Automatic Summarization of Multiple Travel Blog Entries Focusing on Travelers' Behavior

Shumpei Iinuma,
Hidetsugu Nanba, and
Toshiyuki Takezawa
Graduate School of Information Sciences, Hiroshima City University, Hiroshima, Japan
{iinuma, nanba, takezawa}@ls.hiroshima-cu.ac.jp

Abstract

The evolution of information and communication technology now makes it possible to collect travel information in a variety of ways. Social media content that includes blogs is one such useful information source when planning a trip. In this study, we propose a method for generating a summary of multiple travel blog entries that contain images. Our method identifies significant sentences in addition to the images by using a graph-based approach that takes account of travelers' types of behavior. To investigate the effectiveness of our method, we conducted experiments, which demonstrated that our method can outperform some baseline methods. We also implemented a system for generating summaries based on our method.

Keywords: Travel Blog; Multimedia Summarization; Travel Information Processing.

1 Introduction

Travel guidebooks are a useful information source about travel. Guidebooks give basic information about tourist spots, souvenirs, restaurants, and hotels. However, social media content, particularly travel blogs, are another, more recent, information source that provides many bloggers' experiences of travel destinations, tourist spots, and hotels.

Various researchers have investigated travel blogs as an information source for travel. For example, Nanba et al. (2009) proposed a method that can identify travel blog entries automatically from blogs using machine learning technology. Fujii et al. (2016) proposed a method of classifying travel blog entries into the five categories (or content types) shown in Table 1. However, these methods are not useful for travel planning if there are many travel blogs related to the intended destination, because it would take too long to read all of them.

In this paper, we propose a method that summarizes a set of travel blog entries about a destination. Although there are many studies about text summarization, which we will describe in Section 2.2, a notable difference between previous studies and our approach is that our method focuses not only on texts but also on images. Our summarization method is based on LexRank (Erkan et al., 2004), which were proposed for summarizing texts. Main contribution of our work is to expand the LexRank to summarize not only texts but also images. Potential travelers using our system can quickly get essential information about a destination from our system's output summary, which contains both text and image information.

The remainder of this paper is organized as follows. Section 2 discusses related work. Section 3 describes our method. To investigate the effectiveness of our method, we conducted experiments whose results are reported in Section 4. Section 5 shows the system behavior in terms of snapshots. We present some conclusions in Section 6.

Table 1. Content types and their descriptions

Content Type	Description
Watch	Blog entry about sightseeing at tourist spots
Experience	Blog entry about an experience such as scuba diving or dancing
Buy	Blog entry about shopping or souvenir stores
Dine	Blog entry about drinking and dining
Stay	Blog entry about accommodation

2 Related Work

2.1 Travel Information Recommendation

Wu et al. (2008) proposed a system that searched and summarized tourism-related information. When a user (traveler) entered a query, such as “What is the historical background of Tian Tan?” the system searched for and obtained information from Wikipedia, Flickr, YouTube, and official tourism Web sites using the tourist spot name as a query. Their system also classified the query as belonging to one of five categories, namely “general,” “history,” “landscape,” “indoor scenery,” and “outdoor scenery,” to provide users with more relevant information. For example, if a query is classified as belonging to the “history” category, the information is obtained from texts, whereas a query regarding “outdoor scenery” obtains its information from photos and videos. However, even if a query is classified to “history” category, showing a text with images as an answer might be easier to understand than just showing a text. Therefore, we construct a system that can generates summaries comprising multiple sentences and images from a set of travel blog entries.

Hao et al. (2010) proposed a method for mining location-representative knowledge from travel blogs based on a probabilistic topic model (the Location-Topic model). Using this model, they developed three modules, namely a destination-recommendation module for flexible queries, a characteristics-summarization module for a given destination (with representative tags and snippets), and an identification module for informative parts of a travel blog that enriched the recommendations with related images. However, the output summaries do not always match with images, because this system extracts them from different sources. We propose a summarization method that takes account of association between texts and images.

2.2 Text Summarization

Text summarization is a method that identifies important information in a text (or multiple texts), and shows the results as a summary text. Text summarization has been studied since Luhn (1958) and has become a hot topic in the field of natural language processing. Typically, the traditional approach to text summarization is to identify important sentences from a text (or multiple texts) and output them as the result of the summarization. Several methods have been proposed to identify important sentences in texts. We now describe one such well-known method, called LexRank (Erkan et al., 2004), whose effectiveness has been confirmed in other text-summarization research, including the Text Analysis Conference (TAC) (https://tac.nist.gov/), an evaluation workshop for text summarization. In our work, we also adopt LexRank.

