ISSN : 0974 - 7435 Volume 10 Issue 24

2014

BioTechnology
An Indian Journal

FULL PAPER
BTAIJ, 10(24), 2014 [15031-15035]

MR^{2}P: A mutually reinforced relevance propagation model for query-focused multi-document summarization

Jicheng Wei, Libin Yang*, Shuqin Li, Xiaoyan Cai, Shengzhe Wang
College of Information Engineering, Northwest A&F University, YangLing, Shaanxi, 712100, (CHINA)
E-mail : libiny@nwsuaf.edu.cn

ABSTRACT

Query-focused multi-document summarization aims to create a compressed summary biased to a given query. This paper presents a mutually reinforced relevance propagation (MR^{2}P) approach to this summarization task. Experiments are conducted on the DUC 2005 and DUC 2006 data sets and the ROUGE evaluation results demonstrate the advantages of the proposed approach.

KEYWORDS

Manifold ranking; Mutual reinforcement; Query-focused multi-document summarization; Relevance propagation.

INTRODUCTION

Recently, manifold-ranking has been exploited for query-focused summarization [1]. The manifold-ranking based summarization approach constructs a weighted graph that explicitly represents both query and sentences as vertices. The pre-specified positive ranking score of query is then propagated to nearby vertices via the graph iteratively until a global stable state is achieved. At the end, all the sentences are ranked according to their final scores, with a larger score indicating a higher chance to be extracted. However, this approach performed relevance propagation among homogeneous objects (i.e., sentences). The information beyond the sentence level is totally ignored. Actually, in a given document set, document information can also assist the users in understanding the content in the whole document set. So the document information is supposed to have great influence on sentence ranking.

Based on the above analysis, we argue that the ranking score of a sentence depends not only on its relevance to the given query, but also on the relevance of its belonging document to the query. We apply mutual reinforcement principle to query-focused sentence and document ranking, i.e.,

"A sentence should be ranked higher if it is contained in the document which is more relevant to the given query while a document should be ranked higher if it contains many sentences which are more relevant to the given query."

The above principle is similar to the principle by Zha to detect key terms and generic summary sentences [2].

In this paper, we propose a new query-focused multi-document summarization model, which enhances manifold-ranking based relevance propagation via mutual reinforcement between documents and sentences.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 introduces the proposed two-level mutual reinforcement model and presents a relevance propagation-based sentence ranking algorithm. Section 4 then presents experiments and evaluations. Finally, Section 5 concludes the paper.

RELATED WORK

Under the framework of extractive summarization, sentence ranking is the issue of most concern. In recent years, graph-based approaches have been proposed to rank sentences. These approaches modeled a document or a set of documents as a weighted text graph, took into account the global information and recursively calculated sentence significance from the entire text graph rather than only relying on the unconnected individual sentences.

Most existing query-focused multi-document summarization approaches incorporated the information of the given query into the generic summarizers in order to extract the sentences suiting the user's declared information need. Different from the traditional query-focused summarization approaches, which were usually the simple extensions of generic summarizers and did not uniformly fuse the information in the query and the documents, Wan et al. [1] proposed a manifold ranking based approach to make uniform use of sentence-to-sentence and sentence-to-query relationships. A weighted graph was built where the vertices included both the query description and the sentences in the documents. The manifold ranking was employed to iteratively propagate the relevance of the query to nearby vertices via the graph structure. The ranking score of a sentence obtained by this process indicated the topic-biased informativeness of the sentence and those with high ranks are chosen to form the summary.

MUTUALLY REINFORCED RELEVANCE PROPAGATION (MR²P) MODEL

Notation

Let us formally model the linked two-layer graph containing both sentences and documents as $G = \langle V_S, V_D, E_{SS}, E_{DD}, E_{SD} \rangle$, where $V_S = \{s_i\}$ ($1 \le i \le n_s$, n_s is the total number of the sentences) and $V_D = \{d_j\}$ ($1 \le i \le m_d$, m_d is the total number of the documents). $E_{SS} = \{e_{ij} \mid s_i, s_j \in V_S\}$, $E_{DD} = \{e_{ij} \mid d_i, d_j \in V_D\}$ and $E_{SD} = \{e_{ij} \mid s_i \in V_S, d_j \in V_D\}$ correspond to the edges between sentences, the edges between documents and the edges between sentence and documents, respectively. Let $W_{SS} = [w_{s_i s_j}]_{(n_s+1)\times(n_s+1)}$, $W_{DD} = [w_{d_i d_j}]_{(m_d+1)\times(m_d+1)}$ and $W_{SD} = [w_{s_i d_j}]_{(n_s+1)\times(m_d+1)}$ be the sentence-to-sentence, document-to-document and sentence-to-document affinity matrices, where $w_{s_i s_j}$, $w_{d_i d_j}$ and $w_{s_i d_j}$ are the cosine similarity between two sentences, between two documents, and between one sentence and one document, respectively. Let F_S and F_D denote the ranking scores of the sentence set V_S and the document set V_D, respectively.

