Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence

Generating Pictorial Storylines via Minimum-Weight Connected Dominating Set Approximation in Multi-View Graphs

Dingding Wang
Center for Computational Sciences
University of Miami
Miami, FL 33136

Tao Li
School of Computer Science
Florida International University
Miami, FL 33199

Mitsunori Ogihara
Center for Computational Sciences
University of Miami
Miami, FL 33136

Abstract

This paper introduces a novel framework for generating pictorial storylines for given topics from text and image data on the Internet. Unlike traditional text summarization and timeline generation systems, the proposed framework combines text and image analysis and delivers a storyline containing textual, pictorial, and structural information to provide a sketch of the topic evolution. A key idea in the framework is the use of an approximate solution for the dominating set problem. Given a collection of topic-related objects consisting of images and their text descriptions, a weighted multi-view graph is first constructed to capture the contextual and temporal relationships among these objects. Then the objects are selected by solving the minimum-weighted connected dominating set problem defined on this graph. Comprehensive experiments on real-world data sets demonstrate the effectiveness of the proposed framework.

Introduction

The rapid increase of online information often leads to information explosion. Those who search for information on the Internet often encounter the problem of navigating through an overwhelmingly large collection of web documents to extract meanings from the collection. To solve this problem various types of document understanding systems have been recently proposed. For example, generic and query-focused multi-document summarization systems aim to choose from the documents a subset of sentences that collectively conveys the principle idea or the query-related idea. News topic detection and tracking systems usually aim at grouping news articles into a cluster to present an event in a topic and monitor future events related to the topic. More recently timeline generation systems have been proposed, which create summaries in a manner to present the evolution of events in a topic by leveraging temporal information attached or appearing in the documents.

Although these document understanding systems can reduce the information overload problem, they still face two major limitations: (1) Most of the systems focus on highlighting and summarizing events in a topic and lack of the theme structure to capture the event evolution. Although timeline systems present the sequence of events based on the order of time, the linear-structured timelines usually lose comprehensive information of the evolutionary processes. (2) These systems usually deliver texts as summaries but to the reader texts sometimes may look boring and uninteresting.

In this paper, we propose a novel framework that resolves these two problems by generation of pictorial and temporal storylines, with the idea that temporal organization would give the summaries a structure for the reader to follow and presentation of image would make summaries more enjoyable to read and would improve the digestion of the contents by the reader. Input to our system are a topic and a collection of documents/objects relevant to the topic, where each document/object contains images and texts. Our system first constructs a multi-view object graph by text and image analysis and by incorporation of temporal information. Next the system selects a set of nodes using an approximation algorithm for the Minimum-Weight Dominating Set Problem and creates a storyline by the use of a directed Steiner tree algorithm. Our major contributions are as follows. (1) The proposed framework combines image and text processing to improve the semantic analysis and deliver vivid pictorial summaries to readers. (2) We formulate the problem as a graph-based optimization problem and solve the problem utilizing efficient approximation algorithms. (3) The generated storylines achieve both the temporal continuity and the content coherence, which provides richer information and a better result representation to readers.

Related Work

Multi-document summarization is a related topic which compresses a given collection of documents into a summary of much smaller size by extracting either the principle information or the information related to a query associated with the collection. A variety of multi-document summarization methods have been proposed in the literature. The most commonly used methods are centroid based (Radev et al. 2004; Lin and Hovy ; Yih et al.) or graph based (Mihalcea and Tarau ; Erkan and Radev). Other methods such as latent semantic analysis (LSA), non-negative matrix factorization (NMF) and sentence-based topic models have also been used to produce the summaries by selecting semantically and probabilistically important sentences in the documents (Gong and Liu ; Wang et al. 2008; **?**). Most of the existing methods however are geared toward forming short summaries by extracting sentences from the input and thus ignore the temporal or structural information possibly present in the input documents.

Another related topic is topic detection and tracking (TDT) which aims to group news articles based on the topics discussed in them, detect some novel and previously unreported events, and track future events related to the topics. Information retrieval techniques (e.g. information extraction, filtering, and document clustering), are often applied to these problems (All ; Brants, Chen, and Farahat 2003; Kumaran and Allan 2004; Makkonen, Ahonen-Myka, and Salmenkivi 2004; Yang, Pierce, and Carbonell 1998).