LexRank calculates the importance of each sentence based on the idea of centrality in eigenvectors and creates a similarity graph. Figure 1 is an example of such a graph.

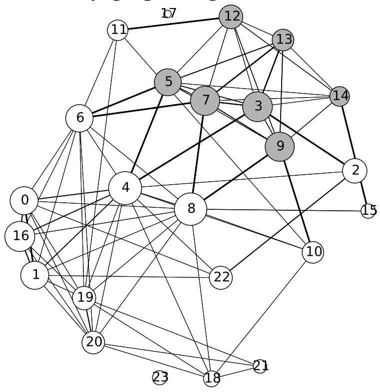

Fig. 1. An example of a similarity graph

In the graph, each node indicates a sentence, with the number in each node identifying the node. If the similarity between two sentences exceeds a threshold value, the two nodes are linked, as shown by the edges in the figure. From this graph, the importance of each node u (a sentence) is calculated using the following equation (1). This equation is similar to the PageRank equation (Brin and Page, 1998).

$$p(u) = \frac{1-d}{N} + d \sum_{v \in adj[u]} \frac{p(v)}{\deg(v)}, \qquad (1)$$

where N indicates the number of nodes (sentences), d is a "dumping factor" (Brin and Page, 1998), *adj[u]* is the set of nodes linked to node u, and *deg(v)* is the order of node v. As for PageRank, sentences that have many links to others tend to obtain higher page-rank scores. Erkan et al. proposed Continuous LexRank, which is an extension of LexRank. LexRank treats all links between sentences equally, whereas Continuous LexRank employs a similarity value between sentences as a weight for each link. Continuous LexRank is defined by equation (2).

$$p(u) = \frac{1-d}{N} + d \sum_{v \in adj[u]} \frac{weight(u,v)}{\sum_{z \ni adj[v]} weight(z,v)} p(v) \qquad (2)$$

In our work, we expand LexRank to generate a summary that comprises a text and images. If we can associate texts with images, we can calculate the importance of each sentence and image at the same time. We consider that the two sentences appearing before and after an image are related to that image, and create one graph from both sentences and images. We then apply LexRank (or Continuous LexRank) to this graph to calculate the importance of each sentence and text. In applying the equation, we also employ a "biased factor" (Otterbacher et al., 2009), which we will describe in Section 3.

3 Summarization of Multiple Travel Blog Entries

Our system assumes that the travel blog entries to be summarized have been geotagged and classified into five categories using Fujii's method (Fujii et al., 2016) in advance. When a user specifies a content type and geographical region, our system generates a summary from all blog entries having the relevant content type within the region. A summary is generated by the following procedure.

- Cluster the travel blog entries to be summarized,
- Calculate the importance of each sentence and image for each cluster,
- Select three to five important sentences and images for each cluster.

Because the blog entries to be summarized might contain multiple topics, we conduct hierarchical clustering when grouping blog entries having a similar topic. We employ the furthest-neighbor method for the clustering. In the clustering process, we merge two clusters if the distance between these clusters is smaller than a threshold value. Here, we express each document as a vector of words, whose weights are calculated by tf*idf (term frequency-inverse document frequency). To calculate the distance between clusters c_i and c_j, we use the function $f(c_i, c_j) = 1-\cos(c_i, c_j)$, where $\cos(c_i, c_j)$ indicates cosine similarity.

3.1 Calculation of the Importance of each Sentence and Image Using LexRank

As discussed in Section 3.2, both LexRank and Continuous LexRank obtain the importance of each sentence by calculating the PageRank score for each node of a sentence-similarity graph. We expand these algorithms by calculating the importance of each sentence-and-image together, using a graph created from sentences and images. We consider that the two sentences appearing before and after an image are related to that image, and create one graph from sentences and images. In creating this concatenation graph, we consider the similarity score between an image and its adjacent sentences as 1. We then use the weights $a_{i,j}$ in equation (3) to calculate the weight between nodes i and j (sentence/image).

$$a_{i,j} = \begin{cases} sim(s_i, s_j) \ \ (type(s_i) = type(s_j)) \\ 1 \ \ (type(s_i) \neq type(s_j) \text{ and } |i-j| = 1), \\ 0 \ \ (\text{otherwise}) \end{cases} \quad (3)$$

where s indicates an element (a sentence or an image) in a blog-entry sentence and s_i indicates the i^{th} element in a blog entry. $type(s_i)$ indicates whether each element is a sentence or an image, and sim is a function that calculates the similarity between two elements. Here, we use different similarity functions for each element type. We will explain the calculation between elements in Section 3.2. In our example of a graph (see Figure 1), white and gray nodes indicate sentences and images, respectively. The size of each node indicates its importance, as calculated using equation (4) in Section 3.3.

