Update Summarization Based on Co-Ranking with Constraints

Xiaojun Wan
Institute of Computer Science and Technology
The MOE Key Laboratory of Computational Linguistics
Peking University, Beijing 100871, China
wanxiaojun@pku.edu.cn

ABSTRACT

Update summarization is an emerging summarization task of creating a short summary of a set of news articles, under the assumption that the user has already read a given set of earlier articles. In this paper, we propose a new co-ranking method to address the update summarization task. The proposed method integrates two co-ranking processes by adding strict constraints. In comparison with the original co-ranking method, the proposed method can compute more accurate scores of sentences for the purpose of update summarization. Evaluation results on the most recent TAC2011 dataset demonstrate that our proposed method can outperform the original co-ranking method and other baselines.

KEYWORDS : Update Summarization, Multi-document summarization, Co-Ranking

Proceedings of COLING 2012: Posters, pages 1291–1300,
COLING 2012, Mumbai, December 2012.

1 Introduction

Update summarization is an emerging new summarization task of creating a short summary of a set of news articles, under the assumption that the user has already read a given set of earlier articles. The purpose of update summarization is to inform the reader of new information about a particular topic. Update summary is very useful for the user to know about a chronic topic. For example, given a topic of "Haiti earthquake", the earlier articles mainly talk about the occurrence of the earthquake and the consequence of the earthquake, and the later articles talk about the consequence of the earthquake and the rescue issues. In this case, the reader will read the later articles to know about the rescue issues after he/she has read the earlier articles. Therefore, an update summary of the later articles may facilitate the reader to grasp the "update" information in a very convenient way.

The update summarization task can be formulated as follows: Given an earlier document set D^A and a later document set D^B about a topic q, the sentence set for D^A is denoted as S^A and the sentence set for D^B is denoted as S^B. The update summary with a predefined length for D^B after reading D^A is denoted as SUM^B, where the sentences in SUM^B must meet the following requirements:

1) The summary sentences must be representative of D^B, i.e., the summary sentences in SUM^B must reflect important information in D^B. Moreover, the sentences must be biased to the topic q.
2) The summary sentences must be the least redundant with the sentences in D^A, i.e., the summary must not contain important information in D^A.

The task of update summarization was piloted in DUC2007, and it has been the fundamental task through TAC2008~TAC2011. The "update" characteristic makes the task more challenging than traditional document summarization tasks. Till now, most existing update summarization methods are adaptations of multi-document summarization methods by considering the redundancy information between the earlier and later document sets (Boudin et al. 2008; Fisher and Roark 2008; Nastase et al. 2008). In addition, several new methods have been proposed for addressing this task (Du et al. 2010; Wang and Li 2010; Li et al. 2008), and graph-based co-ranking is a typical one, where the sentences in the two document sets are ranked simultaneously by considering the sentence relationships across the document sets. Based on the co-ranking framework, Li et al. (2008) propose a graph-based sentence ranking algorithm named PNR^2 for update summarization, and it models both the positive and negative mutual reinforcement between sentences in the ranking process. In addition, Wan et al. (2011) apply the co-ranking algorithm for multilingual news summarization.

In this study, we propose a new co-ranking method, which is inspired by (Wan et al. 2011), to address the update summarization task. The proposed method integrates two co-ranking processes by adding strict constraints. In comparison with the original co-ranking method, the proposed method can compute more accurate scores of sentences for the purpose of update summarization. We perform experiments on the most recent TAC2011 dataset, and the evaluation results demonstrate that out proposed method can outperform the original co-ranking method and a few other baselines.

2 Our proposed method

Given the two document sets D^A and D^B about a topic q, we introduce two kinds of scores for each sentence: an update score and a consistency score. In our proposed method, each kind of score is computed with a co-ranking process, and the two kinds of scores are adjusted by adding strict constraints. Finally, the refined update scores are used for summary extraction.