There also exist a limited number of studies on generating timelines (Shahaf and Guestrin 2010) and storylines. Google news Timeline clusters news articles into groups based on the topics and then lists them in the order of time. In (Alonso, Baeza-Yates, and Gertz 2009) a framework is proposed for generating temporal snippets as complementary to traditional text summaries. These timeline creation methods consider the temporal information as references and represent the results in a linear structure. Very recently an evolutionary timeline summarization approach has been proposed to construct the timeline of a topic by optimizing the relevance, coverage, coherence ,and diversity (Yan et al. 2011) Unlike these existing systems, our framework integrates text, image, and temporal information, and generates storyline-based summaries to reflect the evolution of the given topic.

Methodology

Problem Definition

The problem of generating pictorial temporal storyline can be defined as follows:

Input: A query q and a collection of m objects, $O = \{o_1, o_2, \ldots, o_m\}$, where each object o_i is an image with a text description (e.g., a small paragraph or a sentence) and with a timestamp t_i.

Output: A pictorial storyline which consists of the most representative objects summarizing the evolution of a query-relevant topic.

Below we will formulate this problem a the Minimum-Weight Connected Dominating Set Problem on a multi-view graph, which can be decomposed into two optimization problems: finding a minimum-weight dominating set and connecting the dominators using directed Steiner tree.

System Framework

Figure 1 shows the storyline generation framework. Given a collection of images and their text descriptions, we first construct a weighted multi-view object graph where each vertex is an image associated with short texts describing the image. The graph has two sets of edges, undirected edges representing a certain level of similarity between the objects and directed edges representing a certain type of pairwise temporal relationship. Each vertex is assigned a weight, which is the calculated cdistance between the object and the query. We will run a minimum-weight dominating set algorithm on the graph to find the objects that are collectively the most representative of O and then run a directed Steiner tree approximation algorithm to form a storyline of the chosen objects.

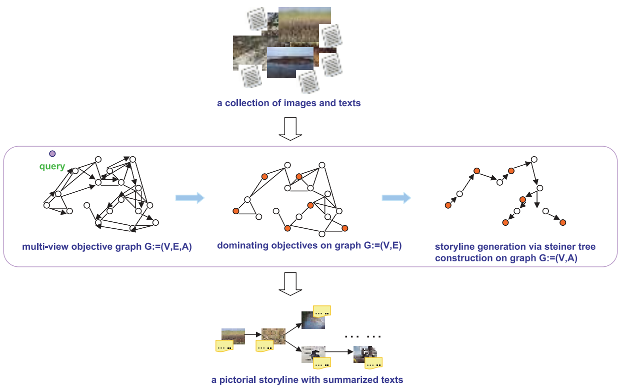

a collection of images and texts
query
multi-view objective graph G:=(V,E,A)
dominating objectives on graph G:=(V,E)
storyline generation via steiner tree construction on graph G:=(V,A)
... ..
... ...
a pictorial storyline with summarized texts

Figure 1: The framework of the storyline generation.

Multi-View Object Graph Construction

Definition. A *multi-view graph* is a triple $G = (V, E, A)$, where V is a set of vertices (nodes), E a set of undirected edges, and A a set of directed edges (arcs).

Given a collection of images and their text descriptions with time stamp, we construct a multi-view object graph by viewing the images as the vertices V, calculating the undirected edges E based on both text and image similarities, and calculating the directed edges A based on difference between time stamps. We use four nonegative real parameters $\alpha, \beta, \tau_1, \tau_2, \tau_1 < \tau_2$ in defining these edges.

For texts, we apply the standard "bag-of-words" representation. For images, we calculate their features from color and texture by adopting Color and Edge Directivity Descriptor (CEDD) (Chatzichristofis and Boutalis 2008). For both feature vectors, we use cosine measure to calculate similarity.

Let o_i and o_j be two objects in V. To define E, we join the two by an edge if and only if the text similarity and the image similairty between the two are greater than respectively α and β. To define A, we draw an arc from o_i and o_j if and only if $\tau_1 \leq t_j - t_i \leq \tau_2$, where t_i and t_j are their respective time stamps. We call $[\tau_1, \tau_2]$ the *temporal window.* Also, for each node o_i, its vertex weight, $w(o_i)$, is $1-$ (the cosine similarity between q and o_i.

Identifying Query-Relevant Dominating Objects via Minimum-Weight Dominating Set

We say that a vertex u of a graph dominates another vertex v of the graph, if u and v are joined by an edge in the graph. A subset S of the vertex set of an undirected graph is a *dominating set* if for each vertex u either u is in S or a vertex in S dominates u. The problem of finding a set of query-relevant objects can be viewed as the minimum-weight dominating set problem on the undirected graph (V, E).

