RIN-Sum: A System for Query-Specific Multi-Document Extractive Summarization

Rajesh Wadhvani
CSE
National Institute of Technology
Bhopal, India

Manasi Gyanchandani
CSE
National Institute of Technology
Bhopal, India

Rajesh Kumar Pateriya
CSE
National Institute of Technology
Bhopal, India

Sanyam Shukla
CSE
National Institute of Technology
Bhopal, India

***Abstract*—In paper, we have proposed a novel summarization framework to generate a quality summary by extracting Relevant-Informative-Novel (RIN) sentences from topically related document collection called as RIN-Sum. In the proposed framework, with the aim to retrieve user's relevant informative sentences conveying novel information, ranking of structured sentences has been carried out. For sentence ranking, Relevant-Informative-Novelty (RIN) ranking function is formulated in which three factors, i.e., the relevance of sentence with input query, informativeness of the sentence and the novelty of the sentence have been considered. For relevance measure instead of incorporating existing metrics, i.e., Cosine and Overlap which have certain limitations, a new relevant metric called as C-Overlap has been formulated. RIN ranking is applied on document collection to retrieve relevant sentences conveying significant and novel information about the query. These retrieved sentences are used to generate query-specific summary of multiple documents. The performance of proposed framework have been investigated using standard dataset, i.e., DUC2007 documents collection and summary evaluation tool, i.e., ROUGE.**

Keywords*—Text summarization; maximum marginal relevance; sentence selection; DUC2007 data collection

I. Introduction

The notion of information retrieval is to locate documents that might contain the relevant information. Generally, when a user fires a query, his desire is to locate relevant information rather than locate a ranked list of documents. The retrieved documents contain the relevant information leaving the user with a massive amount of text. There is a requirement of a tool that shrinks this amount of text in order to comprehend the complete text [1]. The query focused summarization track at Document Understanding Conference (DUC) aims at doing exactly this. Conventional query focused text summarization systems rank and assimilate sentences based on maximizing relevance to the user's information need expressed via query [2]. These systems do not consider the important factor, i.e., informativeness and novelty of the sentence. In this paper, a novel summarization framework to generate a quality summary by extracting Relevant-Informative-Novel (RIN) sentences from topically related document collection called as RIN-Sum has been presented. This framework generates a query focused summary of multiple documents by using three factors, namely: sentence relevance with input query (discussed in section 2), sentence informativeness (discussed in section 3) and sentence novelty (discussed in section 4). In this work, ordering of these factors has been considered to rank the sentences. Firstly, relevance with input query is applied, and then sentence informativeness, and finally sentence novelty. For example, if a sentence is novel and highly informative in the document collection, but if it is not relevant to a user's query, it will not be considered for a final summary.

II. The Relevance Measure

Relevance measures can be divided into two types based on whether the ordering of vectors is taken into account, i.e., symmetric and asymmetric [3] [4]. For two sentence vectors S_i and S_j, a symmetric measure yields the same result regardless of the ordering of the sentence vectors, i.e., $Sim(S_i,S_j) = Sim(S_j,S_i)$. An asymmetric measure yields different results for different orderings of two sentence vectors, i.e., $Sim(S_i,S_j) \neq Sim(S_j,S_i)$. The Cosine measure is the most popular symmetric measure based on VSM for checking the extent of similarity between two texts. In VSM for text summarization, the sentence is usually presented as a vector of weighted terms. Cosine similarity between two weighted sentences $S_i = [w_{1i},\ldots,w_{ni}]$ and $S_j = [w_{1j},\ldots,w_{nj}]$ can be define as:

$$Sim_{cos}(s_i,s_j) = \frac{vec(s_i).vec(s_j)}{|\ s_i\ |\ \ |\ s_j\ |} = \frac{\sum_{k=1}^{n} w_{ki} \times w_{kj}}{\sqrt{\sum_{k=1}^{n} w_{ki}^2} \times \sqrt{\sum_{k=1}^{n} w_{kj}^2}} \quad (1)$$

In Cosine measure, two sentence vectors S_i and S_j are compared on the basis of all terms which appear in S_i and/or S_j. In both sentences discriminative power of each term is well defined. Discriminative power of uncommon terms between S_i and S_j also affects the similarity measure. Hence this type of similarity measure performs well when two texts are compared on the basis of a set of terms appearing in either first text and/or second text.