We assign each sentence an update score to indicate how much the sentence contains significant and new information after knowing about the sentences in the other document set. The update score of each sentence relies not only on the sentences in the same document set, but also on the sentences in the other document set. In particular, the co-ranking method is based on the following assumption:

> *The update score of a sentence is positively associated with the sentences with high update scores in the same document set, and is negatively associated with the sentences with high update scores in the other document set.*

We introduce a consistency score for each sentence to indicate how much a sentence contains important and shared information in the two document sets. In particular, the consistency scores of the sentences can be computed based on the following assumption:

> *The consistency score of a sentence is positively associated with the sentences with high consistency scores in the same document set, and is also positively associated with the sentences with high consistency scores in the other document set.*

Formally, let $G=(S^A, S^B, E^A, E^B, E^{AB})$ be an undirected graph for the sentences in the two document sets D^A and D^B. $S^A=\{s^A_i \mid 1\leq i\leq m\}$ is the set of earlier sentences. $S^B=\{s^B_j \mid 1\leq j\leq n\}$ is the set of later sentences. m, n are the sentence numbers in the two document sets, respectively. Each sentence s^A_i or s^B_j is represented by a term vector $\vec{s}_i^{\,A}$ or $\vec{s}_j^{\,B}$ in the VSM model. E^A is the edge set to reflect the similarity relationships between the sentences in the earlier document set. E^B is the edge set to reflect the similarity relationships between the sentences in the later document set. E^{AB} is the edge set to reflect the similarity or dissimilarity relationships between the sentences in the two different document sets. The following matrices are required to be computed to reflect the three kinds of sentence relationships:

$\boldsymbol{M}^A=[M^A_{ij}]_{m\times m}$: This matrix aims to reflect the similarity relationships between the sentences in S^A. Each entry in the matrix corresponds to the cosine similarity between two sentences, and we let $M^A_{ii}=0$. Then $\boldsymbol{M}^A$ is normalized to $\tilde{\boldsymbol{M}}^A$ to make the sum of each row equal to 1.

$\boldsymbol{M}^B=[M^B_{ij}]_{n\times n}$: This matrix aims to reflect the similarity relationships between the sentences in S^B. Each entry in the matrix corresponds to the cosine similarity between two sentences, and we let $M^B_{ii}=0$. Then $\boldsymbol{M}^B$ is normalized to $\tilde{\boldsymbol{M}}^B$ to make the sum of each row equal to 1.

$\boldsymbol{W}^{AB}=[W^{AB}_{ij}]_{m\times n}$: This matrix aims to reflect the similarity relationships between the two sets of sentences. Each entry W^{AB}_{ij} in the matrix corresponds to the cosine similarity value between the sentence s^A_i and the sentence s^B_j. Then $\boldsymbol{W}^{AB}$ is normalized to $\tilde{\boldsymbol{W}}^{AB}$ to make the sum of each row equal to 1. In addition, we use $\boldsymbol{W}^{BA}=[W^{BA}_{ij}]_{n\times m}$ to denote the transpose of $\boldsymbol{W}^{AB}$, i.e., $\boldsymbol{W}^{BA}=(\boldsymbol{W}^{AB})^{\mathrm{T}}$. Then $\boldsymbol{W}^{BA}$ is normalized to $\tilde{\boldsymbol{W}}^{BA}$ to make the sum of each row equal to 1.

$\boldsymbol{M}^{AB}=[M^{AB}_{ij}]_{m\times n}$: This matrix aims to reflect the dissimilarity relationships between the sentences in S^A and the sentences in S^B. Each entry M^{AB}_{ij} in the matrix corresponds to the dissimilarity between the sentence s^A_i and the sentence s^B_j.

$$M_{ij}^{AB} = \frac{\left\|\vec{s}_i^{\,A} - \vec{s}_j^{\,B}\right\|}{\left\|\vec{s}_i^{\,A}\right\| \times \left\|\vec{s}_j^{\,B}\right\|}$$

Then $\boldsymbol{M}^{AB}$ is normalized to $\tilde{\boldsymbol{M}}^{AB}$ to make the sum of each row equal to 1. In addition, we use $\boldsymbol{M}^{BA}=[M^{BA}_{ij}]_{n\times m}$ to denote the transpose of $\boldsymbol{M}^{AB}$, i.e., $\boldsymbol{M}^{BA}= (\boldsymbol{M}^{AB})^{\mathrm{T}}$. Then $\boldsymbol{M}^{BA}$ is normalized to $\tilde{\boldsymbol{M}}^{BA}$ to make the sum of each row equal to 1.Note that $\tilde{\boldsymbol{M}}^{AB}$ and $\tilde{\boldsymbol{M}}^{BA}$ directly embody the negative association between the sentences in the two sets.

