A New Two-layer Storyline Generation Framework for Disaster Management

Wubai Zhou, Florida International University, Email: wzhou005@cs.fiu.edu
and
Chao Shen, Florida International University, Email: cshen001@cs.fiu.edu
and
Tao Li, Florida International University, Email: taoli@cs.fiu.edu
and
Shu-Ching Chen, Florida International University, Email: chens@cs.fiu.edu
and
Ning Xie, Florida International University, Email: nxie@cs.fiu.edu
and
S. S. Iyengar Florida International University, Email: iyengar@cs.fiu.edu

Disasters, such as hurricanes, earthquakes and environmental emergencies, are serious disruptions of the functioning of a community or a society. To mitigate the social and physical impact of disasters, a critical task in disaster management is to extract situation updates on the disaster from a large number of disaster-related documents, and obtain a big picture of the disaster's trends and how it affects different areas. In this paper, we present a novel two-layer storyline generation framework which generates an overall storyline of the disaster events in the first layer, and provides condensed information about specific regions affected by the disaster (i.e., a location-specific storyline) in the second layer. To generate the overall storyline of a disaster, we consider both temporal and spatial factors, which are encoded using integer linear programming. While for location-specific storylines, we employ a Steiner tree based method. Compared with the previous work of storyline generation, which generates flat storylines without considering spatial information, our framework is more suitable for large-scale disaster events. We further demonstrate the efficacy of our proposed framework through the evaluation on the datasets of three major hurricane disasters.

Keywords: Document summarization, disaster management, storyline generation.

1. INTRODUCTION

Natural disasters such as hurricanes, earthquakes and tsunamis cause inestimable physical destruction, loss of life and property around the world every year. For example, Hurricane Sandy affected east coast of U.S. in 2012 and posed immense threat to businesses, human lives and properties. In order to minimize the consequent loss of the disasters, a critical task in disaster management is to efficiently analyze and understand the disaster-related situation updates. This requires effective information gathering methods to operate on a myriad of web documents, e.g., news and reports that are related to the disasters. The domain experts expect to obtain condensed information about the detailed disaster event description, e.g., the evolutionary tendency of the disaster with respect to different locations (Li and Li, 2014). However, it is a non-trivial task to generate a big picture of the disaster events due to the flood of web documents.

To tackle this problem, various types of document understanding systems have been proposed in the last decade. These systems include (1) summarization-based systems (Li, Li, and Li, 2012; Radev, Jing, Styś, and Tam, 2004; Saggion, Bontcheva, and Cunningham, 2003; Shen and Li, 2010; Wei, Li, Lu, and He, 2008) that choose from multiple documents a subset of sentences conveying the principle idea; (2) topic detection and tracking systems (Allan, 2012) aiming to group documents into different clusters as events and monitor future events related to the corresponding topic; and (3) timeline generation systems (Shahaf, Guestrin, and Horvitz, 2012; Wang, Li, Ogihara, et al., 2012) that create summaries to present the evolution of an event by leveraging temporal information attached to or extracted from the documents. These systems are able to alleviate the so-called *information overload* problem to some extent; however, they suffer from several limitations that may affect the quality of the summarized results. Firstly, most of them focus on summarizing an event via topic evolution over the time, but ignore the spatial information which is important especially for large-scale disaster events. For instance, for a hurricane which affects several states of U.S., we are interested in how these regions are affected, and how the hurricane evolves over different geo-spatial regions. Secondly, these systems usually generate a single layer summarization or storyline to reflect topic switches over the entire event. However, due to the spatial factor, the information evolution over a disaster event is intrinsically hierarchical. In most cases, the domain experts are interested in not only the general picture of the disaster, but also how it affects a particular region.

In this paper, we propose a storyline generation framework that addresses the aforementioned limitations by generating a two-layer storyline that consists of global storylines for cross-location disaster events on the first layer and location-specific storylines for individual events on the second layer. Specifically, in our framework, a disaster event is initially summarized from a large set of documents (e.g., news and reports) with a big picture showing how the disaster affects different regions. It can then be zoomed into a specific location for more detailed location-specific event summarization. In the cross-location layer, integer linear programming is employed to summarize the event via a list of representative locations, each of which is associated with a short description. On the location-specific layer, a Steiner-tree based approach is applied to generate a storyline for each specific location. A demo of our system can be found at `http://bigdata-node01.cs.fiu.edu/HurricaneStoryline/storyline.html`.