The Overlap measure is the asymmetric relevance measure between two texts. It is a relative measure to detect similarity or overlap among texts by making comparison between the current text and any other text with respect to all those terms which appears only in current text. The Overlap measure is computed by comparing the current sentence S_i with any sentence S_j, as define in [5] is given in (2):

$$Sim_{overlaps}(s_i, s_j) = \frac{\sum_{k=1}^{m} w_{ki} \times w_{kj}}{\sum_{k=1}^{m} w_{ki}^2} \quad (2)$$

This metric is a relative measure to detect similarity or overlap among sentences. This mechanism works on the comparison of the relative frequency of the words representing a sentence. One of the limitations with (2) is that it does not compare two sentences irrespective of their sizes. This causes problem in a situation when for a given common term weight in S_j dominates over weight in S_i, resulting in increase of the overlap score in the proportion to differences in their weights. In case of Cosine measure there is no such limitation. For getting the advantages of overlap measure, there is a need to improve (2) over this limitation. Proposed improvement over this metric has been formulated in (3).

$$Sim_{overlaps}(s_i, s_j) = \frac{\sum_{k=1}^{m} [min(w_{ki}, w_{kj})]^2}{\sum_{k=1}^{m} w_{ki}^2} \quad (3)$$

The above mentioned metrics for Overlap measure is used to determine whether sentences are copies of one another or not. One limitation with this metric is that it does not consider the discriminative power of the terms. In next section Overlap based Cosine measure is formulated for identifying all those sentences in which each term of current text appears with high discriminative power.

At the time of sentence extraction, applying Overlap measure technique as relevance measure returns a set of sentences without considering the discriminative power of query terms of those sentences. Here the use of Cosine measure may improve the match quality by considering the discriminative power of query terms in sentence ranking, but at the same time ranking of the sentence are declined in its non-query terms. Matching quality can be improved by adding the properties of Overlap in Cosine measures. In this respect, proposed methodology has been formulated called as Overlapped based Cosine measure which can be abbreviated as C-Overlap measure. In this formulation, at its first step, the terms appear in sentence S_j are decomposed into two groups having common and uncommon terms with respect to S_i. After decomposing sentence S_j into two groups, $S_j' = [w_{1j}, \ldots, w_{mj}]$ and $S_j'' = [w_{(m+1)j}, \ldots, w_{nj}]$ are obtained, where $S_j = S_j' \cup S_j''$. Overlap between S_i and S_j is nothing but cosine similarity between S_i and S_j'. Cosine similarity between two weighted sentences $S_i = [w_{1i}, \ldots, w_{mi}]$ and $S_j' = [w_{1j}, \ldots, w_{mj}]$, can be formulated as follows:

$$Sim_{Coverlap}(s_i, s_j) = Sim_{cos}(s_i, s_j') = \frac{\sum_{k=1}^{m} w_{ki} \times w_{kj}}{\sqrt{\sum_{k=1}^{m} w_{ki}^2} \sqrt{\sum_{k=1}^{m} w_{kj}^2}} \quad (4)$$

$$Sim_{Coverlap}(s_i, s_j) = \frac{vec(s_i).vec(s_j)}{\|s_i\| \sqrt{\sum_{k=1}^{m} w_{kj}^2}} \quad (5)$$

Here in normalization process of vector S_j, uncommon terms are neglected. As a result strength of common terms increases. Hence sentences will be ranked on the basis of discriminative query terms only.