In order to compute the query-biased scores of the sentences, the relevance values of the sentences to the query also need to be computed. We use two column vectors $\boldsymbol{r}^A=[r^A_i]_{m\times 1}$ and $\boldsymbol{r}^B=[r^B_j]_{n\times 1}$ to reflect the query-biased scores, where each entry in $\boldsymbol{r}^A$ corresponds to the cosine similarity between a sentence and the given topic description. Then $\boldsymbol{r}^A$ is normalized to $\tilde{\boldsymbol{r}}^{\boldsymbol{A}}$ to make the sum of all elements equal to 1. Each entry in $\boldsymbol{r}^B$ is computed in the same way, and $\boldsymbol{r}^B$ is normalized to $\tilde{\boldsymbol{r}}^{\boldsymbol{B}}$.

After computing the above matrices and vectors, we can compute the update scores of the sentences in the two sets in a co-ranking process. We use two column vectors $\boldsymbol{u}^{\boldsymbol{A}} =[u^{\boldsymbol{A}}_j]_{m\times 1}$ and $\boldsymbol{u}^{\boldsymbol{B}}=[u^{\boldsymbol{B}}_i]_{n\times 1}$ to denote the update scores of the sentences in S^A and the sentences in S^B , respectively. Based on the first assumption, we can obtain the following equations:

$$u_j^A = \alpha_1 \sum\nolimits_i \tilde{M}_{ij}^{A} u_i^A + \beta_1 \sum\nolimits_i \tilde{M}_{ij}^{BA} u_i^B + \gamma_1 \cdot r_j^A$$
$$u_i^B = \alpha_1 \sum\nolimits_j \tilde{M}_{ji}^{B} u_j^B + \beta_1 \sum\nolimits_j \tilde{M}_{ji}^{AB} u_j^A + \gamma_1 \cdot r_i^B$$

where α_1, β_1, $\gamma_1 \in [0, 1]$ specify the relative contributions to the final scores from different sources and we have $\alpha_1+\beta_1+\gamma_1=1$. Note that since $\tilde{\boldsymbol{M}}^{AB}$ and $\tilde{\boldsymbol{M}}^{BA}$ contain the dissimilarity values between the two sets of sentences, the second terms in the right hands of the above equations actually embody the negative reinforcement between the two sets of sentences. Different from (Li et al. 2008), the addition of all the terms in the right hands of the equations makes the algorithm more convenient to be solved in an iterative way.

We can also compute the consistency scores of the sentences in the two sets in a co-ranking process. We use two column vectors $\boldsymbol{v}^{\boldsymbol{A}} =[v^{\boldsymbol{A}}_j]_{m\times 1}$ and $\boldsymbol{v}^{\boldsymbol{B}}=[v^{\boldsymbol{B}}_i]_{n\times 1}$ to denote the consistency scores of the sentences in S^A and the sentences in S^B , respectively. Based on the second assumption, we can obtain the following equations:

$$v_j^A = \alpha_2 \sum\nolimits_i \tilde{M}_{ij}^{A} v_i^A + \beta_2 \sum\nolimits_i \tilde{W}_{ij}^{BA} v_i^B + \gamma_2 \cdot r_j^A$$
$$v_i^B = \alpha_2 \sum\nolimits_j \tilde{M}_{ji}^{B} v_j^B + \beta_2 \sum\nolimits_j \tilde{W}_{ji}^{AB} v_j^A + \gamma_2 \cdot r_i^B$$

where α_2, β_2, $\gamma_2 \in [0, 1]$ specify the relative contributions to the final scores from different sources and we have $\alpha_2+\beta_2+\gamma_2=1$.

Then, we interconnect the two co-ranking processes based on our key assumption that the update score and the consistency score of each sentence is mutually exclusive. If the update score of a sentence is high, then the sentence contains significant and new information, which is not contained in the other document set; but if the consistency score of a sentence is high, then the sentence contains significant and shared information with the other document set. Therefore, the update score and the consistency score of a sentence are conflicting with each other, and they cannot be high at the same time.

> *The sum of the update score and the consistency score of each sentence is fixed to a particular value.*

This assumption can be used to adjust the inaccurately assigned scores for the sentences.

