DEIM Forum 2012 D5-1

マルチメディアニュース推薦のための補完木

許　レイ†　馬　強†　吉川　正俊†

† 京都大学情報学研究科　〒 606–8501 京都府京都市左京区吉田本町

E-mail: †xu@db.soc.i.kyoto-u.ac.jp, ††{qiang,yoshikawa}@i.kyoto-u.ac.jp

キーワード　マルチメディアニュース、クラスタ、補完木、ニュース推薦

A Complementary Tree for Recommendation of Multimedia News Items

Ling XU†, Qiang MA†, and Masatoshi YOSHIKAWA†

† Graduate School of Informatics, Kyoto University

Yoshida-Honmachi, Sakyo, Kyoto 606–8501 Japan

E-mail: †xu@db.soc.i.kyoto-u.ac.jp, ††{qiang,yoshikawa}@i.kyoto-u.ac.jp

Abstract　To reduce the information redundance and help user acquire comprehensive and detailed information from multimedia news, we proposed a complementary tree organizing the news items reporting a same topic. The complementary tree is constructed by considering two complementary relationships between two news items: 1) same opinion that possibly providing detailed information, 2) different opinion that providing diverse viewpoints. Here, opinion is extracted from the sentiment words in the closed captions of a news item. We also demonstrate the way to news recommendation by using the complementary tree.

Key words　multimedia news, cluster, complementary tree, news recommendation

1. Introduction

Nowadays, people are exposed to a large number of multimedia news every day. When searching for a particular news in the search engine such like Google Video(注1) and Youtube(注2), there come a mass of related results. Despite the relevance with the user need, it is difficult for user to watch or check all the contents, even only browsing the snippet and introduction of the contents. How to quickly and effectively present the contents still remains an issue.

Now it is possible to help official or well-evaluated web page results maintain rank of upper by using some advanced retrieval methods such as PageRank [2]. For multimedia news from television programs or online stream on the Internet, the existing retrieval methods have limits, especially for the new coming news. Thus, we focus on the contents analysis. However, in multimedia news, the contents analysis becomes very complicated, because of the free perspectives, complex editing methods and the relevance of the continuous reports in the story. Considering the complement of the multimedia news, current methods are far from satisfied. We are looking forward to a recommendation method of multimedia news that can help user acquire comprehensive and detailed information as well as know the significant information and the relationships between news.

In this paper, we proposed a complementary tree organizing the multimedia news items. In the tree, every news item is located as a node. In the sub-tree of each node there are the nodes providing the complement information to the parent node. In the binary tree, descendants in the left branch and descendants in the right branch indicate the related nodes providing complementary information in different meanings.

Two complement relationships are: 1) same opinion that possibly providing detailed information, and 2) different opinion that providing diverse viewpoints. Here, opinion is extracted from the sentiment words in the closed captions of a news item.

In our methods, all the nodes are inserted to the tree one by one according to the published time. Latest published news item is inserted last when ensuring the possibly posi-

(注1)：Google Video:http://www.google.com/videohp

(注2)：Youtube:http://www.youtube.com/

tion with higher priority than the others. A complementary tree is marked with features:

（1） Proper sub-tree corresponding to the nodes at 1^{st} level is a binary tree where nodes are the multimedia news reporting related contents,

（2） Any node in the proper sub-tree except the root, provides complement contents to the parent,

（3） Descendants in the left sub-tree provide the detailed information of similar opinion of the parent, and

（4） Descendants in the right sub-tree provide the diverse viewpoints with different opinions.

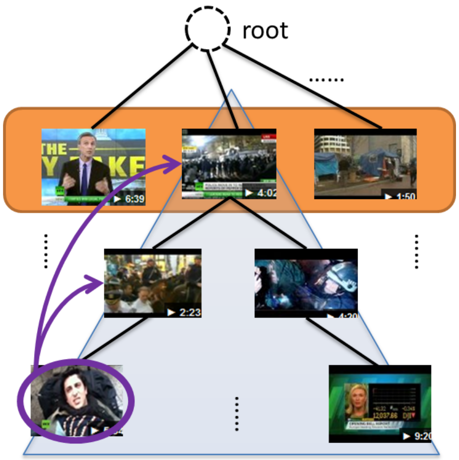


図 1　Complementary Tree of "NYPD, Wall Street, Occupy".