III. The Informativeness Measure

Cosine, Overlap and C-Overlap all are pure relevance measurement techniques which do not consider the sentence informativeness. A ranking metric is required which improves the rank of relevant sentences on the basis of informativeness of the sentence. In this section, a ranking function which measures the informativeness score of the given sentence based on assumed hypothesis, i.e., "within a query relevant sentence, its non-query terms may convey information about the query terms" is formulated, which is defined as follows:

$$informative_{t \in S_i \setminus Q}(S_i) = \sqrt{\sum_{t \in S_i \setminus Q} w_{ki}^2} \quad (6)$$

Here informativeness of sentences S_i is measured by considering the weights of non-query terms only. A score of informativeness of the sentence is equal to L2 norm or Euclidean norm of the weights of discriminative non query terms. In this work, instead of preferring large number of low discriminative terms, small numbers of high discriminative terms are considered. Therefore, L2 norm is preferred over L1 norm as the L1 norm focuses on total weights while L2 norm considers the distribution of weights. Further, the value of the score may be greater than one and to use it with other scores it need to be normalized in the range of 0 to 1 for all sentences in the document collection. To normalize this score, initially score of informativeness for all sentences in document collection is calculated and then maximum score between them is found as:

$$Maxscore = \max[informative_{t \in S_i \setminus Q}(S_1), informative_{t \in S_i \setminus Q} \quad (7)$$

Now to obtain the normalized score, score of each sentence is divided with Maxscore and can be written as:

$$informative_{t \in S_i \setminus Q}(S_i) = \frac{\sqrt{\sum_{t \in S_i \setminus Q} w_{ki}^2}}{Maxscore} \quad (8)$$

Besides this, a ranking function for informativeness is used to formulate sentence informativeness based relevant metric.

This approach measures relevance and informativeness of the sentence separately and then uses a linear combination of the two to produce a single score for the ranking of a sentence. The informativeness based relevant metric can be formulated as:

$$Sim_{Relevance-informative}(Q,S_i) = \beta\, Sim_{relevance}(Q,S_i) + (1-\beta)\, informative_{t\in S_i \setminus Q}\,(S_i) \quad (9)$$

In this metric any one of Cosine, Overlap and C-Overlap can be used for relevance measurement. β is tuning factor and its theoretical value lies between 0 to 1. A sentence of our interest is primarily relevant to user query and then informative. To accomplish this in (9) relevant metric should get more weight as compare to informative metric. So practically, value of β should be close to one.

IV. The Novelty Measure

In automatic text summarization, precision of results will be increased by being very selective about the sentences and retaining only those in summary that are considered to be surely relevant. Therefore, necessary condition to retain a sentence in the summary is its relevance with input query. Along with precision a good coverage is required for improving recall, but at the same time another constrains with summary is that it is bounded in length [6]. Optimizing these three constrains, namely: relevance, coverage, and summary length is a challenging task. One of the solutions to maximize the coverage of summary by confirming its length is trying to include those relevant sentences which are novel to the sentences already retained in summary.

A. Maximum Marginal Relevance (MMR)

Carbonell et al. [7] encouraged Maximal Marginal Relevance (MMR) which considers novelty along with relevance to rank the text. Using this technique, partial or full duplicate information is prevented from being retrieved. In particular, MMR has been widely used in text summarization because of its simplicity and effectiveness, and it has shown a consistently good performance. MMR uses the Retrieval Status Value (RSV) as a parameter to measure the diversity among the sentences. The RSV value of the newly retrieved sentence is decided by sentences which have been already retrieved. It prevents the similar sentences by lowering their RSV value and as a result, it boosts up dissimilar sentences. The final score of given sentence S_i is calculated as follows:

$$MMR(R,S) = \underset{S_i \in R\setminus S}{argmax}\left[\lambda\{Sim_1(Q,S_i)\} - (1-\lambda)\left\{\underset{S_j \in S}{max} Sim_2\,(S_i,S_j)\right\}\right] \quad (10)$$

Where R stands for the ranked list of sentences, S represents the sentences that have been extracted into the summary, Q denotes the query and S_i indicates a sentence. Sim_1 and Sim_2 are similarity measures, which can either be same or different. Different similarity measures have been explored in next session. λ is tuning factor which lies between 0 to 1.