Till now, the update scores and the consistency scores are computed by using a co-ranking process separately. Based on our new assumption, we can add the following constraints to interconnect the two co-ranking processes:

$$u_j^A + v_j^A = \varepsilon_j^A \qquad u_i^B + v_i^B = \varepsilon_i^B$$

where ε_j^A and ε_i^B are the specified fixed sum values for the sentences $s^A{}_i$ and $s^B{}_i$. The values for different sentences may be different since they are unequally important in the document sets. In this study, we use the generic informativeness score of each sentence as the fixed sum value for the sentence. The generic informativeness score of a sentence is computed by using the basic graph-based ranking algorithm. Taking a sentence $s^B{}_i$ in S^B as an example, the value can be computed in a recursive form as follows:

$$\varepsilon_i^B = \mu \cdot \sum_{all\, j \neq i} \varepsilon_j^B \cdot \widetilde{M}_{ji}^B + \frac{(1-\mu)}{\text{n}}$$

where μ is the damping factor usually set to 0.85, as in the PageRank algorithm. The generic informativeness score of a sentence in S^A can be computed based on $\widetilde{\boldsymbol{M}}^A$ in the same way.

In order to add the constraints to interconnect the two co-ranking processes, the constraints are executed as a normalization step. In particular, the following steps are iteratively performed until convergence. Note that all the scores are simply initialized to 1, and (t+1), (t) means the (t+1)-th and (t)-th iterations, respectively.

1) Compute the update scores of the sentences with the following equations:

$$(u_j^A)^{(t+1)} = \alpha_1 \sum\nolimits_i \widetilde{M}_{ij}^A (u_i^A)^{(t)} + \beta_1 \sum\nolimits_i \widetilde{M}_{ij}^{BA} (u_i^B)^{(t)} + \gamma_1 \cdot r_j^A$$

$$(u_i^B)^{(t+1)} = \alpha_1 \sum\nolimits_j \widetilde{M}_{ji}^B (u_j^B)^{(t)} + \beta_1 \sum\nolimits_j \widetilde{M}_{ji}^{AB} (u_j^A)^{(t)} + \gamma_1 \cdot r_i^B$$

$$(\boldsymbol{u}^A)^{(t+1)} = (\boldsymbol{u}^A)^{(t+1)} / \left\|(\boldsymbol{u}^A)^{(t+1)}\right\| \qquad (\boldsymbol{u}^B)^{(t+1)} = (\boldsymbol{u}^B)^{(t+1)} / \left\|(\boldsymbol{u}^B)^{(t+1)}\right\|$$

2) Compute the consistency scores of the sentences with the following equations:

$$(v_j^A)^{(t+1)} = \alpha_2 \sum\nolimits_i \widetilde{M}_{ij}^A (v_i^A)^{(t)} + \beta_2 \sum\nolimits_i \widetilde{W}_{ij}^{BA} (v_i^B)^{(t)} + \gamma_2 \cdot r_j^A$$

$$(v_i^B)^{(t+1)} = \alpha_2 \sum\nolimits_j \widetilde{M}_{ji}^B (v_j^B)^{(t)} + \beta_2 \sum\nolimits_j \widetilde{W}_{ji}^{AB} (v_j^A)^{(t)} + \gamma_2 \cdot r_i^B$$

$$(\boldsymbol{v}^{\boldsymbol{A}})^{(t+1)} = (\boldsymbol{v}^{\boldsymbol{A}})^{(t+1)} / \left\|(\boldsymbol{v}^{\boldsymbol{A}})^{(t+1)}\right\| \qquad (\boldsymbol{v}^{\boldsymbol{B}})^{(t+1)} = (\boldsymbol{v}^{\boldsymbol{B}})^{(t+1)} / \left\|(\boldsymbol{v}^{\boldsymbol{B}})^{(t+1)}\right\|$$

3) Add the constraints on the update scores and the consistency scores of the sentences by normalization with the following equations ($\boldsymbol{\zeta}^A$, $\boldsymbol{\eta}^B$ are temporary vectors):

$$\xi_j^A = (u_j^A)^{(t+1)} + (v_j^A)^{(t+1)} \qquad \eta_i^B = (u_i^B)^{(t+1)} + (v_i^B)^{(t+1)}$$

$$(u_j^A)^{(t+1)} = \varepsilon_j^A \cdot \frac{(u_j^A)^{(t+1)}}{\xi_j^A} \qquad (v_j^A)^{(t+1)} = \varepsilon_j^A \cdot \frac{(v_j^A)^{(t+1)}}{\xi_j^A}$$

$$(u_i^B)^{(t+1)} = \varepsilon_i^B \cdot \frac{(u_i^B)^{(t+1)}}{\eta_i^B} \qquad (v_i^B)^{(t+1)} = \varepsilon_i^B \cdot \frac{(v_i^B)^{(t+1)}}{\eta_i^B}$$