3.2 Similarity between Sentences and Images

Each sentence is expressed as a vector, where each element of the vector is the tf*idf score for a word. We use cosine similarity for the calculation between vectors. Each image is expressed by two kinds of vector, namely a color histogram and a bag of visual words (BoVW). We calculate the similarity between two images for both types of vector, and take the average of the two similarity values as the similarity between the two images. The color histogram is created by projecting all elements in an image to an HSV color space divided into 160 areas (H, S, and V are divided into 10, 4, and 4, respectively) and then counting the number of elements in each area. The BoVW represents images by vectors of the frequency of appearance of local features. These features were obtained by extracting from images using an algorithm, such as SIFT (Lowe, 1999), and by clustering them. BoVW was originally applied in natural language processing as a "bag of words" (BoW) that represented documents by vectors of appearance frequency for its constituent words.

3.3 Expansion of LexRank Using Content Types

In each travel blog entry, there are sentences that have a strong relationship with a given content type, and we expect these sentences to be in the summary. To achieve this, we employ biased LexRank, which consists of the following two steps (Otterbacher et al., 2009). First, we calculate the degree of association of each sentence with a given content type. Second, we use these degrees to calculate the importance of each node using the following equation, which is an extension of equation (2).

$$p(u) = (1-d)\frac{typeScore(u)}{\sum_v typeScore(v)} + d \sum_{v \in adj[u]} \frac{weight(u,v)}{\sum_{z \ni adj[v]} weight(z,v)} p(v) \quad (4)$$

In equation (4), *typeScore(u)* indicates the degree of association of a sentence u with a given content type. *typeScore(w)*, as defined by equation (5), is calculated by the degree of association of each word in a sentence u with a given content type.

$$typeScore(w) = IG(w,t) \log\left(1 + \frac{count(w, D_t)}{count(w, D)}\right) \quad (5)$$

Here, *IG(w, t)* indicates information gain of word w for a content type t. *count(w, D)* indicates the number of blog entries that contain the word w in the set of travel blog entries D. D_t is the number of travel blog entries associated with content type t. If the word w often appears only in travel blog entries with content type t, values of both $IG(w, t)$ and $count(w, D_t)$ will become large, and as a result, $typeScore(w)$ will also become large. Now, we consider $typeScore(w)$, which has the highest value among all words in sentence u, as $typeScore(u)$.

3.4 Reduction of Redundancy

After applying LexRank to a sentence-similarity graph, there might be sentences having high importance scores that are very similar to each other. As a result, there would be redundancy in the summary generated. To resolve this problem, Radev et al. (2000) reranked sentences by taking account of their containment relationship. In the same way, our method rejects sentences and images that are sufficiently similar (their similarity values exceed a threshold value) to sentences or images that have already been chosen as part of the summary.

4 Experiments

4.1 Experimental Settings

We used the travel blog entries collected using Nanbas' method (Nanba et al., 2009), and the content types were assigned using Fujiis' method (Fujii et al., 2016). These entries and their content types were manually checked. We selected 20 areas in Japan, and chose approximately 10 blog entries for each area. We then created correct summaries manually by choosing three images and five sentences for each content type. Finally, we obtained 47 human-produced summaries, which we used in the evaluation of our system.

4.1.1 Evaluation

We conducted both automatic and manual evaluations.

Automatic evaluation:
For the evaluation of the text part of computer-produced summaries, we employed ROUGE-N (Lin, 2004), which is widely used as an evaluation metric in text-summarization research and projects such as TAC (https://tac.nist.gov). ROUGE-N is calculated by dividing the number of N-word-grams that are contained in both human-produced and computer-produced summaries by the number of word N-grams in a human-produced summary. That is, ROUGE-N is a metric of how well a computer-produced summary covers the word N-grams in a human-produced summary. For the values of N, we used N = 1 and N = 2.

