTANCE(IMPR) AND SENTENCE RELEVANCE(REL)

The sentence importance is the main score in our study; it is calculated as linear combination of the document features. Liu et al. [3] computed the sentence importance also as linear combination of some different features.

$$IMPR(S_i)=avg(WSS(S_i)+SC(S_i)+SS\_NG(S_i)+sim\_fsd(S_i)+kwrd(S_i)) \qquad (9)$$

Where WSS: word sentence score, SC: sentence centrality, SS_NG: average of THS and THSRS features, Sim_fsd: the similarity of the sentence s_i with the first document sentence and $kwrd(S_i)$ is the key word feature.

The sentence relevance between two sentences is calculated in [3] based on degree of the semantic relevance between their concepts, but in our study the sentence relevance between two sentences is calculated based on the shared friends, the shared n-grams and the similarity between those two sentences:

$$\mathrm{Re}l(s_i,s_j)=avg(n-friends(s_i,s_j)+n-grams(s_i,s_j)+sim(s_i,s_j)) \qquad (10)$$

4. DOCUMENT - SENTENCE TREE BUILDING (DST)

The first stage for building the document-sentence tree is to cluster the document sentences into a number of clusters which is determined automatically by the summary length (number of sentences in the final summary), to select the initial centroids, from the sentences list, the sentence with higher number of similar sentences (sentence friends) is selected and form a sentence list for that sentence and its friends, where the number of friends are selected is equal to the total number of document sentences divided by the number of clusters, to calculate the sentence similarity between two sentences $s_i and\ s_j$, we use *TF-ISF* and *cosine* similarity measure as in eq. 11 [19]:

$$sim(s_i,s_j)=\frac{\sum_{w_i \in s_i,s_j} tf(w_i,s_i)tf(w_i,s_j)\left[1-\frac{\log(sf(w_i)+1)}{\log(n+1)}\right]^2}{\sqrt{\sum_{w_i \in s_i}\left(tf(w_i,s_i)\left[1-\frac{\log(sf(w_i)+1)}{\log(n+1)}\right]\right)^2}\times\sqrt{\sum_{w_i \in s_j}\left(tf(w_i,s_j)\left[1-\frac{\log(sf(w_i)+1)}{\log(n+1)}\right]\right)^2}} \tag{11}$$

Where tf is term frequency of term w_i in the sentence $s_i or\ s_j$, sf is number of sentences containing the term w_i in the document, n is number of sentences in the document.

From the sentence list, the highest important sentence is selected as initial centroid and remove all sentences in the sentence list from document sentence list. The next centroid is selected from the remaining sentences in the document sentence list using the same procedure; this process is repeated until the required number of centroids achieved. For each sentences cluster, one binary tree or more is built, the sentence with higher number of friends (higher number of similar sentences) is selected with its friends and they are removed from the sentence cluster, the selected sentence with its friends used to build a binary tree, where the top level in the binary tree contain one sentence which is a sentence has highest score, the score of the sentence in the binary tree building process is calculated based on its importance and its friends number,- this is to balance between the importance and the centrality (a number of high important friends)- as the following:

$$Score_{BT}(s_i)=impr(s_i)+(1-(1-impr(s_i)\times friendsNo(s_i))) \tag{12}$$

Where $Score_{BT}(s_i)$ is the score of the sentence s_i in the binary tree building process, $impr(s_i)$ is importance of the sentence s_i and $friendsNo(s_i)$ is the number of sentence friends. Each level in the binary tree will contain 2^{ln} of the higher score sentences, where ln is the level number, ln=0, 1, 2, ...,n. if there are sentences remaining in the sentences cluster, the same procedure is repeated.

5. METHODOLOGY

The proposed method for summary generation depends on the extraction of the highest important sentences from the original text, we introduce a modified version of MMR, and we called it MMI (maximal marginal importance). MMR approach depends on the relevance of the document to the query, and it is for query based summary, in our modification we have tried to release this restriction by replacing the query relevance with sentence importance for presenting the MMI as generic summarization approach.

Most features used in this method are accumulated together to show the importance of the sentence, the reason for including the importance of the sentence in the method is to emphasize on the high information richness in the sentence as well as high information novelty. We use the tree for grouping the most similar sentences together in easy way, and we assume that the tree structure can take part in finding the diversity.