Finally, we obtain the update scores $\boldsymbol{u}^{\boldsymbol{B}}$ for the sentences in the later document set D^B, and we apply the simple greedy algorithm in (Wan et al. 2007) to remove redundant sentences and select summary sentences until the summary length reaches the given limit. Note that in the experiments, the iteration number of the above algorithm is mostly around10, which is very efficient.

3 Empirical evaluation

3.1 Evaluation setup

In this study, we used the most recent update summarization task on TAC 2011 for evaluation purpose. NIST selected 44 topics, and two sets of 10 documents (set A and set B) were provided for each topic. The update task aims to create a 100-word summary of 10 documents in set B, with the assumption that the content of the first 10 documents in set A is already know to the reader. For each document set, NIST assessors have created 4 human summaries as reference (model) summaries. The sentences have already been split for the documents.

For each topic, we only used the topic title as the topic description. As a pre-processing step, we removed the very long or very short sentences, which are usually not good summary sentences. We also polished some sentences to make them more concise by applying simple rules, e.g. removing some clauses in the sentences. The sentences in the documents were then stemmed by using Porter's stemmer. Our proposed summarization method and the baseline methods were performed on the pre-processed document sets.

We used the ROUGE-1.5.5 toolkit[1] for evaluation, which was officially adopted by DUC for automatic summarization evaluation. The toolkit measures summary quality by counting overlapping units such as the n-gram, word sequences and word pairs between the candidate summary and the reference summary. The ROUGE toolkit reports separate scores for 1, 2, 3 and 4-gram, and also for longest common subsequence co-occurrences. In this study, we show three ROUGE scores in the experimental results[2]: ROUGE-1 (unigram-based), ROUGE-2 (bigram-based), and ROUGE-SU4 (based on skip bigram with a maximum skip distance of 4).

[1] http://www.berouge.com

[2] We used the options: -n 4 -w 1.2 -m -2 4 -u -c 95 -r 1000 -f A -p 0.5 -t 0 -a -l 100.

In the experiments, our proposed method is compared with the following baseline methods:

Lead: This baseline is provided by NIST, and it returns all the leading sentences (up to 100 words) in the most recent document. Baseline 1 provides a lower bound on what can be achieved with a simple fully automatic extractive summarizer.

Mead: This baseline is also provided by NIST, and it uses the MEAD automatic summarizer with all default settings, to produce summaries.

MMR: This baseline is based on the MMR criterion for selecting summary sentences with new information.

SinkManifoldRank: This baseline is a new graph based ranking method for update summarization, which is based on manifold ranking with sink points (Du et al. 2010).

CoRank: This baseline is the basic co-ranking method for update summarization, and it directly uses the co-ranking algorithm to compute the update score for each sentence, without considering the constraints between the update score and the consistency score.

For the baseline co-ranking method, we let $\gamma_1 = 0.15$, as in the PageRank algorithm, and thus we have $\alpha_1+\beta_1 = 0.85$, and $\alpha_1:\beta_1$ is empirically set to 0.7:0.3. For our method, we also let $\gamma_1=\gamma_2=0.15$. Thus we have $\alpha_1+\beta_1 =\alpha_2+\beta_2= 0.85$, and $\alpha_1:\beta_1$ is set to 0.5:0.5 and $\alpha_2:\beta_2$ is set to 0.7:0.3.

3.2 Evaluation results

First, our proposed method is compared with the baseline methods, and the comparison results are shown in Table 1. In the table, the 95% confidence interval of each ROUGE score is given in brackets, which is reported by the ROUGE toolkit.

We can see from the table that our proposed method outperforms all the baseline methods over all three metrics. In particular, the baseline co-ranking method performs better than other baselines, and our proposed method can achieve better performance than the co-ranking method. The results demonstrate the good effectiveness of our proposed method.