Problem 1. The Minimum-Weight Dominating Set Problem (MWDS) is the problem of finding, given a vertex-weighted undirected graph G, from all dominating sets of G the one whose total vertex weight is the smallest.

MWDS is known to be NP-hard (Raz and Safra 1997). We consider the following straightforward greedy algorithm for obtaining an approximate solution (Algorithm 1). This algorithm views that the weight of a newly added vertex is evenly shared among its newly covered neighbors and selects the node that minimizes this share at each round of iteration.

Algorithm 1 Greedy MWDS Approximation

```
Input:   $G = (V, E)$: vertex weighted object graph
         $W$: maximal number of dominators
Output: dominating set $S$
$S = \emptyset$
$T = \emptyset$
while $|S| < W$ and $S! = V$ do
    for $v \in V - S$ do
    $s(v) = \|\{v'|(v', v) \in E\} \setminus T\|$
    $v* = \text{argmin}_v \frac{w(v)}{s(v)}$
    $S = S \cup \{v*\}$
    $T = T \cup \{v''|(v'', v*) \in E\}$
```

The approximation rate of this algorithm is $1 + \log(\Delta\|OPT\|)$, where Δ is the maximal degree of G and OPT is the optimal dominating set (Shen and Li 2010).

Generating Storylines by Connecting Dominating Objects via Directed Steiner Tree

Once we select the most representative objects using the dominating set approximation, we need to generate a natural storyline capturing the temporal and structural information of the query-relevant events. To study this problem we use the concept of Steiner trees. A Steiner tree of a graph G with respect to a vertex subset X is the edge-induced substree of G that contains all the vertices of X having the minimum total cost, where the cost is often the size of the tree, which is the number of vertices of the tree minus 1. Here we use as the cost the total weight of the vertices and define the Steiner problem as follows.

Problem 2. Given a directed graph $G = (V, A)$, a set X of vertices (called terminals), and a root $v_0 \in X$ from which every vertex of X is reachable in G, find the subtree G rooted at v_0 containing X with the smallest total vertex weight.

The problem is known to be NP-hard since the undirected version is already NP-hard. While the undirected Steiner tree problem has been well studied, much less work has been on the directed graph version (Charikar and Chekuri 1999). A straightforward solution for this problem is to find the shortest path from the root to each of the terminal and merge the paths. Of course, combining lightest paths does not guarantee the minimum total cost.

This observation suggests Algorithm 2 from (Charikar and Chekuri 1999). The algorithm takes a level parameter $i \geq 1$ and takes as input the target terminal set Y, the root r, and the required number of nodes in Y to cover, ℓ. In the case where $i = 1$ the algorithm defaults to the straightforward algorithm; i.e., it selects ℓ vertices in Y closest to r and returns the union of the shortest paths to the ℓ vertices. The length of an arc $(u, v) \in A$ is the vertex weight of u. We will make the initial call of $A_i(k, v_0, X)$ with X set to the dominating set calculated by the previous algorith, v_0 set to the vertex among X with the earliest time stamp, and k set to $\|X\|$. We will interpret the output tree as the storyline transitioning from the root object to all the other dominating objects. For a constant i, the algorithm is known to run in polynomial time and produce an $O(k^{1/i})$ approximation (Charikar and Chekuri 1999).

Algorithm 2 $A_i(G, k, r, X)$

```
Input:    G=(V,A): vertex-weighted directed graph
          X: target vertex set $X$
          $r \in X$: the root
          $k \geq 1$: the target size
Output:   T: a Steiner tree rooted at $r$
          covering at least $k$ vertices in $X$
$T = \emptyset$
while $k > 0$
    $T_{best} \leftarrow \emptyset$
    cost$(T_{best}) \leftarrow \infty$
    for each vertex $v$, $(v_0, v) \in A$, and $k'$, $1 \leq k' \leq k$
    $T' \leftarrow A_{i-1}(k', v, X) \bigcup \{(v_0, v)\}$
    if cost$(T_{best}) >$ cost$(T')$ then $T_{best} \leftarrow T'$
    $T \leftarrow T \cup T_{best}$
    $k \leftarrow k - \|X \cap V(T_{best})\|$
    $X \leftarrow X \setminus V(T_{best})$
return $T$
```

Experiments and Evaluation

Evaluating the generated storylines is not an easy job because the task itself is very subjective. Thus, in the experiments, we first evaluate the summarization ability of the overall texts contained in the storylines by ignoring the temporal and structural information. Then we evaluate the subtasks of our method and compare our approaches with other alternatives. And finally we conduct a user study to evaluate the overall user satisfaction.