Figure 1 shows the tree of “NYPD[注3] Wall Street Occupy”. Triangle in the middle shows the sub-tree of news items reporting related contents. Nodes in red region are the news items reporting different events. From left to right, they are talking about:

（1） Bloomberg said the media was kept away but nobody seems to be buying that defence,

（2） Police begun clearing out “Occupy” protester camps in New York, and

（3） “Occupy Boston” said they fears the same fate.

The remainder of this paper is organized as follows: Section 2 discusses related work. Section 3 describes the features of the complementary tree. Section 4 is our novel method of the construction. Section 5 concludes this paper with a brief summary and mentions future work.

2. Related Work

In 2011, Y. Lv et al. [9] study the problem of post-click news recommendation and characterize relatedness across four aspects. Our analysis of relatedness is based on the comparison of contents in the multimedia news, not the user activities.

(注3)：New York Police Department

There also exist contents-based methods. I. Ide et al. analyzed the inter-structure of news video [3], [5]. Y. Yang et al. use similarity-based functions to search interfaces with the goal to retrieve content related to the query document [7]. Q. Ma et al. proposed a content-representation model “topic-structure” model [6] for extracting the topic structures and searching for complement information from various perspectives. A. Ahmed et al. proposed a new integrated algorithm for estimating propensity of interacting content [11].

In our method, we focus on the multimedia news and compare them in both material and opinion. With enough related news data, we can construct a complementary tree for presenting and recommending the news items. The complementary tree also reminds the different relations between news items in the recommendation.

3. Complementary Tree for Recommendation

Supposing D is a collection of multimedia news items. Each multimedia news item d consists of the contents of video and text (=closed captions in the multimedia news content). The complementary tree T is initiated at the request of the user (e.g., searching for the multimedia news with key words of “NYPD, Wall Street, Occupy”).

$$T = (V, E) \tag{1}$$

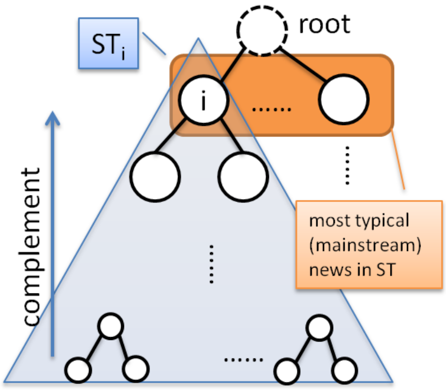


図 2　Complementary Tree and Sub-tree ST.

As shown in Figure 2, T satisfies the following properties:

（1） Root node is a virtual item indicating a topic. As shown in Figure 1, root node indicates the user request of “NYPD, Wall Street, Occupy”. According to the events in the topic, root node has one or multiple child nodes.

（2） Each node at level 2 is a root node of a binary tree ST_i, where i is the id of the root node. In each ST_i, node n_i indicates the most typical reported multimedia news item in the sub-tree. In the specific process,

(3) In each ST_i, any n_i's descendants complement its ancestor(s). For example, in Figure 3, n_j is one of the nodes in ST_i. Supposing the left sub-tree of n_j is st_j^L and the right sub-tree of n_j is st_j^R. Every node in st_j^L indicates the multimedia news providing detailed information complement to n_j. Every node in st_j^R indicates the multimedia news providing contents from broaden information from different views complement to n_j.

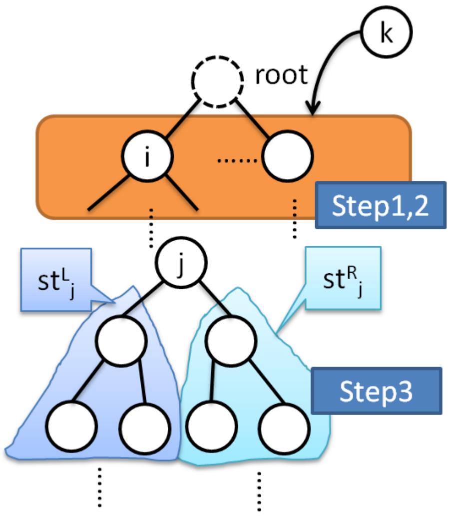


図 3 News Items Sharing Similar Materials/Opinions.

4. Construction of Complementary Tree by Using A Material-Opinion Model

In our method, complementary tree T is initiated at the request of the user. Root node is a virtual item indicating a topic and its descendants are the multimedia news items related to the topic. Possibly there are multiple events in the topic. Proper sub-tree ST_i corresponding to the nodes at 1^{st} level indicates the event.

