Estimation of the Representative Story Transition in a Chronological Semantic Structure of News Topics

Kosuke Kato
Nagoya University, Graduate School of Information Science
Furo-cho, Chikusa-ku, Nagoya, 464-8601 Japan
katok@murase.m.is.nagoya-u.ac.jp

Ichiro Ide
Nagoya University, Graduate School of Information Science
Furo-cho, Chikusa-ku, Nagoya, 464-8601 Japan
ide@is.nagoya-u.ac.jp

Daisuke Deguchi
Nagoya University, Info. & Comm. Headquarters
Furo-cho, Chikusa-ku, Nagoya, 464-8601 Japan
ddeguchi@nagoya-u.jp

Hiroshi Murase
Nagoya University, Graduate School of Information Science
Furo-cho, Chikusa-ku, Nagoya, 464-8601 Japan
murase@is.nagoya-u.ac.jp

ABSTRACT

It is important to track the flow of topics to thoroughly understand the contents. Accordingly, a method that structures the chronological semantic relations between news stories, namely a "topic thread structure" has been proposed. It allows the comprehensive understanding of a topic by chronologically tracking stories one by one from the initial story. However, this task imposes a user to watch many stories when it contains various sub-topics. Thus, we propose a method that estimates the representative story transition in a topic thread structure. In the proposed method, features obtained from a story and those from the topic thread structure are used for the estimation. We confirmed the effectiveness of the proposed method by comparing the results obtained from the proposed method to the ground truth obtained from votes in a subjective experiment.

Categories and Subject Descriptors

H.3.1 [Information Systems Applications]: Abstracting methods

General Terms

Algorithms

Keywords

News video, news video archive, topic thread structure

ICMR'14 , April 01 - 04 2014, Glasgow, United Kingdom
Copyright 2014 ACM 978-1-2782-4/14/04 ...$15.00.

1. INTRODUCTION

Recently, the diffusion of large capacity of storage has enabled us to easily store a large amount of broadcast videos. Among various kinds of broadcast videos, news videos are important information sources because they report events in the real world. When we consider news videos archived during a long period as a target, it is important to track the flow of topics to understand the news contents comprehensively. However, it is very difficult to manually track news videos related to an event comprehensively when they are stored in a large amount.

In regard to the described background, researches that analyze the semantics of news videos and construct a topic structure have been conducted. Duygulu et al. [1] proposed a method that stories[1] related to a certain topic are linearly connected in chronology order. Wu et al. [2] proposed a method that constructs a binary tree corresponding to the change of a subject and context of news stories in a cluster which consists of stories related to a certain topic. However, these methods do not represent parallel flows in a topic structure. Meanwhile, Ide et al. [3] proposed a method that represents semantic relations between news stories as a graph structure called a "topic thread structure", so that it can express parallel flows in a topic structure. However, some topic thread structures have many branches because it contains various sub-topics that lead to different endings. In such case, a user has to track many topic threads to understand the news contents, even if tracking one representative topic thread would have served the purpose.

In this paper, we propose a method to estimate the representative story transition in a topic by choosing a topic thread in a topic thread structure. It consists of two phases: 1) The estimation of a representative ending story in a topic thread structure. 2) The estimation of the most representative topic thread in case there are multiple topic thread candidates after Phase 1.

[1] A semantically minimum unit in a news video.

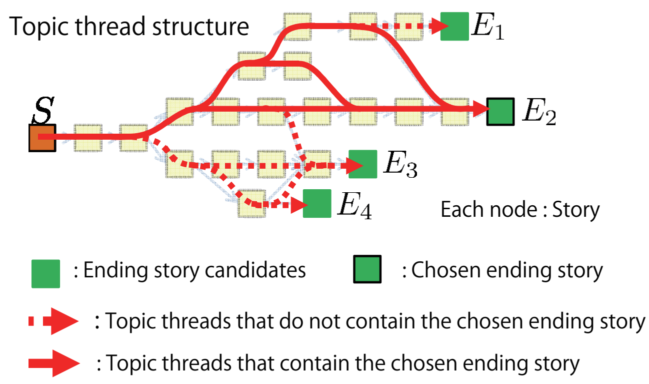

Figure 1: Example of a topic thread structure with multiple ending stories.

This paper is organized as follows. In Section 2, we describe the proposed method. In Section 3, we report the result of an experiment. Finally, we conclude the paper in Section 4.