Data Sets and Annotation

Data Sets The data sets are manually collected 355 images with text descriptions from Flickr, ABC News, Reuters, AOL News, and National Geographic to build the standard data sets. And there are four topics contained in the data as shown in Table 1. Figure 2 shows one example object for each of the four topics, and Table 2 provides the text description and the time stamp associated with each example image. We call an image and its description an object.

Topics	# of objects
Flooding in hurricane Katrina	101
Building Collapse in hurricane Katrina	101
Damage to sea grass in oil spill	53
Damage to animals in oil spill	100

Table 1: Description of the data sets

Data Annotation Given four target queries related to the four topics in the data sets, we hire 8 human labelers to manually pick representative objects to construct the storylines from the data. Each of the annotators is assigned two

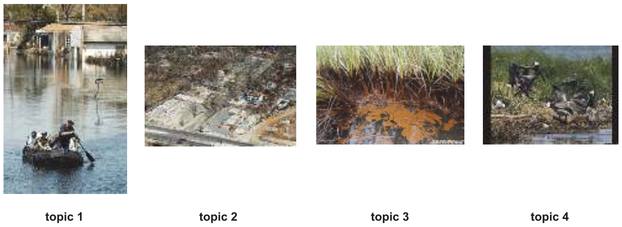


Figure 2: Sample images from each topic.

topic	text	time
1	New Orleans- U.S. Navy Sailors assigned to the dock landing ship USS Tortuga search flooded New Orleans neighborhoods for survivors of Hurricane Katrina.	09/06/2005
2	Gulf Coast of Mississippi-Mississippi suffered extensive damage from the storm surge and high winds of hurricane Katrina.	09/30/2005
3	Heavy crude oil from the Deepwater Horizon wellhead has penetrated the wetland grasses on an island in Bay Batiste, Louisiana. The grass will die as a result and erosion of the low lying island will be accelerated.	06/16/2010
4	Brown pelicans and seagulls are seen at a rookery near an absorbent boom soaked with oil from the Deepwater Horizon spill.	06/12/2010

Table 2: Text description and time of sample images.

queries, thus each query has four human generated storylines from different annotators. The four queries are: (1). What are the effects of Hurricane Katrina in New Orleans? (2). How the buildings were destroyed and rebuilt during Hurricane Katrina? (3). What are the damages to the wetlands and sea grasses in BP oil spill? (4). How BP oil spill affects animals such as pelicans and turtles?

Evaluation Metrics

Summarization Performance We compare the human created storyline summaries with the summaries generated by different systems. The ROUGE (Lin and E.Hovy) toolkit (version 1.5.5) is used in the experiments to measure each summarization system, which is widely applied for document summarization performance evaluation. We also examine the precision and recall of the representative objects selected by different methods with the human picked ones. In experiments, we usually require the number of objects selected by various systems comparable with the number of representative objects picked by human labelers, thus the precision and recall scores are consistent in general. Therefore, we only report the precision scores in the experimental results.

An Illustrative Example

Figure 3 illustrates an example storyline generated by our proposed method where the query is "How BP oil spill affects animals such as pelicans and turtles?". With the pictorial temporal storyline, users can easily capture the event evolution.