Mutually Reinforced Relevance Propagation (MR²P) Algorithm

The MR²P algorithm performs the internal relevance propagation in the sentence set and the document set separately until the stable states of both are achieved. The obtained sentence and document ranking scores are the updated via external mutual reinforcement until all the scores are converged. The basic manifold ranking algorithm is applied as the solution technique to the relevance propagation problem:

$$\begin{cases} Y_S(k+1) = \alpha L_{SS} Y_S(k) + (1-\alpha) Y_S(0) \\ Y_D(k+1) = \alpha L_{DD} Y_D(k) + (1-\alpha) Y_D(0) \end{cases} \quad (1)$$

where $Y_S(0)$ is defined as a constant vector $[1,0,0\ldots0]^T_{(n_S+1)\times 1}$ in relation to the sentence layer, in which the first term denotes the initial relevance of the query sentence and the rest terms denote the relevance of all the sentence in the documents. Similarly, we have the document layer vector $Y_C(0) = [1,0,0\ldots0]^T_{(m_d+1)\times 1}$. $0 < \alpha < 1$ is a weighting parameter. $L_{SS} = A_{SS}^{-1/2} W_{SS} A_{SS}^{-1/2}$ and $L_{DD} = A_{DD}^{-1/2} W_{DD} A_{DD}^{-1/2}$ are graph Laplacian [3], which denote symmetrically normalized W_{SS} and W_{DD}, respectively, where A_{SS} and A_{DD} are the diagonal matrix with the (i,i) element equal to the sum of the ith row of W_{SS} and W_{DD}, respectively. The theorem in [4] guarantees that $Y_S(n+1)$ and $Y_D(n+1)$ converge to

$$\begin{cases} Y_S^* = (I - \alpha L_{SS})^{-1} Y_S(0) \\ Y_D^* = (I - \alpha L_{DD})^{-1} Y_D(0) \end{cases} \quad (2)$$

The resulting sentence and document ranking scores are then refined each other in a mutual reinforcement manner. Let $F_S(0)$ and $F_D(0)$ denote the initial ranking vectors of the sentences and the documents during the reinforcement process, respectively. They are directly derived from Y_S^* and Y_D^* by removing the query dimension. In other words, query does not play any role during the update processing in the MR2P algorithm. The n_S -dimensional vector $F_S(0)$ and the m_D-dimensional vector $F_D(0)$ are then updated by the connections between sentences and documents iteratively according to

$$\begin{cases} F_S(k+1) = \theta F_S(0) + (1-\theta) L_{DS}^T F_D(k) \\ F_D(k+1) = \theta F_D(0) + (1-\theta) L_{DS} F_S(k) \end{cases} \quad (3)$$

where L_{DS} is the normalized document-to-sentence adjacency matrix W_{DS}. $F_S(i)$ and $F_D(i)$ are the refined $F_S(0)$ and $F_D(0)$ at the ith iteration. θ is a weighting parameter ranging from 0 to 1.

$$L_{DS} = A_{DSr}^{-1} W_{DS} A_{DSc}^{-1} \quad (4)$$

where A_{DSr} is the diagonal matrix with its (i,i)-element equal to the sum of the ith row of W_{DS}. A_{DSc} is the diagonal matrix with its (i,i)-element equal to the sum of the ith column of W_{DS}.

$L_{DS}^T F_D$ and $L_{DS} F_S$ in Eq. (3) reveal the influence of document ranking on the sentence ranking and the influence of sentence ranking on document ranking through the connections between sentences and documents. Therefore, the first equation in Eq.(3) addresses the F_S update reinforced by F_D, while the second equation address the F_D update reinforced by F_S. In certain degree the algorithm still keeps confidence on the original sentence and document ranking results that are obtained based on query relevance propagation. The convergence proof of MR^{2}P algorithm is omitted due to space limit.

We can use the closed form which is obtained in the proof process to compute the ranking scores of sentences and documents directly. In large-scale real-world problems; however, we prefer using iteration algorithm due to computational efficiency. Usually, the convergence of the iteration algorithm is achieved when the maximal difference between sentence scores or document scores computed at two successive iterations for any point falls below a given threshold (0.0001 in this study).