2. ESTIMATION OF THE REPRESENTATIVE STORY TRANSITION

2.1 Approach

We solve the task of estimating the representative story transition for a given topic by selecting a topic thread in a topic thread structure. Here, we assume that a topic thread structure is given by Ide et al.'s method [3]. A path that connects the initial story (S) to any ending story candidate (E_i) is called a topic thread. The proposed method estimates the representative topic thread in two phases. The first phase estimates the ending story from the ending story candidates in a topic thread structure. Figure 1 shows an example of this phase. The second phase estimates the representative topic thread from the multiple (if any) topic threads selected in the first phase. Figure 2 shows an example of this phase.

2.2 Phase 1: Choosing an ending story

The following five features are used as features to choose the ending story. Features 4 and 5 were proposed by Sawai et al. in [4] for a different purpose.

1. Similarity of proper nouns in the initial and ending stories

   We assumed that the core people, organizations, areas and so on of a topic are consistent throughout the representative story transition of a specific topic. Here, we refer to the first α (= 3 in the experiment) sentences from the closed caption accompanying the initial story and the ending story candidates. First, morphological analysis is applied to these sentences. Next, proper nouns and unknown words are extracted and a frequency vector of these words are constructed for each story. Finally, we calculate the cosine similarity between the initial story and each ending story candidate. The similarity of proper nouns $v_{\mathrm{relevance},i}$ is defined as Eq. (1), where $\boldsymbol{W}_S$ is a word frequency vector of the initial story S and $\boldsymbol{W}_{E_i}$ is that of the i-th ending story candidate E_i.

$$\begin{cases} v_{\mathrm{relevance},i} = \dfrac{\sum_{j=1}^{n} W_{S,j} W_{E_i,j}}{\sqrt{\sum_{j=1}^{n} W_{S,j}^2}\sqrt{\sum_{j=1}^{n} W_{E_i,j}^2}} \\ \boldsymbol{W}_S = (W_{S_1}, \cdots, W_{S_n}) \\ \boldsymbol{W}_{E_i} = (W_{E_{i,1}}, \cdots, W_{E_{i,n}}) \end{cases} \quad (1)$$

2. Elapse of days between the initial story and ending stories

   We assumed that a topic thread whose elapse of days from the initial day shorter than that of the others, is not the main flow. The elapse of days $v_{\mathrm{elapse},i}$ is defined as Eq. (2), where the date of the initial story S is D_S and that of the i-th ending story candidate E_i is D_{E_i}

$$v_{\mathrm{elapse},i} = D_{E_i} - D_S \quad (2)$$

3. Length of a topic thread

   We assumed that a representative topic thread has a longer path length between the initial story and the ending story than that of the others. The length of a topic thread $v_{\mathrm{length},i}$ is defined as Eq. (3), where $L_{i,j}$ is the number of stories in the j-th topic thread which starts from the initial story S and ends at the i-th ending story candidate E_i.

$$v_{\mathrm{length},i} = \max_j (L_{i,j}) \quad (3)$$

4. Importance of the ending story by its broadcasted order

   The importance of the broadcasted order $v_{\mathrm{order},i}$ is defined as Eq. (4), where N_{T_i} is the number of stories broadcasted in the news show which contains the i-th ending story candidate E_i, and L_{T_i} is its broadcasted order.

$$v_{\mathrm{order},i} = \frac{N_{T_i}}{L_{T_i}} \quad (4)$$

5. Importance of the ending story by video length

   The importance of a video length $v_{\mathrm{interval},i}$ is defined as Eq. (5), where I_{E_i} is the video length of the i-th ending story candidate.

$$v_{\mathrm{interval},i} = T_{E_i} \quad (5)$$

The ending story of a representative topic thread is estimated by evaluating these features comprehensively. These features are standardized so that their average equals to 0 and their variance equals to 1. The score for the i-th ending story candidate E_i is defined as Eq. (6), where m_j is a weight for a standardized feature v_j and $\sum_j m_j = 1$.

$$\begin{cases} \mathrm{Score}_i = \sum_j m_j \cdot v_{j,i} \\ \sum_j m_j = 1 \end{cases} \quad (6)$$

Here, m_j is determined by learning. The ending story candidate with the highest score is chosen as the ending story for the representative topic thread.