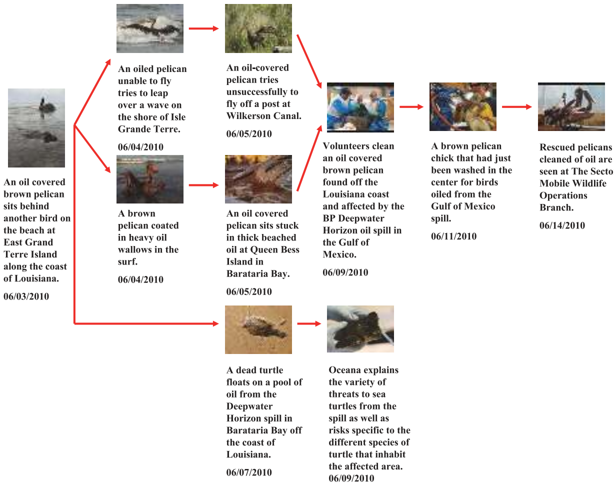

An oil covered brown pelican sits behind another bird on the beach at East Grand Terre Island along the coast of Louisiana.
06/03/2010
An oiled pelican unable to fly tries to leap over a wave on the shore of Isle Grande Terre.
06/04/2010
A brown pelican coated in heavy oil wallows in the surf.
06/04/2010
An oil-covered pelican tries unsuccessfully to fly off a post at Wilkerson Canal.
06/05/2010
An oil covered pelican sits stuck in thick beached oil at Queen Bess Island in Barataria Bay.
06/05/2010
A dead turtle floats on a pool of oil from the Deepwater Horizon spill in Barataria Bay off the coast of Louisiana.
06/07/2010
Volunteers clean an oil covered brown pelican found off the Louisiana coast and affected by the BP Deepwater Horizon oil spill in the Gulf of Mexico.
06/09/2010
Oceana explains the variety of threats to sea turtles from the spill as well as risks specific to the different species of turtle that inhabit the affected area.
06/09/2010
A brown pelican chick that had just been washed in the center for birds oiled from the Gulf of Mexico spill.
06/11/2010
Rescued pelicans cleaned of oil are seen at The Sector Mobile Wildlife Operations Branch.
06/14/2010

Figure 3: An illustrative Example.

Overall Summarization Evaluation

There exist few previous studies on storyline generation, however, there are quite a number of document summarization approaches focusing on generating a short summary given a collection of texts. By ignoring the temporal and structural information in the generated storyline, the texts describing the dominating objects naturally form a summary. Thus, in this set of experiments, we compare the summarization performance of our method with existing query-relevant document summarization methods.

We implement several most widely used query-based document summarization systems as the baselines. They are (1) LexPageRank, which first constructs a sentence connectivity graph based on cosine similarity and then selects important sentences based on the concept of eigenvector centrality (Erkan and Radev); (2) NMF, which performs NMF on terms by sentences matrix and ranks the sentences by their weighted scores (Lee and Seung); (3)TMR which incorporates the query information into the topic model, and uses topic based score and term frequency to estimate the importance of the sentences (Tang, Yao, and Chen 2009); (4)Seman, which calculates sentence-sentence similarities by semantic role analysis, clusters the sentences via symmetric non-negative matrix factorization, and extracts the sentences based on the clustering result (Wang et al. 2008); (5)MultiMR, which uses a manifold-ranking algorithm by considering the within-document sentence relationships (Wan and Xiao 2009).

Given the queries and the collection of objects, each of these systems generates a query-based summary consisting of certain number of text pieces contained in the objects. The number of selected text pieces is comparable to the number of objects in the human created storylines.

Experimental Results Figure 4 shows the overall summarization performance, and Figure 5 demonstrates the average precision of different methods. In Figure 4, we also

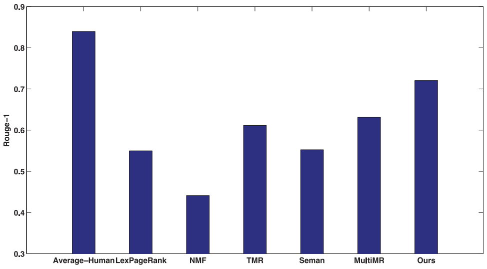


Figure 4: Rouge-1 Results of Different Systems.

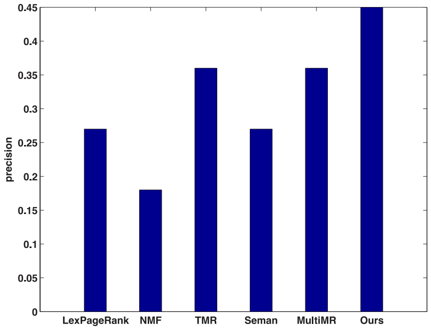


Figure 5: Average Precision of the Results of Different Systems.

show the average scores of results generated by human labelers to demonstrate the inner agreement of human created storylines. From the results, we clearly observe that the objects contained in our generated storylines have good summarization performance, and the results outperform recent query-based summarization methods. The effective results of our method mainly benefit from the following facts of the proposed algorithms. (1) The proposed method incorporates both image and text analysis to enrich the semantic similarity representation. (2) The selected representative objects can effectively capture the main concepts contained in the topic, and these dominators are representative and diverse. (3) Our method considers the temporal information in the objects to achieve the smoothness and coherence.