For the automatic evaluation of the image part of computer-produced summaries, we used recall and precision as evaluation metrics. We asked human subjects to choose representative images for each topic, and used them for the evaluation of the image part of the computer-produced summaries.

Manual evaluation:
We evaluated computer-produced summaries manually from the following two viewpoints:

- Are the essential points of blog entries contained in the computer-produced summary? (MANUAL-TEXT).
- Do the images and texts in each computer-produced summary match? (MANUAL-IMAGE-TEXT).

In this evaluation, we asked human subjects to evaluate summaries according to the following procedure.

1. Read all blog entries to be summarized.
2. Read a human-produced summary and five computer-produced summaries (to be described later), evaluating in terms of a five-point scale (MANUAL-IMAGE-TEXT).
3. Rank six summaries (MANUAL-TEXT). (This ranking-based evaluation was employed in the evaluation workshop NTCIR-2 Text Summarization Challenge (Fukushima et al., 2002).) If the qualities of two summaries were considered the same, we allowed the human subjects to rank these summaries equally.

Alternatives:

We performed evaluations for the following five methods. For each method, we adjusted the length of the summary and the number of images to be the same as the human-produced summaries.

- **Lead (baseline)**: extract sentences and images from the head of each blog entry.
- **LR (baseline)**: construct a sentence similarity graph and an image similarity graph, and then apply LexRank to each graph (equation (2)).
- **LR+IMG (our method)**: construct one graph by connecting the sentence similarity and image similarity graphs, and apply LexRank (equations (2) and (3)).
- **LR+TYPE (our method)**: apply LexRank, while taking account of content types (equation (4))
- **LR+IMG+TYPE (our method)**: apply the LR+TYPE method to a similarity graph in the LR+IMG method.

In conducting hierarchical clustering, we used 0.9 as a threshold value for merging two clusters (see Section 3). We standardized on 16 pixels, and sampled every 8 pixels to obtain the SIFT features. Here, SIFT is an algorithm to detect and describe local features in images (Lowe, 1999). We then calculated the BoVW vectors by conducting K-means clustering, fixing 1,000 for the cluster numbers (see Section 3.2). For the extraction of SIFT features, we used OpenCV software (http://opencv.jp). We employed 0.85 as the dumping factor for LexRank (equation 2). To reduce the redundancy in computer-produced summaries, we used threshold values of cosine distance = 0.9 for sentences and 0.5 for images (see Section 3.4).

We obtained content-type-related words using 1,836 travel blog entries (purchase: 147, watch: 1,145, experience: 119, stay: 38, and dine: 693). None of these travel blog entries included entries that were used in the summarization.

4.2 Results and Discussion

The evaluation results using ROUGE (automatic evaluation of text parts) are shown in Table 2. Note that both LR+TYPE and LR+IMG+TYPE outperformed two of the baseline methods. We conducted a t-test, which confirmed a significant difference between our methods and these baseline methods.

Table 2. Evaluation results using ROUGE-N (automatic evaluation of text parts)

	ROUGE-1	ROUGE-2
Lead (baseline)	0.318	0.222
LR (baseline)	0.316	0.207
LR+IMG	0.331	0.227
LR+TYPE	**0.345**	**0.240**
LR+IMG+TYPE	0.340	0.237

Highest values among all systems are shown in bold.

We show another set of evaluation results (automatic evaluation of image parts) in Table 3. Here, our LR+IMG+TYPE method outperformed other methods, but we could not confirm the statistical significance of this result.

Table 3. Evaluation results using precision/recall (automatic evaluation of image parts)

	Precision	Recall
Lead (baseline)	0.351	0.359
LR (baseline)	0.341	0.341
LR+IMG	0.338	0.342
LR+TYPE	0.351	0.359
LR+IMG+TYPE	**0.372**	**0.367**

Highest values among all systems are shown in bold.

We show the results for manual evaluation in Tables 4 and 5. Table 4 gives the average rank of each summary by two human subjects, where a smaller average rank indicates a better summarization. Table 5 gives a five-point scale values of each method, where a larger value indicates a better summarization. We conducted a t-test ($p < 0.05$), which confirmed that there was a significant difference between our methods (LR+IMG, LR+TYPE, and LR+IMG+TYPE) and the Lead method.

Table 4. Evaluation results by human subjects (MANUAL-TEXT)

	Average rank
Human-produced summaries	1.28
Lead (baseline)	4.01
LR (baseline)	3.09
LR+IMG	**2.85**
LR+TYPE	3.22
LR+IMG+TYPE	2.99

The highest value among all systems is shown in bold.