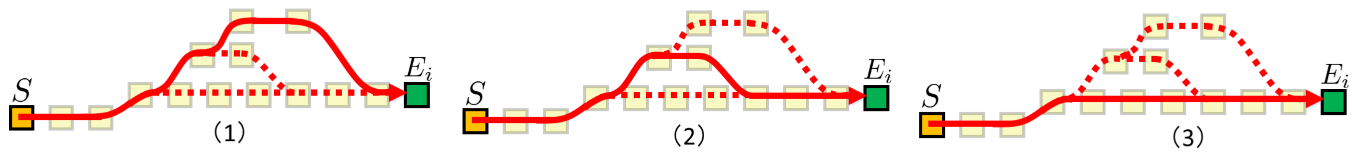

Figure 2: Example of multiple topic threads connecting an initial story S and the ending story E_i. One of them is selected as the representative topic thread.

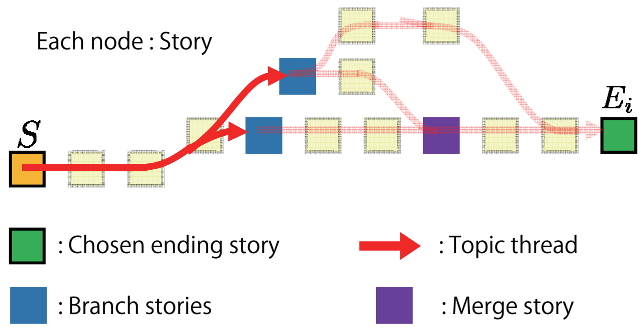

Figure 3: Replacing ending stories with branch stories in Phase 2.

2.3 Phase 2: Solution of the multi-path problem

As shown in Fig. 2, when there are multiple topic threads (paths) that connect the initial story and the chosen ending story in Phase 1, we need to choose one of them as the representative topic thread. This is called the multi-path problem. The multi-path problem is solved in the same way as in Phase 1. However, instead of applying the method to ending stories, it applies it to branch story candidates. As shown in Fig. 3, a branch story is the first story after a branch in the topic thread structure. A merge story is a story where multiple topic threads merge. Accordingly, the length of a (sub-)topic thread is redefined as follows.

- Length of a (sub-)topic thread

  The length of a (sub-)topic thread $v'_{\mathrm{length},i}$ is defined as Eq. (7), where L_i is the length from the i-th branch story to a merge story.

$$v'_{\mathrm{length},i} = L_i \tag{7}$$

3. EXPERIMENT AND DISCUSSIONS

3.1 Dataset

In this experiment, we used stories obtained from a Japanese daily news show "NHK News7" broadcasted during March 16th, 2001 and December 31st, 2012 as inputs, and twelve topic thread structures constructed from them. Table 1 shows their statistics.

Table 1: Statistics on the topic thread structures used in the experiment.

	Avg.	Min.	Max.
Number of topic threads	7.8	4	19
Length of a topic thread [stories]	7.0	2	24
Number of stories	15.8	10	31
Number of ending story candidates	4.4	3	6

3.2 Creation of ground truth

In order to obtain the ground truth both for the learning of the weights m_i in Eq. (6) and for the evaluation, we performed a subjective experiment. The subjects were 16 students all specializing in computer science in their twenties. Their average age was 22.9 and their variance was 0.72. They consisted of 15 men and 1 woman. The subjects were asked to choose a representative topic thread in two phases.

Phase 1: The subjects were only shown the initial story and the ending story candidates and asked to choose one ending story.

Phase 2: The subjects were asked to choose the most representative topic thread if there was a multi-path problem in the topic thread structure. There were twelve such cases in the experiment.

As a result of the subjective experiment, we obtained the most representative ending story for each topic thread structure for Phase 1, and the most representative (sub-)topic thread for Phase 2.

3.3 Evaluation methods using individual features

We evaluated both the ending story chosen in Phase 1 and the topic thread chosen in Phase 2.

For Phase 1, we checked the correspondence of the most voted ending story candidate in the subjective experiment with the ending story that obtained the highest score by the proposed method.

For Phase 2, we checked the correspondence of the most voted branch story in the subjective experiment with the branch story that obtained the highest score by the proposed method.

These correspondences were evaluated by both three-fold cross validation and the R method (Re-substitute method). We compared the accuracy of the proposed method with that using individual features alone.

Table 2: Accuracy of each method in Phase 1.

Method	Accuracy
Proposed (R method)	67% (8/12)
Proposed (Avg. of 3-fold cross valid.)	50% (2.0/4)
Similarity of proper nouns	42% (5/12)
Elapse of days	42% (5/12)
Length of a topic thread	17% (2/12)
Broadcasted order	17% (2/12)
Video length	33% (4/12)

Table 3: Accuracy of the proposed method in Phase 1 by 3-fold cross validation.