For each ST_i, descendants in the left sub-tree provide the detailed information of similar opinion to the parent. Descendants in the right sub-tree provide the diverse viewpoints with different opinions.

For the comparison of two multimedia news, we proposed a material-opinion model for the contents analysis. In the model, material is the visual description of each video that filmed by the camera as a reflection of the real world, which is supposed to be believable despite the aggrandizement or forgery in films. Opinion is extracted from the textual description of each news, consisting of the sentiment words put forward by the news agency. For a multimedia news c_k,

$$c_k = (M_k, O_k) \tag{2}$$

$$M_k = \{seg_p | seg_p \in V_k\} \tag{3}$$

$$O_k = \{word_q | word_q \in S_k\} \tag{4}$$

where V_k is the visual description of c_k; S_k are the sentences in the textual descriptions. p is the number of segments. M_k and O_k are the material and opinion of c_k. seg_p is a segment in video of c_k.

When we compare two news items, we compare their materials and opinions for computing the material dissimilarity and opinion dissimilarity. We use the dissimilarities for clustering the news items into different branches in the tree indicating the different complementary relationships. We also use the dissimilarities for computing the typical score t_s for deciding the most mainstream, typical and representative news in the sub-tree.

In the computation of material and opinion dissimilarities, we focus on the important participants in the news story. Currently we focus on the participants who are persons in the news. After extract the participants related segments and sentences from the news video and text (closed captions), we compute the dissimilarities as follows.

- Material Dissimilarity

The first question when comparing the material of two multimedia news item is how to select and construct their features. Video clips in news items offer not only the streams of the fix-sized pixel information but also the appearance of the concepts (objects shown in the frames). Furthermore, there are also abundant connotations included in the appearing patterns and particulars that make the differences between two items. In our method, we compare the two news items by using the visual descriptions presented related to the important participants to compute the material dissimilarity.

We think that both of the frequency (exposure degree) and appearing sequence in the video are important in the presentation. Different appearing patterns indicate different enumerate and cause different effects. We also detected the announcer-appearing segments because announcer always appears with brief introduction in the beginning of the news. The announcer-appearing segments require different treatment when comparing the materials.

Our method of calculating material dissimilarity is as follows.

– For each segment in the video, we make an important participants set S_p consisting of participants(persons) extracted from the contents. p is the number of the target segment. If there is no participants extracted, $S_p = \emptyset$.

– As shown in Figure 4, we transfer each video to a character string as a chain of the repeated Sid_p. Sid is an character identifying the different participant set. Same participant set shares the same Sid in all the news items reporting a same event. If there is no participant extracted in the segment p, Sid_p = "N". Because of the announcer faces extracted in the segments are different from each other because of the different post-casters, we define Sid_p = "−" if in the segment p there is no participant but an announcer making the introduction. The repeating counts rely on the frame amount of each segment with settled time interval (plus 1 per 500ms in or current implement). The character string of each video is the ordered chain of the repeated Sid_p of the segments. Our comparison of two videos is actually the comparison of the two transferred character strings.

– The material dissimilarity $diss^M(c_1, c_2)$ of items c_1 and c_2 is calculated as follows.

$$diss^M(c_1, c_2) = \frac{|LCS(c_1, c_2)|}{ll(c_1, c_2)} \tag{5}$$

where $LCS(c_1, c_2)$ is the Longest Common Subsequence [1] of multimedia news c_1 and c_2. ll is the longest length of the character string of the two videos. Respectively, $0 \leqq diss^M(c_1, c_2) \leqq 1$.

- Opinion dissimilarity

Similar with how we compute the material dissimilarity, opinion dissimilarity is calculated by using the participant-related descriptions in the contents. To compute the opinion dissimilarity, we use a method of extracting descriptive polarities from the sentences and make feature vectors by using the sentiment polarities of participant-related descriptions in the event.

First we use CaboCha [4] to build a grammar tree of each sentence in the text. If any of the nodes in the tree contains the name of a participant s, the sub-tree started from this node's parent would be the s-related phrase. For each s-related phrase in the news item, we use a sentiment dictionary evaluating each word with positive, negative and objective scores [8]. We multiply the score of nodes in main clause (node where s exists, sibling nodes and parent node) with coefficient = 1 and multiply the polarity of rest with coefficient = 0.5. After adding the products up, for each multimedia news c, we can extract the polarity scores as follows.

$$\begin{aligned} pls = &(pol(s_1, P), pol(s_1, N), pol(s_1, O), \\ &..., pol(s_n, P), pol(s_n, N), pol(s_n, O) \end{aligned} \tag{6}$$

where $pol(s, P)$, $pol(s, N)$, $pol(s, O)$ are the positive, negative and objective sum of the textual description on the participant in a multimedia news.