Table 5. Evaluation results by human subjects (MANUAL-IMAGE-TEXT)

	Average rank
Human-produced summaries	4.33
Lead (baseline)	2.80
LR (baseline)	3.09
LR+IMG	**3.12**
LR+TYPE	2.96
LR+IMG+TYPE	3.05

The highest value among all systems is shown in bold.

Next, we compared our methods with LexRank (LR) for each summarization type. Tables 6 and 7 show the number of cases for which our methods are better/worse than LR for each content type in TEXT and IMAGE-TEXT evaluations, respectively. For example, "LR < LR+IMG" indicates that lR+IMG outperformed LR, and the shaded cells indicate where our method outperformed or matched LR. Table 6 shows that our LR+IMG method outperformed or matched LR for all content types. However, LR+TYPE could not match LR for any content type. Table 7 shows that all of our methods outperformed or matched LR for content type "Watch," whereas our methods could not match LR for content type "DINE." Through further investigation, we found that these results relate to the amount of text associated with each image.

Table 6. Comparison of LexRank (LR) with our methods for each content type (MANUAL-TEXT)

	Watch	Dine	Exp., Buy, Stay	Total
LR < LR+IMG	12	12	4	28
LR > LR+IMG	7	6	3	16
LR = LR+IMG	23	12	15	50
LR < LR+TYPE	7	9	2	18
LR > LR+ TYPE	14	9	6	29
LR = LR+TYPE	21	12	14	47
LR < LR+IMG+TYPE	13	15	7	35
LR > LR+IMG+TYPE	13	11	6	30
LR = LR+IMG+TYPE	16	4	9	29

Table 7. Comparison of LexRank (LR) with our methods for each content type (MANUAL-IMAGE-TEXT)

	Watch	Dine	Exp., Buy, Stay	Total
LR < LR+IMG	8	10	4	22
LR > LR+IMG	5	10	3	18
LR = LR+IMG	29	10	15	54
LR < LR+TYPE	6	6	2	14
LR > LR+ TYPE	4	11	7	22
LR = LR+TYPE	32	13	13	58
LR < LR+IMG+TYPE	12	9	6	27
LR > LR+IMG+TYPE	9	12	6	27
LR = LR+IMG+TYPE	21	9	10	40

Table 8. Comparison of LexRank (LR) with our methods for each content type, in terms of the number of characters associated with each image

	Less than 100 characters	Over 100 characters
LR < LR+IMG	18	10
LR > LR+IMG	10	6
LR = LR+IMG	30	20
LR < LR+TYPE	11	7
LR > LR+ TYPE	20	9
LR = LR+TYPE	27	20
LR < LR+IMG+TYPE	21	14
LR > LR+IMG+TYPE	22	8
LR = LR+IMG+TYPE	15	14

Using the same approach, we investigated the relationship between the amounts of text associated with images and the quality of summaries. Table 8 gives the results of a comparison between the LR method and our methods for different numbers of characters per image. From Table 8, we see that only the LR+IMG method can match the LR method when the number of characters is less than 100, whereas both the LR+IMG and LR+IMG+TYPE methods outperform the LR method when the number of characters is more than 100.

5 System Behavior

In this section, we introduce our system's behavior in terms of the travel blog entries collected and classified by the Nanbas' method (Nanba et al., 2009) and the Fujiis' method (Fujii et al., 2016), which were mentioned in the previous section. Figure 2 shows a map that summarizes multiple travel blog entries, as generated by our system. In this figure, blog entries are shown as icons. If we push one of the buttons "watch," "experience," "purchase," "dine," or "stay" (as listed in Table 1), the blog entries corresponding to this category are shown on the map. Clicking an icon on the map produces a list of the blog entries related to that point.

After clicking a button on the bottom left of the map, a computer-generated summary appears in a pop-up window, which summarizes all the blog entries shown in the map. The blog entries shown in the map will refer to the particular content type that the system user has specified. If the user chooses a different content type, the system will quickly generate another summary from the travel blogs related to that content type. Figures 3 and 4 are the computer-generated summaries for content types "watch" and "dine", respectively. Although both summaries were generated for the same location, namely "Miyajima" (one of the most famous tourist spots in Japan), the summaries refer to different aspects of a visit to Miyajima.