MMI is used to select one sentence from the binary tree of each sentence cluster to be included in the final summary. In the binary tree, each level sentences get level penalty which is 0.01 times the level number, the purpose of the level penalty is to reduce the noisy sentences score, where the noisy sentences will exist in the low levels where the level penalty is higher, this is to allow the sentence with high importance and high centrality to get the chance to be a summary sentence, this idea is supported by the idea of PageRank used in Google [22] where the citation (link) graph of the web page or backlinks to that page is used to determine the rank of that page. The summary sentence is selected from the binary tree by traversing all levels and applying MMI on each level sentences.

$$MMI(S_i) = Arg \max_{S_i \in CS \setminus SS} \left[(Score_{BT}(S_i) - \beta(S_i)) - \max_{S_j \in SS} (\mathrm{Re}l(S_i, S_j)) \right] \quad (13)$$

Where $\mathrm{Re}l(S_i, S_j)$ is the relevance between the two competitive sentences, S_i is unselected sentence in the current binary tree, S_j is already selected sentence, SS is the list of already selected sentences, CS is the competitive sentences of the current binary tree and β is penalty level.

In MMR, the parameter λ is very important, it controls the similarity between already selected sentences and unselected sentences, and where setting it to incorrect value may cause creation of low quality summary. Our method pays more attention for the redundancy removing by applying MMI in the tree structure-it used for grouping the most similar sentences in one cluster-, so we didn't use the parameter λ because we just select one sentence from each binary tree and leave the other sentences.

Our method is intent to be used for single document summarization as well as multi-documents summarization; it has ability to get rid of the problem of some information stored in single document or multi-documents which inevitably overlap with each other, and can extract globally important information. In addition to that advantage of the proposed method, it maximizes the coverage of each sentence by taking into account the sentence relatedness to all other document sentences. The best sentence based on our method policy is that sentence that has higher importance in the document, higher relatedness to most document sentences and less similar to the sentences already selected as candidates for inclusion in the summary.

6. EXPERIMENTAL DESIGN

The Document Understanding Conference (DUC) [10] data collection became as standard data set for testing any summarization method; it is used by most researchers in text summarization. We have used DUC 2002 data to evaluate our method for creating a generic 100-word summary, the task 1 in DUC 2001 and 2002, for that task, the training set comprised 30 sets of approximately 10 documents each, together with their 100-word human written summaries. The test set comprised 30 unseen documents. A part of this data is used in our experiment which is document set D061.

ROUGE (Recall-Oriented Understudy for Gisting Evaluation) toolkit [25] is used for evaluating the proposed method, where ROUGE compares a system generated summary against a human generated summary to measure the quality. ROUGE is the main metric in the DUC text summarization evaluations. It has different variants, in our experiment, we use ROUGE-N (N=1 and 2) and ROUGE-L, the reason for selecting these measures is what reported by same study [25] that those measures work well for single document summarization.

The ROUGE evaluation measure (version 1.5.5[2]) generates three scores for each summary: recall, precision and F-measure (weighted harmonic mean, eq. 14), in the literature, we found that the recall is the most important measure to be used for comparison purpose, so we will concentrate more on the recall in this evaluation.

$$F=\frac{1}{\left(alpha\times\left(\frac{1}{P}\right)+\left(1-alpha\right)\times\left(\frac{1}{R}\right)\right)} \tag{14}$$

[2] http://haydn.isi.edu/ROUGE/latest.html

Where P, R are precision and recall respectively, alpha is parameter to balance between precision and recall, we set this parameter to 0.5.

7. EXPERIMENTAL RESULTS

The similarity threshold play very important role in our study where the most score of any sentence depends on its relation with other document sentences therefore we must pay more attention to this factor, one experiment is run for this purpose. The data set is used in this experiment is DUC 2001, document set d01a containing eleven documents, each document accompanied with its model or human generated summary. We have experimented with 21 different similarity threshold values ranging from 0.01 to 0.2, 3 by stepping 0.01. We found that the best average recall score can be gotten using the similarity threshold value 0.16 but this value doesn't do well with each document separately, so we have examined each similarity threshold value with each document, we found that the similarity threshold value that can perform well with all documents is 0.03, therefore we decided to run our summarization experiment using the similarity threshold 0.03.

We have run our summarization experiment using DUC 2002 document set D061 which contains two model or human generated summaries for each document, we called those two model summaries H1 and H2, H1 used to evaluate our proposed method summary against it and the human summary H2 used as benchmark to measure the quality of our method summary. Beside the human with human benchmark, we use also two more benchmarks which are baseline (outperformed all systems participated in DUC 2001 and DUC 2002 in creating 100 words summary, [26]) and MS word summarizer, the baseline is the first 100 words from the beginning of the document as determine by DUC 2002.