Summary Generation and Redundancy Control

In multi-document summarization, the number of the documents to be summarized can be very large. This makes information redundancy problem appear to be more serious in multi-document summarization than in single-document summarization. Redundancy control becomes an inevitable process. Since our focus in this study is the design of effective (sentence) ranking algorithms, we apply a straightforward but effective sentence selection principle. We incrementally add into the summary the highest ranked sentence of concern if it doesn't significantly repeat the information already included in the summary until the word limitation of the summary is reached.

EXPERIMENT

We conduct the experiments on the data sets from the DUC 2005 and the DUC 2006. In these two years, query-focused multi-document summarization is the only task. According to the task definitions, systems are required to produce a concise summary for each document set and the length of summaries is limited to 250 English words.

A well-recognized automatic evaluation toolkit ROUGE [5] is used for evaluation. It measures summary quality by counting the overlapping units between system-generated summaries and human-written reference summaries. We report three common ROUGE scores in this paper, namely ROUGE-1, ROUGE-2 and ROUGESU4 which base on Uni-gram match, Bi-gram match, and unigram plus skip-bigram match with maximum skip distance of 4. Documents and queries are pre-processed by segmenting sentences and splitting words. Stop words are removed and the remaining words are stemmed using Porter stemmer.

In the experiments, the proposed MR2P algorithm is compared with the two baselines employed in the DUC. They are the lead baseline and the coverage baseline. The lead baseline takes the first sentences one by one in the last document in the collection, where documents are assumed to be ordered chronologically. The coverage baseline takes the first sentence one by one from the first document to the last document. We also present the results of top three systems with the highest ROUGE scores that participate in the DUC 2005 and the DUC 2006 for comparison.

For further comparison of the MR^{2}P algorithm, we also implement the basic manifold ranking algorithm as proposed in [1].

Tables 1 and 2 show the comparison results on the DUC 2005 and 2006 data sets respectively. The parameters of manifold ranking based approaches are set as $\alpha = 0.6$ and $\theta = 0.75$. S15 and S4 etc. in the tables are the IDs of those top performing systems participated in the DUC.

TABLE 1 : Experimental Results on the Data of DUC2005

Systems	ROUGE-1	ROUGE-2	ROUGE-SU4
MR^{2}P	0.38591	0.07498	0.13397
Wan’s	0.38523	0.07496	0.13353
S15	0.37665	0.07381	0.13260
S4	0.37484	0.07003	0.12798
S17	0.36930	0.07256	0.12977
Coverage Baseline	0.34659	0.06013	0.09275
Lead Baseline	0.30583	0.04875	0.08154

TABLE 2 : Experimental Results on the Data of DUC2006

Systems	ROUGE-1	ROUGE-2	ROUGE-SU4
MR^{2}P	0.41715	0.10285	0.17429
Wan’s	0.41685	0.10279	0.17401
S12	0.41611	0.10276	0.17399
S23	0.41505	0.10800	0.17834
S24	0.41020	0.10727	0.17431
Coverage Baseline	0.36753	0.08132	0.14596
Lead Baseline	0.33574	0.06942	0.12439

CONCLUSION AND FUTURE WORK

In this paper, we propose a new model to enhance manifold- ranking based relevance propagation via mutual reinforcement between sentences and documents. Based on it, we develop a new sentence ranking algorithms for the application of query-focused multi-document summarization. Experimental results demonstrate the effectiveness of the proposed algorithm. In the future, we will further explore how to combine sentence-to-document relationship in multi-modality relevance propagation model.

ACKNOWLEDGEMENT

The work described in this paper was partially support by “Twelfth Five Year” National Science and Technology Support Program (No. 2012BAH30F01) and National Natural Science Foundation of China (Project No. 61303125, 61303226).

REFERENCES

[1] X.J.Wan, J. W. Yang, and J. G. Xiao; Manifold-ranking based topic-focused multi-document summarization. in Proc.18th IJCAI Conf., pp.2903–2908,(2007).
[2] H.Y.Zha. Generic summarization and key phrase extraction using mutual reinforcement principle and sentence clustering. in Proc. 25th SIGIR Conf., pp. 113–120, (2002).
[3] U.V.Luxburg. A tutorial on spectral clustering.Statist. Comput., vol.17, no.4.(2007).
[4] D.Y.Zhou, J.Weston, A.Gretton, O.Bousquet, and B.Scholkopf. Ranking on data manifolds. in Proc. 17th NIPS Conf., pp. 169–176.(2003).
[5] Lin, C.Y.. ROUGE: A package for automatic evaluation of summaries. In Proc. of ACL. (2004).