In this approach, summaries are created using greedy sentence-by-sentence selection. At each selection step, the greedy algorithm is constrained to select the sentence that is maximally relevant to the user query and minimally redundant with sentences which have been already included in the summary. MMR measures relevance and novelty separately and then uses a linear combination of the two to produce a single score for the importance of a sentence in a given stage of the selection process. Xie et al. [8], Forst et al. [9] and Chowdary et al. [10] encouraged the concept of "relevant novelty", which claim that a sentence of input text will be retained in a summary if it is relevant to the user and should not convey the information which is already covered by the current summary sentences.

B. Relevant-Informative-Novelty (RIN) metric for sentence selection

Relevance, informativeness and novelty are the three basic measures which have been considered in the ranking during sentence extraction. Considering only relevance measure for generating the summary does not give the guarantee of novelty in the summary. In this section, a ranking metric is formulated which improves the rank of relevant and informative sentences based on their diversity with other sentences. In this formulation, MMR has been used. A ranking function which measures a novelty score of the given sentence with respect to current summary sentences is formulated. This formulation is based on the following assumptions:

- Those sentences in the current summary are put under considerations which are diverse on the basis of conveyed information.
- Sentences are retained in the current summary if they convey novel information about the query.
- Within a query relevant sentence, its non-query terms may convey information about the query terms.

Thus, novelty of given sentence with respect to current summary sentence can be measured in term of amount of overlap between non query term of given sentence S_i and current summary sentence S_j. This can be calculated as follows:

$$Sim_{novelty}(S_i,S_j) = Sim_{overlaps}\,(S_i^{\prime},S_j) : if\ k \in S_i \cap Q\ then\ w_{ki}^{\prime} = 0\ else\ w_{ki}^{\prime} = w_{ki}\ , \quad where\ w_{ki}^{\prime}\ is\ k^{th}\ term\ in\ S_i^{\prime} \quad (11)$$

Now using linear combination of relevant and novelty metric final score is obtained. Relevant-novelty metric can be given as:

$$Score(R,S) = \underset{S_i \in R\setminus S}{argmax}\left[\lambda\{Sim_{Relevant}(Q,S_i)\} - (1-\lambda)\left\{\underset{S_j \in S}{max} Sim_{novelty}\,(S_i,S_j)\right\}\right] \quad (12)$$

In case, when informativeness of the sentence is considered, RIN metric can be given as:

$$Score(R,S) = \underset{S_i \in R \setminus S}{argmax} \left[\lambda\{Sim_{Relavant-informative}(Q,S_i)\} - (1-\lambda)\left\{\underset{S_j \in S}{max} Sim_{novelty}(S_i,S_j)\right\}\right] \quad (13)$$

In this metric, novelty of sentence S_i is measured in terms of amount of overlap between non query term of given sentence S_i and current summary sentence S_j. λ is tuning factor and its theoretical value lies between 0 to 1. More weight is given to informativeness based relevant metric because a sentence is significant if primarily relevant to the user query then it should be informative and finally it should be a novel.

V. RIN-SUM METHODOLOGY

To provide the methodology of sentence extraction from unstructured text to generate its query-specific summary, following are the steps that RIN-Sum takes to construct query-specific summary of multiple documents.

1) Select a query and set of associated documents for which summary is to be generated. These documents and the query constitute the input to RIN-Sum.

2) Each document in the collection is analyzed to obtained its structured representation using following steps:

- Firstly, each document is pre-processed to generate sentence set.
- Each sentence in the resultant set is represented by vector in dimensions of content terms of pre-processed document.
- Each sentence vector is weighted for content terms.