The proposed method and the three benchmarks are used to create a summary for each document in the document set used in this study, each system created good summary compared with the reference (human) summary, the results using the ROUGE variants (ROUGE-1, ROUGE-2 and ROUGE-L) demonstrate that our method performs better than the three benchmarks. Although the recall score is the main score used for comparing the text summarization methods when the summary length is limited[3], we found that our method outperforms all three benchmarks for all average ROUGE variants scores. The overall analysis for the results is concluded in Table-1 and the MMI average recall at the 95%-confidence interval is shown in Table-2:

[3] http://haydn.isi.edu/ROUGE/latest.html

	ROUGE-1			ROUGE-2			ROUGE-L		
Method	**Avg-R**	**Avg-P**	**Avg-F**	**Avg-R**	**Avg-P**	**Avg-F**	**Avg-R**	**Avg-P**	**Avg-F**
Baseline	0.44008	0.44979	0.44456	0.18023	0.18596	0.18291	0.41241	0.42149	0.41660
MS Word Summarizer	0.43681	0.52798	0.47356	0.21578	0.25889	0.23315	0.40328	0.48754	0.43729
H1-H2	0.47379	0.48641	0.47993	0.17955	0.18494	0.18218	0.44018	0.45230	0.44608
MMI	**0.53484**	**0.55043**	**0.54243**	**0.29655**	**0.30536**	**0.30085**	**0.50129**	**0.51574**	**0.50832**

Table-1: MMI, Baseline, MS Word Summarizer and H1-H2 comparison: Recall, Precision and F-measure using ROUGE-1, ROUGE-2 and ROUGE-L

Metric	95%-confidence interval
ROUGE-1	0.47519 - 0.60689
ROUGE-2	0.20742 - 0.39742
ROUGE-L	0.43929 - 0.57523

Table-2: MMI average recall at the 95%-confidence interval.

For ROUGE-1 average recall score, our method performance is better than the three benchmarks by: 0.06105, 0.09803 and 0.09476 for H1-H2, MS word summarizer and baseline respectively. For ROUGE-2 average recall score, our method performance is better than the three benchmarks by: 0.117, 0.08077 and 0.11632 for H1-H2, MS word summarizer and baseline respectively. For ROUGE-L average recall score, our method performance is better than the three benchmarks by: 0.06111, 0.09801 and 0.08888 for H1-H2, MS word summarizer and baseline respectively. The results obtained demonstrated that our proposed method - despite its simplicity where it doesn't make use of any deep natural language processing – it is effective in creating extracts.

8. CONCLUSION AND FUTURE WORK

In this paper we have presented an effective diversity based method for single document summarization, two ways were used for finding the diversity: the first one is as preliminary way where the document sentences are clustered based on the similarity- similarity threshold is 0.03 determined empirically- and all resulted clusters are presented as a tree containing a binary tree for each group of similar sentences. The second way is to apply the proposed method on each branch in the tree to select one sentence as summary sentence. The introduced method has advantages such simplicity, it doesn't use external resource except the original document given to be summarized and deep natural language processing is not required. Our method has shown good performance comparing with the benchmark methods used in this study. For future work, our research is still going on to extend the proposed method for multi document summarization and using a large data set.

REFERENCES

[1] Carbonell Jaime and Jade Goldstein, "The Use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries". In the proceedingsof the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'98), 1998.

[2] Mori T., Nozawa M., and Asada Y., "Multi-Answer-Focused Multi-document Summarization Using a Question-Answering Engine". ACM Transactions on Asian Language Information Processing, Vol. 4, No. 3, September 2005, Pages 305–320.

[3] Liu D., Wang Y., Liu C., and Wang Z., "Multiple Documents Summarization Based on Genetic Algorithm". Springer-Verlag Berlin Heidelberg 2006, pp. 355–364.

[4] Zajic David M., Dorr Bonnie J., Schwartz Richard and Lin Jimmy, "Sentence Compression as a Component of a Multi-Document Summarization System". In Proceedings of the 2006 Document Understanding Workshop, New York.

[5] Filippova Katja, Mieskes Margot, Nastase Vivi, Ponzetto Simone Paolo, and Strube Michael, "Cascaded Filtering for Topic-Driven Multi-Document Summarization". In Proceedings of the Document Understanding Conference. Rochester, N.Y., 26-27 April 2007, pages 30-35.