Comparison on Different Graph Generation Approaches

In this set of experiments, we compare our graph construction approach which considers both text and image similarities with traditional image or document analysis methods which base on either image graphs or text graphs. In our methods, we use two parameters (α and β) as the thresholds for generating edges in the graph. Thus, we gradually adjust the values of α and β to empirically evaluate the sensitivity of the proposed method with respect to the threshold setting. We use the same algorithms in the dominating object selection and the storyline generation processes. Figure 7 shows the Rouge-1 scores when α and β vary respectively, and Figure 6 demonstrates the precision results of the methods only considering text similarity or image similarity.

From the results we observe (1) the methods performing on text only graph outperform the methods on image only graph, however, both of these methods can not achieve satisfactory results in general; (2) texts and images are complimentary to each other so that combing the similarity analysis on them achieves better performance than using one of them solely; (3) when the thresholds are very small, edges fail to represent the similarity of the objects, while if they are set too large, the graph will be very sparse.

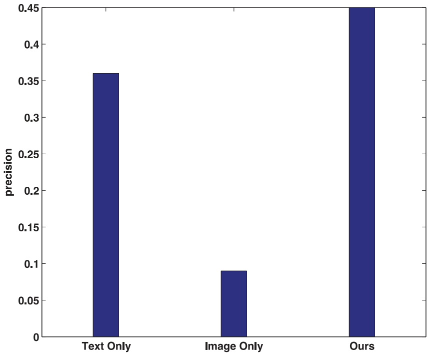
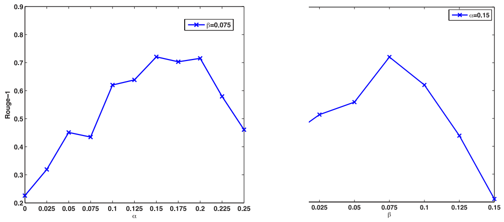


Figure 6: Precision Results of Text Only and Image only Methods.



(a) Text Similarity Threshold (α)

(b) Image Similarity Threshold (β)

Figure 7: Rouge-1 Results of Different Similarity Thresholds.

Comparison on Different Representative Object Selection Approaches

In this set of experiments, we evaluate the minimum-weight dominating set approximation used in our dominating objects discovery process. In order to identify the dominating objects, the typical alternative approaches could first filter the objects using the given queries and then conduct clustering algorithms on the query-relevant objects. The centroid objects in each cluster could be treated as the dominating objects. We implement various clustering algorithms such as traditional K-means, Spectral Clustering with Normalized Cuts (Ncut) (Shi and Malik 2000), and Nonnegative Matrix Factorization (NMF) (Lee and Seung) to compare with the dominating set algorithm (MWDS) used in our method. It is not feasible to simply combine the text and image similarities since they are not in the same scale. Thus in this set of experiments we only perform these clustering algorithms on the text graph. In order to make a fair comparison, we also report the results of our method on the text only graph.

Experimental Results Figure 8 and Figure 9 demonstrate the Rouge-1 and precision results of the alternative approaches respectively. We clearly observe that the MWDS algorithm outperforms all the other clustering algorithms.

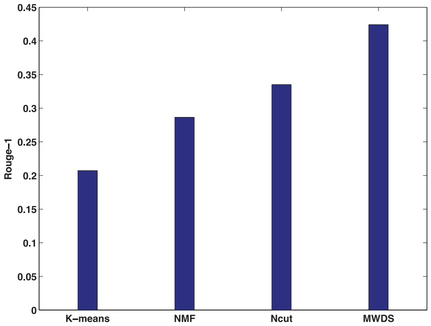


Figure 8: Rouge-1 Scores of Various Clustering Algorithms on Text Only Graph.

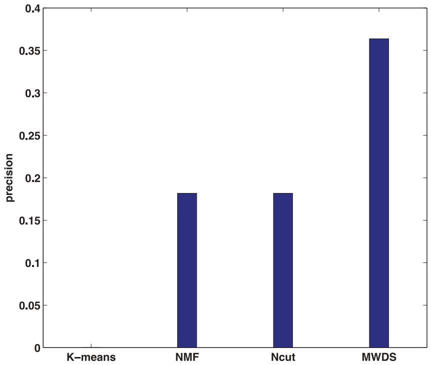


Figure 9: Precision Scores of Various Clustering Algorithms on Text Only Graph.

The good results of MWDS benefit from the follow factors. (1) Similarity between each object and the given query is assigned as the node weight in MWDS, which provides a more comprehensive way of incorporating query information. (2) Objects selected by MWDS are more representative because each object is either within the minimum dominating set or connected in the set. (3) The redundancy is minimized since the dominating set is of minimum size.