3) Finally a cluster of unstructured sentences is generated as a final summary by extracting salient and non-redundant sentences from given document collection. This process consists of following steps:

- Firstly, Sentence vectors are ranked by applying proposed C-Overlap measure based relevant metric to produces a cluster of relevant sentences.
- Resultant cluster sentence vectors are again ranked through proposed Relevant-Informative metric to produces a cluster of relevant and informative sentences.
- Finally, to retrieve sentences conveying novel information about query from group of identified relevant-informative sentences, a Relevant-Informative-Novelty (RIN) ranking function is used.

4) Further the performance of the proposed framework has been investigated using standard dataset, i.e., DUC2007 documents collection and summary evaluation tool, i.e., ROUGE, and simulation strategy of proposed methodology and analysis of results have been performed.

Thus RIN-Sum uses topically related documents to produce a summary. These summaries are deemed relevant to a user query. For example, to satisfy the user's information need about given topically related documents collection, a summary which contains user's intended information on that topic will be generated.

VI. EXPERIMENTS

DUC2007 dataset has been used for evaluation and it is available through [11] on request. A total of 45 documents were constructed by NIST assessors based on topics of interest and for each topic four reference summaries were produced by human experts to create gold collection for evaluation purposes. For performance evaluation ROUGE-1, ROUGE-2 and ROUGE-SU metrics of ROUGE-1.5.5 package [12] has been used. ROUGE-1 compares the unigram overlap between the candidate summary and the reference summaries. ROUGE-2 compares the bigram overlap between the candidate summary and the reference summaries. ROUGE-SU is an extended version of ROUGE-2 that match skip bigrams, with skip distance up to 4 words. Performance is measured in terms of Recall, Precision and F-score. Several experiments have been conducted in which for text representation the standard sequence of steps have been followed, which are:

1) Generate sentence set by separating sentences of DUC2007 document collection.

2) Remove functional and grammatical words of the sentences using stop word list, provided with DUC document collection.

3) For each sentence, apply stemming algorithm on each word of the sentence with the help of well-known Porter Stemmer [13] in order to find related words.

4) Calculate weight of each word within the sentence using standard tf.idf weighting scheme [14].

As an output of the above steps, sentences of each document are represented as sentence-terms weighted vector. Now using sentence ranking function as formulated in (13), sentences are ranked and then extracted to get a final summary. With ranking function, different experiments are performed for different relevance measure i.e. Cosine, overlap, C-Overlap. For informativeness and novelty measure, fixed measure as defined in (8) and (11) respectively are used. Experiments were performed in three different phases. In each phase four different ranking functions were used which are:

- Relevant Metric
- Relevant-Informative Metric
- Relevant-Novelty Metric
- Relevant-Informative-Novelty(RIN) Metric

Here results were obtained for different ROUGE metrics in term of Precision, Recall and F-Score.

Phase I: In this phase results were obtained for above four ranking functions. In these experiments Cosine measure was used as relevant metric and for informativeness and novelty measure fixed metrics was used as defined in (8) and (11) respectively. The results are as shown in Table (1).

TABLE. I. EVALUATION RESULTS USING COSINE MEASURE BASED (A) RELEVANT RANKING FUNCTION; (B) RELEVANT-INFORMATIVE RANKING FUNCTION; (C) RELEVANT NOVELTY RANKING FUNCTION AND (D) RELEVANT-INFORMATIVE-NOVELTY RANKING FUNCTION

a

	Recall	Precision	F-score
ROUGE-1	0.42037	0.38742	0.40248
ROUGE-2	0.10729	0.10046	0.10369
ROUGE-SU	0.16919	0.14283	0.15391

b

	Recall	Precision	F-score
ROUGE-1	0.42203	0.39064	0.40494
ROUGE-2	0.10789	0.10155	0.10455
ROUGE-SU	0.16997	0.14686	0.15682

c

	Recall	Precision	F-score
ROUGE-1	0.42638	0.39512	0.40938
ROUGE-2	0.10775	0.10073	0.10397
ROUGE-SU	0.17403	0.14873	0.15935

d

	Recall	Precision	F-score
ROUGE-1	0.43557	0.40243	0.41786
ROUGE-2	0.11719	0.10980	0.11329
ROUGE-SU	0.18374	0.15546	0.16745

Phase II: In this phase results were obtained for above four ranking functions. In these experiments overlap measure was used as relevant metric and for informativeness and novelty measure fixed metrics was used as defined in (8) and (11) respectively. The results are as shown in Table (2).