Method	**ROUGE-1**	**ROUGE-2**	**ROUGE-SU4**
Our Method	**0.36795** [0.35673 - 0.37893]	**0.08838** [0.07894 - 0.09812]	**0.12716** [0.11931 - 0.13544]
CoRank	0.36143 [0.34755 - 0.37588]	0.07994 [0.06960 - 0.09048]	0.12164 [0.11274 - 0.13151]
SinkManifoldRank	0.31112 [0.29678 - 0.32543]	0.06198 [0.05456 - 0.06946]	0.10106 [0.09373 - 0.10878]
MMR	0.34724 [0.33493 - 0.36005]	0.07450 [0.06548 - 0.08367]	0.11529 [0.10780 - 0.12326]
Mead	0.28347 [0.27062 - 0.29696]	0.05903 [0.05037 - 0.06781]	0.09132 [0.08444 - 0.09850]
Lead	0.29378 [0.27684 - 0.30969]	0.05685 [0.04769 - 0.06680]	0.09449 [0.08637 - 0.10289]

TABLE 1 – Comparison results

Second, our proposed method is compared with the participating systems on TAC 2011. On TAC 2011, NIST received 48 runs from 24 participating teams. We rank the runs based on the ROUGE-2 scores, and list the top five runs for comparison. In addition, we also compute the average ROUGE scores. The comparison results are shown in Table 2.

We can see from the table that our proposed method ranks 4th out of all the runs over the ROUGE-2 metric. The performance of our proposed method is comparable with the top run's performance. We can also see that the performance values of our proposed method are much better than the average scores. Note that the top runs have leveraged world knowledge or various features, for example, the NUS1 run has used category knowledge in a supervised machine learning approach. However, our proposed method only uses the similarity/dissimilarity relationships between sentences in an unsupervised approach. The comparison results demonstrate that our proposed method is a competitive method for the update summarization task.

ID & Run Name	**ROUGE-2**	**ROUGE-SU4**
43 (NUS1)	0.09581	0.13080
25 (CLASSY2)	0.09259	0.12759
17 (NUS2)	0.08855	0.12792
Our Method	0.08838	0.12716
24 (PolyCom2)	0.08643	0.12803
35 (SIEL_IIITH2)	0.08538	0.12376
TAC Average	0.07053	0.11009

TABLE 2 – Comparison with top five runs (out of 48 runs, ranked by ROUGE-2) on TAC 2011

Conclusion and future work

In this paper, we propose a new method for update summarization, and it improves the basic co-ranking method by adding strict constraints and interconnecting two co-ranking processes. Evaluation results on the most recent TAC 2011 dataset demonstrate the good effectiveness of the proposed method, which can outperform a few baseline methods, and the performance is comparable to the top participating systems on TAC 2011.

In this study, we only use the topic title as the topic description, however, the title is usually very short, and we will investigate query expansion techniques to get a clearer topic description. We will also make use of the topic category specific features (i.e. the guided information) to improve our update summarization method in future work.

Acknowledgments

The work was supported by NSFC (61170166), Beijing Nova Program (2008B03) and National High-Tech R&D Program (2012AA011101).

References

Boudin, F., El-Bѐze, M., and Torres-Moreno, J.-M. (2008). The LIA update summarization systems at TAC-2008. In *Proceedings of TAC2008*.

Du, P., Guo, J., Zhang, J., and Cheng, X. (2010). Manifold ranking with sink points for update summarization. In *Proceedings of CIKM2010*.

Fisher, S. and Roark, B. (2008). Query-focused supervised sentence ranking for update summaries. In *Proceeding of TAC2008*.

Li, W., Wei, F., Lu, Q., and He, Y. (2008). PNR^2: Ranking sentences with positive and negative reinforcement for query-oriented update summarization. In *Proceedings of COLING2008*.

Nastase, V., Filippova, K., and Ponzetto, S. P. (2008). Generating update summaries with spreading activation. In *Proceedings of TAC2008*.

Wan, X., Jia, H., Huang, S., and Xiao, J. (2011). Summarizing the Differences in Multilingual News. In *Proceedings of SIGIR2011*.

Wan, X., Yang, J., and Xiao, J. (2007). Manifold-ranking based topic-focused multi-document summarization. In *Proceedings of IJCAI-07*.

Wang, D. and Li, T. (2010). Document update summarization using incremental hierarchical clustering. In *Proceedings of CIKM2010*.