Comparison on Different Approaches for Connecting Dominating Objects

In this set of experiments, we evaluate our directed Steiner tree (DST) approximation for connecting the dominating objects to form the storylines. Instead of performing DST on the temporal graph ($G = (V, A)$), an alternative approach is to apply an algorithm of undirected Steiner tree (ST) on the similarity graph ($G = (V, E)$). Here, we conduct experiments to compare the two approaches.

Figure 10 shows the Rouge-1 scores of the alternative ST approach and our DST approach. From the results, we observe that the DST approach considering the temporal information outperforms the approach only taking into account the content similarity. This observation confirms us that our approach can achieve the continuity and smoothness of the story evolution, which is consistent with human perspectives.

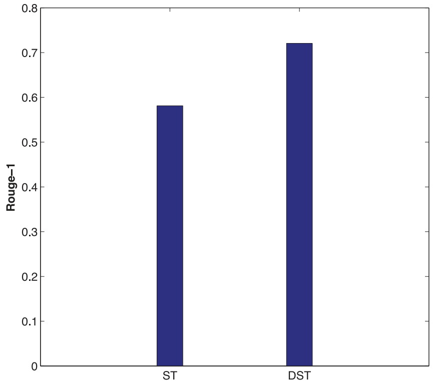
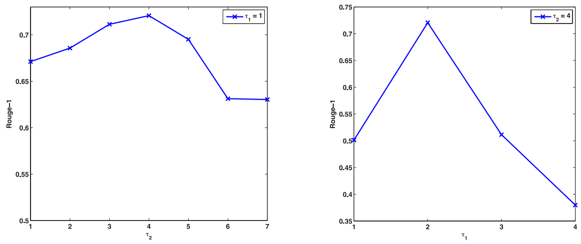


Figure 10: Rouge-1 Results of Different Approaches for Connecting Dominating Objects



(a) $\tau_1 = 1$ (b) $\tau_2 = 4$

Figure 11: Rouge-1 Results of Different Time Intervals.

Parameter Tuning

Recall that we introduce a directed edge from object i (o_i) to j (o_j) if $t_i + \tau_1 \leq t_j \leq t_i + \tau_2$. In this experiment, we examine the effects of the window size setting in our method. Intuitively, if the window size is too large, the arcs fail to represent the time continuity, while if the window size is too small, the graph is too sparse and too many intermediate nodes may be involved to form the storyline. Although different stories have different evolution paces, which may affect the window size setting, for most of the popular news events, typical time intervals may vary from one day to one week. Figure 11 demonstrates the Rouge-1 results when the time interval parameters vary.

From the figure, we can observe that when τ_1 is one (day) and τ_2 varies from one to seven days, the results are relatively stable and achieve the best performance when τ_2 is 4. However, when τ_1 is set to zero, the performance drops dramatically because the system may select many objects in the same day and make the redundancy of the results very high. And also when τ_1 is too large, the performance is poor because the results may lose time continuity.

A User Study

Since storyline generation is a subjective process, to better evaluate the summaries and structures of the created storylines, we conduct a user survey. The subjects of the survey are 15 students at different levels and from various majors of a university. In this survey, we use the same queries and data sets as described in Section , and ask each participant to read two topics and compare the results from different systems. A score of 1 to 5 needs to be assigned to each system according

DocSum	TL1	TL2	SL1	SL2	Ours
2.2	2.3	3.3	2.5	3.2	4.0

Table 3: Survey: User ratings on different systems based on their satisfaction.

to the user's satisfaction of the results. A rank of 5 (1) indicates that results of the system is most (least) helpful. We implement the following systems for comparison. (1) DocSum: one of the most widely used traditional document summarization methods (the LexPageRank method). (2) TL1: ordering the objects in DocSum based on their time stamps. (3) TL2: organizing the objects in our generated storyline in a timeline structure based on the time stamps associated with the objects. (4) SL1: performing NMF as the clustering algorithm in the dominating object selection procedure, and the rest approaches are the same with our proposed storyline generation method. (5) SL2: conducting the ST algorithm to connect the dominating objects to form Steiner trees. (6) Ours: our method as proposed in Section .

Table 3 shows the average ratings that the participants assign to each method. From the results, we have the following observations: (1) Users prefer a pictorial and structural summary to a traditional one consisting of plain texts. (2) Timeline2 and Ours are generated using the same algorithms, however, Timeline2 provides the simple linear structure and Ours uses the graph structure to reflect the story development. Thus, from the survey we confirm that the storyline structure provides more information and is more helpful than the timeline structure. (3) The approximation algorithms used in our method are more effective than those alternatives, thus our method meets the users needs better and achieve high satisfaction from the users.