Fig. 2. Travel blog entries in a map for the content type "watch"

6 Conclusions

We considered that travel blog entries are useful information source for travel, because, they provide many bloggers' experiences of travel destinations. Therefore, we have proposed a method for summarizing multiple travel blog entries. Our method is an extension of LexRank to enable generation of a summary text containing images.

We conducted experiments that demonstrated that one of our methods, LR+IMG, can outperform baseline methods. Finally, we constructed a summarization system that can summarize multiple travel blog entries in terms of content types for any given geographical region. The system is available on our web site (http://165.242.101.30/blogMap/). Our system can generate a summary very quickly if the number of travel blog entries are less than 100. However, when the number of entries is over 100, our system will not generate a summary, because it is quite time-consuming. How to decrease the processing time is our future work.

- This is a firework festival that I wanted to watch before I moved to Hiroshima.
- I went to Miyajima firework festival on August 14.
- My proposal was rejected by my husband and my daughter, and finally we watched firework on a ship.

- Miyajima island is known as one of three most scenic spots in Japan, Itsukushima shrine, and world heritage.
- This is a shrine in Itsukushima (Miyajima island) in Hiroshima.
- In future, we might know the detail of Miyajima a thousand years ago.

Fig. 3. A summary generated from travel blogs on Miyajima for the content type “watch”

- Last February, our family attended Miyajima oyster festival.
- We went to “the fried oyster” line.
- We could eat grilled oyster (free), oyster with rice, fried oyster, udon noodle with oyster, and fresh seafood.

- We arrived at Miyajima port, but it does not look like a port.
- We arrived at Miyajima island by ferry. This island is famous for a world heritage, Itsukushima shrine.
- We arrived at Miyajima at 11:00 am.

Fig. 4. A summary generated from travel blogs on Miyajima for the content type “dine”

References

Brin, S. & Page, L. (1998). The Anatomy of a Large-scale Hypertextual Web Search Engine. In Proceedings of World Wide Web Conference. 107-117.

Csurka, G., Dance, C. R., Fan, L., Willamowski, J., & Bray, C. (2004). Visual Categorization with Bags of Keypoints. In Proceedings of ECCV International Workshop on Statistical Learning in Computer Vision. 1-22.

Erkan, G. & Radev, D. R. (2004). LexRank: Graph-based Lexical Centrality as Salience in Text Summarization. Journal of Artificial Intelligence Research. 457-479.

Fujii, K., Nanba, H., Takezawa, T., & Ishino, A. (2016). Enriching Travel Guidebooks with Travel Blog Entries and Archives of Answered Questions. In Proceedings of ENTER 2016.

Fukushima, T., Okumura, M., & Nanba, H. (2002). Text Summarization Challenge 2 / Text Summarization Evaluation at NTCIR Workshop 3. In Working Notes of the 3rd NTCIR Workshop Meeting, PART V. 1-7.

Hao, Q., Cai, R., Wang, C., Xiao, R., Yang, J.-M., Pang, Y., & Zhang, L. (2010). Equip Tourists with Knowledge Mined from Travelogues. In Proceedings of World Wide Web Conference.

Lin, C. Y. (2004). ROUGE: a Package for Automatic Evaluation of Summaries. In Proceedings of Workshop on Text Summarization Branches Out, 74-81. Wu, X., Li, J., and Neo, S.-Y. (2008). Personalized Multimedia Web Summarizer for Tourist. In Proceedings of World Wide Web Conference.

Lowe, D. G. (1999). Objective Recognition from Local Scale-Invariant Features. In Proceedings of the International Conference on Computer Vision. 1150-1157.

Luhn, H. P. (1958). The Automatic Creation of Literature Abstracts. IBM Journal of Research and Development, 2(2).

Nanba, H., Taguma, H., Ozaki, T., Kobayashi, D., Ishino, A., & Takezawa, T. (2009). Automatic Compilation of Travel Information from Automatically Identified Travel Blogs. In Proceedings of the Joint Conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing. 205-208.

Otterbacher, J., Erkan, G., & Radev, D. R. (2009). Biased LexRank: Passage Retrieval Using Random Walks with Question-based Priors. Information Processing & Management. 45(1): 42-54.

Radev, D. R., Jing, H., & Budzikowska, M. (2000). Centroid-based Summarization of Multiple Documents: Sentence Extraction, Utility-based Evaluation, and User Studies. In Proceedings of the NAACL-ANLP Workshop on Automatic summarization. 4: 21-30.