TABLE. II. EVALUATION RESULTS USING OVERLAP MEASURE BASED (A) RELEVANT RANKING FUNCTION; (B) RELEVANT-INFORMATIVE RANKING FUNCTION; (C) RELEVANT NOVELTY RANKING FUNCTION AND (D) RELEVANT-INFORMATIVE-NOVELTY RANKING FUNCTION

a

	Recall	Precision	F-score
ROUGE-1	0.43321	0.40360	0.41715
ROUGE-2	0.11283	0.10536	0.10876
ROUGE-SU	0.17934	0.15429	0.16467

b

	Recall	Precision	F-score
ROUGE-1	0.44941	0.40532	0.42548
ROUGE-2	0.12231	0.11124	0.11635
ROUGE-SU	0.18958	0.15488	0.16949

c

	Recall	Precision	F-score
ROUGE-1	0.44848	0.41063	0.42826
ROUGE-2	0.12129	0.11273	0.11677
ROUGE-SU	0.19158	0.15975	0.17334

d

	Recall	Precision	F-score
ROUGE-1	0.45678	0.41583	0.43449
ROUGE-2	0.12971	0.12109	0.12511
ROUGE-SU	0.19761	0.16549	0.17886

***Phase III*:** In this phase results were obtained for above four ranking functions. In these experiments C-Overlap measure was used as relevant metric and for informativeness and novelty measure fixed metrics was used as defined in (8) and (11) respectively. The results are as shown in Table (3).

TABLE. III. EVALUATION RESULTS USING C-OVERLAP MEASURE BASED (A) RELEVANT RANKING FUNCTION; (B) RELEVANT-INFORMATIVE RANKING FUNCTION; (C) RELEVANT NOVELTY RANKING FUNCTION AND (D) RELEVANT-INFORMATIVE-NOVELTY RANKING FUNCTION

a

	Recall	Precision	F-score
ROUGE-1	0.44777	0.40909	0.42666
ROUGE-2	0.12581	0.11597	0.12048
ROUGE-SU	0.19081	0.16002	0.17262

b

	Recall	Precision	F-score
ROUGE-1	0.45646	0.41197	0.43217
ROUGE-2	0.13073	0.12003	0.12493
ROUGE-SU	0.19883	0.16304	0.17800

c

	Recall	Precision	F-score
ROUGE-1	0.45935	0.41621	0.43597
ROUGE-2	0.13176	0.12197	0.12643
ROUGE-SU	0.20092	0.16575	0.18005

d

	Recall	Precision	F-score
ROUGE-1	0.46487	0.41969	0.44042
ROUGE-2	0.13568	0.12304	0.12879
ROUGE-SU	0.20821	0.16813	0.18449

Graphically the F-scores (ROUGE-1, ROUGE-2 and ROUGE-SU) results are depicted in figures (1) - (3) respectively.

In these figures, while observing the curves of Relevant and Relevant-Informative Ranking, it can be concluded that in all cases, i.e., Cosine, Overlap and C-Overlap the performance of Relevant-Informative Ranking is better as compared to Relevant Ranking. While observing the curves of Relevant and Relevant-informative-Novelty Ranking it can be concluded that in all cases, i.e., Cosine, Overlap and C-Overlap the performance of Relevant-Informative-Novelty Ranking is better as compared to Relevant Ranking.