Because we compute the polarity scores of each participant in item by adding up the scores of the descriptive semantic words in the phrases, standardization is required for the large scores caused by the long texts.

$$diss^O(c_1, c_2) = \sqrt{1 - \frac{(pls(c_1) \cdot pls(c_2))^2}{|pls(c_1)|^2 |pls(c_2)|^2}} \tag{7}$$

where $pls(c_1)$ and $pls(c_2)$ are polarity value vectors of news items c_1 and c_2 respectively.

図 4　Comparing Two Multimedia News in Material.

Material and opinion in the same multimedia news interact with each other when reporting a news. Because of the synchronism between video and text in multimedia news, we think the opinions are expressed closely connected with the materials. Considering that the textual descriptions always work as the complementary or the conclusion of the visual descriptions, we think there is relationship between material and opinion.

In this paper, we consider the complementary information of both material and opinion. In our methods, we can find two different complement relationships by checking if they are 1) similar in opinion but different in material, or 2) similar in material but different in opinion.

When computing the typical score of news item, we also consider the relationship between material and opinion in the news.

In the example of "NPYD, Wall Street, Occupy", supposing

（1） n_α indicates the news with "'Occupy Boston' is interviewed (material) + 'Occupy Boston' fears the rude raids" (opinion),

（2） n_β indicates the news with "'Occupy Wall Street' are fighting with the police (material) + a protester said 'They cannot pull wool over our eyes' after the police twice pepper-sprayed him in the face (opinion)".

Here, n_α and n_β have dissimilar materials and similar opinions in the contents. In the complementary tree, if n_α is a descendant of n_β (vice visa), $n_\alpha \in st^L_\beta$. Furthermore, if n_α is the most typical news item in st^L_β, n_α will be the left child node of n_β.

Supposing n_k is the new node we want to inserted to the tree. After determining the corresponding ST (Section: 4. 1) and decide the root node in ST (Section: 4. 2), we call the iteration of clustering the complementary news items and deciding the most typical one in two sub-trees (Section: 4. 3).

Specific processes are as follows.

4. 1 Select Corresponding *ST*

Because different STs in the tree indicate different reporting coverage, the first step is to determine which sub-tree n_k belongs to.

Supposing the set of our target news items (the relevant multimedia news of query) is the "topic" of news, we can cluster the news into ST in which the news are not only sharing keywords but also sharing some of the materials or opinions. News in same ST should have similar materials or similar opinions. As shown in Figure 5, we call the purple zone as "event" that multimedia news in the same "event" will be constructed into a same ST. News on same event possibly report a series of news that the complementary information between them are more valuable to the user.

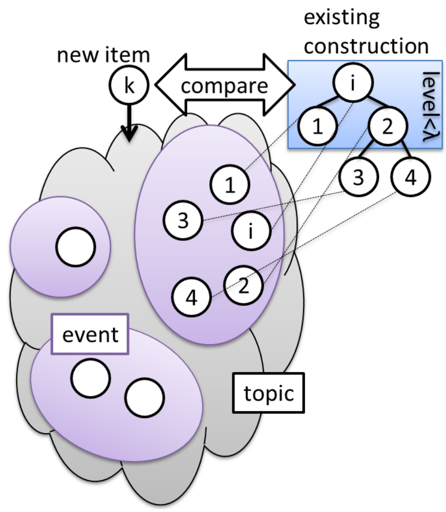


図 5 Selecting Corresponding Event of n_k.

As shown in Figure 5, we compare n_k with the typical nodes in existing ST.

$$diss^M_m(n_k, N(i,\lambda)) = Min(diss^M(n_k, n_p)) \tag{8}$$

$$diss^O_m(n_k, N(i,\lambda)) = Min(diss^O(n_k, n_p)) \tag{9}$$

where, $N(i,\lambda)$ is the set of nodes which belong to sub-tree ST_i and its level $< \lambda$. $n_p \in N(i,\lambda)$.