Conclusion

In this paper, we propose a novel framework to summarize the evolution of topics in a pictorial and structural way. We utilize both text and image analysis and formalize the problem into a graph-based optimization problem and solve the problem using approximation algorithms of minimum-weight dominating set and directed Steiner tree. Comprehensive experiments and a user survey are conducted on real-world web data sets.

Acknowledgments

This work is supported in part by NIH/NIA grant 1P30DA027828-01A1, DHS grant 2009-ST-062-000016, NSF grant DBI-0850203, and Army Research Office grant W911NF-10-1-0366.

References

Alonso, O.; Baeza-Yates, R.; and Gertz, M. 2009. Effectiveness of Temporal Snippets. In *Proceedings of WWW 2009*. ACM.

Brants, T.; Chen, F.; and Farahat, A. 2003. A system for new event detection. In *Proceedings of SIGIR'03*, 330–337. New York, NY, USA: ACM.

Charikar, M., and Chekuri, C. 1999. Approximation algorithms for directed steiner problems. *J. Algorithms* 33:73–91.

Chatzichristofis, S. A., and Boutalis, Y. S. 2008. Cedd: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval. In *Proceedings of ICVS'08*, 312–322. Berlin, Heidelberg: Springer-Verlag.

Erkan, G., and Radev, D. Lexpagerank: Prestige in multi-document text summarization. In *Proceedings of EMNLP 2004*.

Gong, Y., and Liu, X. Generic text summarization using relevance measure and latent semantic analysis. In *Proceedings of SIGIR 2001*.

Kumaran, G., and Allan, J. 2004. Text classification and named entities for new event detection. In *Proceedings of SIGIR'04*, 297–304. New York, NY, USA: ACM.

Lee, D. D., and Seung, H. S. Algorithms for non-negative matrix factorization. In *NIPS 2001*.

Lin, C.-Y., and E.Hovy. Automatic evaluation of summaries using n-gram co-occurrence statistics. In *Proceedings of NLT-NAACL 2003*.

Lin, C.-Y., and Hovy, E. From single to multi-document summarization: A prototype system and its evaluation. In *Proceedings of ACL 2002*.

Makkonen, J.; Ahonen-Myka, H.; and Salmenkivi, M. 2004. Simple semantics in topic detection and tracking. *Inf. Retr.* 7:347–368.

Mihalcea, R., and Tarau, P. A language independent algorithm for single and multiple document summarization. In *Proceedings of IJCNLP 2005*.

Radev, D.; Jing, H.; Stys, M.; and Tam, D. 2004. Centroid-based summarization of multiple documents. *Information Processing and Management* 919–938.

Raz, R., and Safra, S. 1997. A sub-constant error-probability low-degree test, and a sub-constant error-probability pcp characterization of np. In *Proceedings of STOC '97*, 475–484. New York, NY, USA: ACM.

Shahaf, D., and Guestrin, C. 2010. Connecting the dots between news articles. In *KDD*.

Shen, C., and Li, T. 2010. Multi-document summarization via the minimum dominating set. In *Proceedings of COLING '10*, 984–992. Stroudsburg, PA, USA: Association for Computational Linguistics.

Shi, J., and Malik, J. 2000. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 22:888–905.

Tang, J.; Yao, L.; and Chen, D. 2009. Multi-topic based query-oriented summarization. In *SDM*, 1147–1158.

Wan, X., and Xiao, J. 2009. Graph-based multi-modality learning for topic-focused multi-document summarization. In *IJCAI*, 1586–1591.

Wang, D.; Li, T.; Zhu, S.; and Ding, C. 2008. Multi-document summarization via sentence-level semantic analysis and symmetric matrix factorization. In *In SIGIR'08*, 307–314. ACM.

Yan, R.; Wan, X.; Otterbacher, J.; Kong, L.; Li, X.; and Zhang, Y. 2011. Evolutionary timeline summarization: a balanced optimization framework via iterative substitution. In *SIGIR*, 745–754.

Yang, Y.; Pierce, T.; and Carbonell, J. 1998. A study of retrospective and on-line event detection. In *Proceedings of SIGIR '98*, 28–36. New York, NY, USA: ACM.

Yih, W.-T.; Goodman, J.; Vanderwende, L.; and Suzuki, H. Multi-document summarization by maximizing informative content-words. In *Proceedings of IJCAI 2007*.