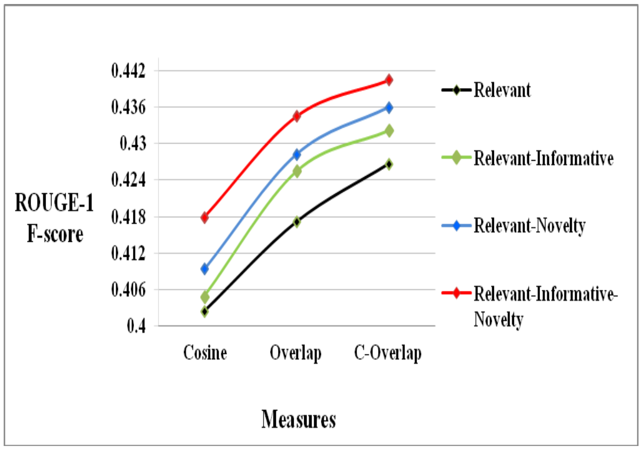

Fig. 1. ROUGE-1 F-score results comparison for Relevant, Relevant-Informative, Relevant-Novelty and Relevant-Informative-Novelty metrics over Cosine, Overlap and C-Overlap measures

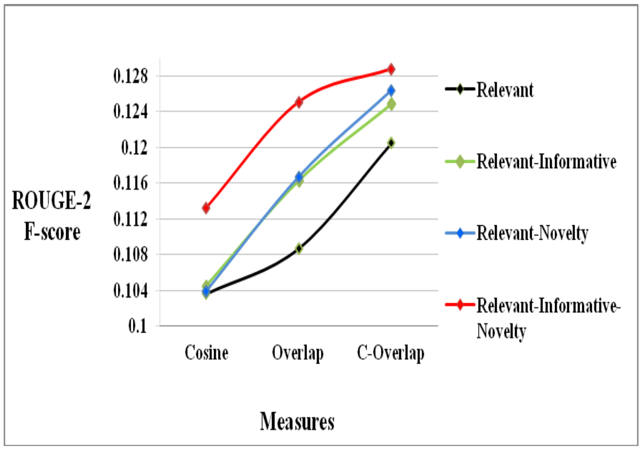

Fig. 2. ROUGE-2 F-score results comparison for Relevant, Relevant-Informative, Relevant-Novelty and Relevant-Informative-Novelty metrics over Cosine, Overlap and C-Overlap measures

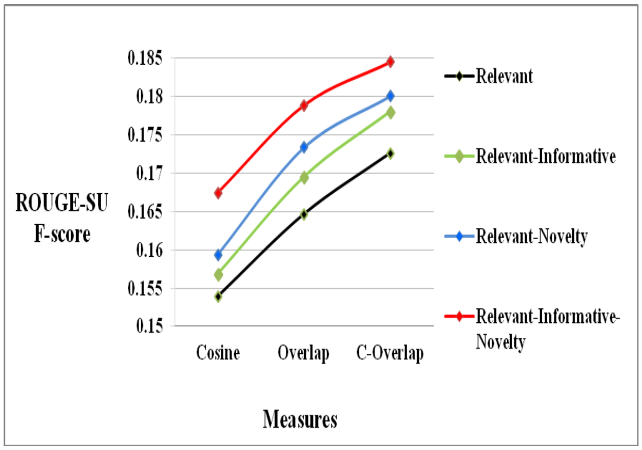

Fig. 3. ROUGE-SU F-score results comparison for Relevant, Relevant-Informative, Relevant-Novelty and Relevant-Informative-Novelty metrics over Cosine, Overlap and C-Overlap measures

Justification for this improvement is that ranking of the sentence based on proposed C-Overlap relevance measure does not consider the significance of non-query terms. When Informative Metric is applied in sentence ranking, it considers the significance of non-query terms also. As a result, this technique tries to retrieve all sentences having significant query terms as well as significant non-query terms. Also, in sentence ranking when Novelty Metric is applied, it prevents the retrieval of partial or full duplicate information and improves the coverage of bounded length summary. As a result, performance in terms of recall value increases.