Supposing θ_{st} is the threshold for the difference between n_k and the representative node in each ST. If for all the existing ST_i, $diss^M_m(n_k, N(i,\lambda) + diss^O_m(n_k, N(i,\lambda)) > \theta_{st}$, n_k will be inserted to a new sub-tree ST_k. Otherwise, we will insert n_k into ST_i where $diss(n_k, N(i,\lambda))$ is the smallest of all.

4. 2 Decide the Root of *ST*

If the determined ST is not a new-created one, we will select the most typical news item as the root of ST. Two reasons are:

（1） for recommending the most typical news item (which can be regarded as the most representative in ST),

（2） for reducing the possibility of disconnected node in ST.

A typical score t_s will be calculated for the selection. Because of the new node n_k, all the t_s should be re-computed. If the re-computed t_s shows that the most typical item is not the formal one, all the nodes will be re-inserted. t_s of a node p is computed as follows.

$$t_s(p) = (Maj^M_p + Maj^O_p) \cdot (div^M_p + div^O_p) \tag{10}$$

$$maj^M_p = \overline{1 - diss^M(n_p, n_{q^M})} \cdot \frac{|G^M_{p,I}|}{|n|} \tag{11}$$

$$div^M_p = \overline{diss^O(n_p, n_{q_M})} \tag{12}$$

$$maj^O_p = \overline{1 - diss^O(n_p, n_{q^O})} \cdot \frac{|G^O_{p,I}|}{|n|} \tag{13}$$

$$div^O_p = \overline{diss^M(n_p, n_{q_O})} \tag{14}$$

where, $G^M_{p,I}$ is the node set consisting of node p and other nodes sharing similar materials with n_p in set I. $G^O_{p,I}$ is the node set consisting of node p and other nodes sharing similar opinions with n_p in set I. Here, I is the set of all nodes in ST. $n_p, n_{q^M} \in G^M_{p,I}$, $n_p, n_{q^O} \in G^O_{p,I}$. $|G^M_{p,I}|$ and $|G^O_{p,I}|$ are the counts of nodes in the group. $|n|$ is the count of all nodes.

4.3 Call the Iteration of Clustering the Complementary News and Deciding the Most Typical News

After deciding the root of ST, if n_k is the new root, all the other nodes in ST will be re-inserted according to the published time. Otherwise, we call the iteration as shown in Figure 6. Supposing current $st_j \leftarrow ST_i$, $n_j \leftarrow n_i$.

- Step 1:

Compare n_k with n_j to decide the complementary type of n_k to n_j.
If n_k provides detailed information, n_k belongs to st^L_j.
If n_k provides broaden information, n_k belongs to st^R_j.

- Step 2:

For all the news in st^L_j or st^R_j where n_k is, compute the typical score in the st.
Here, typical score t_s is computed where I is the set of multimedia news in st^L_j or st^R_j.

- Step 3:

If n_k is the most typical news in the sub-tree st^L_j or st^R_j, n_k will be the child of n_j. All the other nodes in st^L_j or st^R_j will be re-inserted.
Otherwise, supposing n_{next} is the most typical node, let $n_j \leftarrow n_{next}$, $st_j \leftarrow st_{next}$ and back to step 1.
Here, all the typical score t_s are computed among the news items in the current sub-tree, as shown in Equation 10. I is the set of multimedia news items in the sub-tree corresponding to node n_j.

In the system we first compute the dissimilarities of any two of items in the topic for reducing cost and time.

5. Experimental Plan

We prepared the experiments of using complementary tree for user recommendation. After collecting the multimedia news reporting a same topic from Google Video and Youtube, we generate the complementary tree and provide user watching sequence (seen as the path in the sub-tree) for news understanding and complementary information. Different relations between news items will be highlighted so the user can choose the preferred news.

6. Conclusion and Future Work

We proposed a complementary tree organizing the multimedia news items. In the tree, the news items are located considering their relative typical degrees and complementary to ancestors. During the construction, we compare the news items considering if they provide detailed information or diverse viewpoints. Our further work is to propose the application and carry our the experiments.

7. Acknowledgements

This research is partly supported by the Scientific Research Grant (No.20300042, No.20300036) made available by MEXT, Japan.

文 献

[1] T. Jiang and M. Li, "On the Approximation of Shortest Common Supersequences and Longest Common Subsequences", *Proc. of 21st Int. Colloquium on Automata, Languages and Programming*, Lecture Notes in Comput. Sci., vol. 820, pp. 191–202, 1994.