Test dataset	1	2	3
Accuracy	50% (2/4)	75% (3/4)	25% (1/4)

Table 4: Weights (m_i) that maximized the accuracy for each dataset and those learned by the R method.

	$m_{\text{relevance}}$	m_{elapse}	m_{length}	m_{order}	m_{interval}
Dataset 1	0.43	0.15	0.01	0.00	0.41
Dataset 2	0.03	0.47	0.28	0.00	0.22
Dataset 3	0.35	0.33	0.01	0.01	0.30
R method	0.37	0.25	0.03	0.01	0.34

3.4 Experimental results and discussions

3.4.1 Phase 1

Table 2 shows the accuracy of each method in Phase 1, and Table 3 shows that of the proposed method by three-fold cross validation for each dataset.

The proposed method using the R method and the three-fold cross validation showed higher accuracy than those by the comparative methods using individual features alone. Thus we confirmed the effectiveness of combining the features.

Table 4 shows the learned weights for each feature in the case of the R method and also the same datasets as those in Table 3. We can see the weights varied according to each dataset. For example, the weights of the proper nouns and the video length for Dataset 1 were higher than those for the others. From this result, we can infer that Datasets 2 and 3 contained ending stories where these features were important. We considered that the optimum weights for each feature vary according to the type of a topic. The estimation accuracy may be improved if the weights can be switched adaptively to a topic category, for example, the importance of video length in politics is high and that of thread length in an accident is high.

3.4.2 Phase 2

Table 5 shows the results of each method in Phase 2. Here again, the proposed method using the R method and the three-fold cross validation showed higher accuracy than those of the comparative methods using individual features alone.

From Tables 2 and 5, we can see the accuracy using proper nouns in Phase 1 was high, that in Phase 2 was lower. One reason for this is that the similarity of proper nouns becomes

Table 5: Accuracy of each method in Phase 2.

Method	Accuracy
Proposed (R method)	92% (11/12)
Proposed (Avg. of 3-fold cross valid.)	92% (3.7/4)
Similarity of proper nouns	25% (3/12)
Elapse of days	67% (8/12)
Length of a thread	42% (5/12)
Broadcasted order	50% (6/12)
Video length	50% (6/12)

0 after standardization when they are equal. Another reason is the lack of proper nouns due to short news videos. It is necessary to change the number of extracted sentences α or to analyze other stories than the initial story and the ending stories.

4. CONCLUSIONS

We proposed a method to estimate the representative story transition of a news topic by choosing the representative topic thread from a topic thread structure. The proposed method consisted of two phases. In Phase 1, it chose the representative transition of news stories from the ending story candidate. In Phase 2, it chose the representative topic thread if the result from the first phase involved multiple paths.

We conducted an experiment using the votes by subjects as ground truth and confirmed the effectiveness of the proposed method.

Future work is to improve the proposed method by changing weights adaptively depending on the topic category. As an application, we are considering to automatically edit a summarized video by selecting video clips from a representative story transition so that we can understand about a topic effectively.

5. ACKNOWLEDGMENTS

We would like to thank the National Institute of Informatics, Japan for allowing us the usage of the news videos in the NII TV-RECS archive. Parts of this research were supported by MEXT, Grant-in-Aid for Scientific Research.

6. REFERENCES

[1] P. Duygulu, J. Pan, and D.A. Forsyth, “Towards auto-documentary: Tracking the evolution of news stories”, Proc. 12th ACM Int. Conf. Multimedia, pp.820–827, Oct. 2004.

[2] X. Wu, C. Ngo, and Q. Li, “Threading and autodocumenting news videos”, IEEE Signal Processing Mag., Vol.23, No.2, pp.59–68, Mar. 2006.

[3] I. Ide, T. Kinoshita, T. Takahashi, H. Mo, N. Katayama, S. Satoh, and H. Murase, “Efficient tracking of news topics based on chronological semantic structures in a large-scale news video archive”, IEICE Trans. Information & Systems, Vol.E95-D, No.5, pp.1288–1300, May 2012.

[4] R. Sawai, H. Senoo, and Y. Shishikui, “Proposal and evaluation of a method for calculating news value for creating news digest (In Japanese)”, IPSJ Trans. Databases, Vol.2, No.2, pp.158–172, Jun. 2009.