VII. Conclusion

In this paper, a novel technique to query specific extractive text summarization for multiple documents has been presented. The utility of the approach is examined on DUC2007 dataset collection. In the proposed method, with the aim to retrieve user's relevant significant sentences conveying novel information, ranking of structured sentences has been carried out. A new method of sentence ranking has been developed which identifies the relevant, significant and novel sentences from a large volume of input text. To achieve this, RIN metric is formulated for sentence ranking depending on three factors, i.e., the relevance of sentences with input query, informativeness of the sentence as well as the novelty of the sentence. For relevance measurement, a new measure formally known as C-Overlap (Overlapped based Cosine measure) has been proposed with the aim to overcome the limitations of existing relevance measures, i.e., Cosine and Overlap measure. Experimentally it has been proved that C-Overlap measure outperformed the previous ones.

Finally, sentences in document collection were extracted using RIN ranking metric. Results were compared with the other standard sentence ranking functions, i.e., Relevant, Relevant-Informative, Relevant-Novelty and Relevant-Informative-Novelty, using ROUGE-1.5.5. It has been observed that in each case results of proposed function are found to be better as compared to other three ranking functions. Experimentally, it is also observed that Relevance alone is not a good choice as a ranking function.

References

[1] Mani I and Maybury M T 1999 Advances in Automatic Text Summarization. MIT Press, Cambridge

[2] Kumar Y J and Salim N 2011 Automatic Multi Document Summarization Approaches. In: Journal of Computer Science, Volume 8, Issue 1, pp: 133–140

[3] Tsai S F S, Tang W and Chan K L 2010 Evaluation of novelty metrics for sentence-level novelty mining. In: Information Sciences, Volume 180, Number 12,pp: 2359–2374

[4] Zhang Y, Tsai F S and Kwee A T 2011 Multilingual sentence categorization and novelty mining. In: Information Processing and Management, Volume 47, pp: 667–675

[5] Alguliev R M, Aliguliyev R M and Isazade N R 2013 MR&MR-SUM: Maximum Relevance and Minimum Redundancy Document Summarization Model. In: World Scientific Publishing Company, International Journal of Information Technology and Decision Making, Volume 12, Number 3, pp: 361–394

[6] Alguliev R M, Aliguliyev R M, Hajirahimova M S and Mehdiyev C A 2011 MCMR: Maximum coverage and minimum redundant text summarization model. In: Elsevier, Expert Systems with Applications, Volume 38, pp: 14514-14522.

[7] Carbonell J and Goldstein J 1998 The use of MMR, diversity-based Reranking for reordering documents and producing summaries. In Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp: 335–336

[8] Xie S and Liu Y 2008 Using corpus and knowledge-based similarity measure in Maximum Marginal Relevance for meeting summarization. In IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp: 4985-4988

[9] Forst J F, Tombros A and Roelleke T 2009 Less is More: Maximal Marginal Relevance as a Summarization Feature. In Proceedings of the 2nd International Conference on the Theory of Information Retrieval, Lecture Notes in Computer Science, Volume 5766, pp: 350–353

[10] Guohua WU and Yutian G 2016 Using Density Peaks Sentence Clustering For Update Summary Generation. In Proceedings of 2016 IEEE Canadian Conference on Electrical and Computer Engineering.

[11] Document Understanding Conference: <http://duc.nist.gov>.

[12] Lin C-Y 2004 ROUGE: A package for automatic evaluation summaries. In Proceedings of the ACL Text Summarization Branches Out Workshop, Barcelona, Spain, pp: 74–81

[13] Porter M 2006 The Porter Stemming Algorithm, Official home page for distribution of the Porter Stemming Algorithm. <http://tartarus.org/~martin/PorterStemmer/ index.html>

[14] Polettini N 2004 The Vector Space Model in Information Retrieval Term Weighting Problem. http://sra.itc.it/people/polettini/PAPERS/ Polettini Information Retrieval.pdf