[2] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd "The PageRank Citation Ranking: Bringing Order to the Web", Technical Report, Stanford InfoLab. 1999.

[3] Ichiro Ide, Hiroshi Mo, and Norio Katayama "News Video Topics", Proc. of Int. Workshop on Multimedia Information Retrieval, pp. 239–246, 2003.

[4] K. Taku. "CaboCha: Yet Another Japanese Dependency Structure Analyzer [online]", *Technical report*, Nara Institute of Science and Technology, 2004.

[5] Ichiro Ide, Hiroshi Mo, Norio Katayama, and Shinichi Satoh "Threading for Structuring a Large-scale News Video Archive", Proc. of Int. Conf. on Image and Video Retrieval, LNCS 3115, pp. 123–131, 2004.

[6] Qiang Ma, Akiyo Nadamoto, and Katsumi Tanaka "Complementary Information Retrieval for Cross-media News Content", Information Systems, Vol. 31(7), pp. 659–678, 2006.

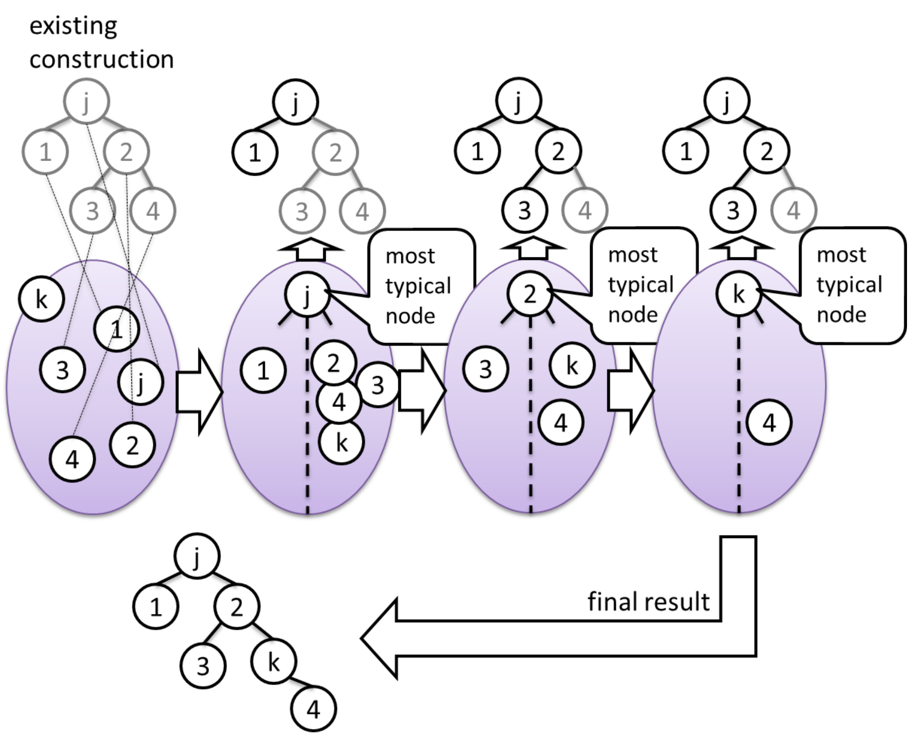
[7] Yin Yang, Nilesh Bansal, Wisam Dakka, Panagiotis Ipeirotis, Nick Koudas, and Dimitris Papadias "Query by document", Proc. of the Second ACM International Conference on Web Search and Data Mining, pp. 34–43, 2009.

[8] S. Ishida, "Comparative Analysis of News Agencies by Extracting Entity Descriptions and Its Application", *Thesis Paper of Master Course*, 2011.

[9] Yuanhua Lv, Taesup Moon, Pranam Kolari, Zhaohui Zheng, Xuanhui Wang, and Yi Chang "Learning to Model Relatedness for News Recommendation", Proc. of the 20th international conference on World wide web, pp. 57–66, 2011.

[10] Ling Xu, Qiang Ma, and Masatoshi Yoshikawa "Credibility-Oriented Ranking of Multimedia News Based on a Material-Opinion Model", Proc. of WAIM2011, LNCS 6897, pp. 290–301, 2011.

[11] Amr Ahmed, Alex Smola, Choon Hui Teo and S. V. N. Vishwanathan "Fair and Balanced: Learning to Present News Stories", Proc. of WSDM2012, LNCS 6897, pp. 333–342, 2012.



]

図 6 Iteration of Inserting n_k.