[6] Ye Shiren, Qiu Long, Chua Tat-Seng, and Kan Min-Ye, "NUS at DUC 2005: Understanding documents via concept links". In Document Understanding Conference.

[7] Lin Ziheng, Chua Tat-Seng, Kan Min-Yen, Lee Wee Sun, Qiu Long and Ye Shiren, " NUS at DUC 2007: Using Evolutionary Models of Text".

[8] Kraaij W., Spitters M., and Heijden M. van der, "Combining a mixture language model and naive bayes for multi-document summarization". In Notebook papers of the SIGIR 2001/ DUC 2001 workshop.

[9] Zhu Xiaojin, Goldberg Andrew B, Gael Jurgen Van, and Andrzejewski David, "Improving diversity in ranking using absorbing random walks". HLT/NAACL, 2007.

[10] The Document Understanding Conference (DUC). http://duc.nist.gov

[11] Neto J. L., Santos A. D., Kaestner C. A. A., Freitas, A. A., "Document Clustering and Text Summarization". In: Proc. of the 4th International Conference on Practical Applications of Knowledge Discovery and Data Mining, 2000.

[12] Luhn, H., "The automatic creation of literature abstracts". IBM Journal of Research and Development 2(92) (1958) 159-165.

[13] Murray Gabriel, Renals Steve, and Carletta Jean,. "Extractive summarization of meeting recordings' .In Proceedings of the 9th European Conference on Speech Communication and Technology (EUROSPEECH '05), Lisbon, Portugal, 4-8 September 2005.

[14] Neto Joel Larocca, Freitas Alex A., and Kaestner Celso A. A., "Automatic Text Summarization using a Machine Learning Approach". In Proc. 16th Brazilian Symp. On Artificial Intelligence (SBIA-2002). Lecture Notes in Artificial Intelligence 2507, 2002, pp205-215. Springer-Verlag.

[15] MAN`A-LO`PEZ Manuel J, and ´ OMEZ-HIDALGO JOS ´E M. G, "Multi-document Summarization: An Added Value to Clustering in Interactive Retrieval". New York: ACM Transactions on Information Systems, Vol. 22. (2004) 215–241.

[16] Zajic David M., "Multiple Alternative Sentence Compressions As A Tool For Automatic Summarization Tasks", University of Maryland, PhD theses, 2007.

[17] Villatoro-Tello Esaú, Villaseñor-Pineda Luis and Montes-y-Gómez Manuel, "Using Word Sequences for Text Summarization". Petr Sojka, Ivan Kopeˇcek and Karel Pala (Eds.): TSD 2006, LNAI 4188, pp. 293–300. Springer-Verlag Berlin Heidelberg.

[18] Ribeiro Ricardo and Matos David Martins de., "Extractive Summarization of Broadcast News: Comparing Strategies for European Portuguese". V. Matouˇsek and P. Mautner (Eds.): TSD 2007, LNAI 4629, pp. 115–122. Springer-Verlag Berlin Heidelberg.

[19] Erkan G., Radev D. R., "LexRank: Graph-based Lexical Centrality as Salience in Text Summarization". Journal of Artificial Intelligence Research (JAIR), July 2004.

[20] Zha Hongyuan, "Generic summarization and key phrase extraction using mutual reinforcement principle and sentence clustering". In proceedings of 25th ACM SIGIR, 2002, pp113-120.

[21] Ganapathiraju, Madhavi K., "Relevance of Cluster size in MMR based Summarizer": A Report 11-742: Self-paced lab in Information Retrieval, November 26, 2002.

[22] Brin S. and Page L., "The anatomy of a large-scale hypertextual Web search engine". Computer Networks and ISDN Systems, 1998, 30(1–7):107–117.

[23] Jaruskulchai Chuleerat and Kruengkrai Canasai, "Generic Text Summarization Using Local and Global Properties" Proceedings of the IEEE/WIC international Conference on Web Intelligence, 2003.

[24] Kiani -B Arman and Akbarzadeh –T M. R., "Automatic Text Summarization Using: Hybrid Fuzzy GA-GP". 2006 IEEE International Conference on Fuzzy Systems, Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada July 16-21, 2006.

[25] Lin C. Y., "Rouge: A package for automatic evaluation of summaries". In WAS 2004: Proceedings of the Workshop on Text Summarization Branches Out, July 25–26, 2004, Barcelona, Spain.

[26] Nenkova Ani "Automatic text summarization of newswire: Lessons learned from the document understanding conference". In Proceedings of American Association for Artificial Intelligence, Pittsburgh, USA, 2005.