In summary, the contribution of this work is three-fold:

—We present a novel two-layer summarization framework to summarize multiple disaster-related documents. The first layer provides an overall summary of the disaster events, while the second layer gives condensed information on how specific locations/regions were affected by the disaster.

—We consider both temporal and spatial factors when generating summaries for the disaster events, and these two factors enable us to reason on the evolution of events over time and locations. The generated summaries can be naturally represented as a storyline.

—We conduct quantitative experiments and case studies on crawled web documents related to three major hurricane disasters, and the results demonstrate the efficacy of our proposed framework in generating readable and understandable summaries.

The rest of the paper is organized as follows. After discussing related work in Section 2, we first define our problem in Section 3. In Section 4, an overview of our proposed framework is introduced. Detailed description of how to generate a global storyline and a local storyline is presented in Section 5 and Section 6 respectively. We evaluate our system in Section 7 and finally conclude our work and discuss potential extensions of the proposed framework in Section 8.

2. RELATED WORK

In this section, we highlight some previous research results that are most relevant to this work in the following three directions: multi-document summarization, topic detection and tracking, and storyline generation. We will also discuss several useful disaster situation-specific tools.

Multi-document summarization is a mechanism which addresses the information overload problem by compressing a given collection of documents into a concise summary. In general, it can be categorized into extractive and abstractive summarization (Mani and Maybury, 2001). Extractive summarization (Radev, Hovy, and McKeown, 2002) selects important sentences from the original documents to form a summary, while abstractive summarization (Radev et al., 2002) paraphrases the corpus using new sentences. The latter usually employs natural language generation techniques such as information fusion, sentence compression and reformulation. Our work is more related to extractive summarization. Various multi-document summarization methods have been proposed in the last decade, including centroid-based (Radev, Jing, and Budzikowska, 2000), graph-based (Erkan and Radev, 2004; Shen and Li, 2010), knowledge-based (Li and Li, 2014; Li, Wang, Shen, and Li, 2010), etc. Other methods, such as non-negative matrix factorization, latent semantic analysis and sentence-based topic models, have also been applied to generate the summaries by important sentences in the documents selecting semantically and probabilistically (Wang, Li, Zhu, and Ding, 2008; Shen, Li, and Ding, 2011). Most existing methods were proposed to generate short summaries by selecting sentence from the input; however, they often ignore the implicit temporal, spatial and structural information possibly presented in the documents.

Topic detection and tracking (TDT) is a research program initiated by DARPA (Defense Advanced Research Projects Agency) for finding and following the new events in streams that broadcast news stories[1]. It consists of three major technical tasks: tracking known events, detecting unknown events, and segmenting a news source into stories. Many promising approaches have been proposed and identified during the TDT evaluation, in particular within the information retrieval and natural language processing communities (Allan, 2012; Lavrenko, Allan, DeGuzman, LaFlamme, Pollard, and Thomas, 2002; Makkonen, Ahonen-Myka, and Salmenkivi, 2004). However, previous research only focused on detecting the flat structure of events, and fails to consider the hidden hierarchies of topics.

Storyline generation aims to obtain a sequence of summaries that describe how an event evolves over time, and has attracted great attention recently. For example, Google News Timeline clusters incoming articles into groups based on topics and lists the generated groups in chronological order. Alonso et al. (Alonso, Baeza-Yates, and Gertz, 2009) proposed a framework for generating temporal snippets to improve user search experience. These methods consider the temporal information as references and represent the results in chronological order. Recently, Wang et al. (Wang et al., 2012) proposed a framework that integrates text, image, and temporal information to generate storyline-based summaries in order to reflect the evolution of the given topic. Shahaf et al. (Shahaf et al., 2012) proposed a methodology called *metro map* for creating structural summaries of documents which optimizes relevance, coherence, coverage and connectivity simultaneously. Jiang et al. (Jiang, Perng, and Li, 2011) proposed an temporal event summarization solution to summarizes the temporal dynamics of the event sequences using the inter-arrival information. Unlike these existing systems, our framework takes into account the spatial information and generates storyline-based summaries to reflect the evolution of a given topic over different geo-spatial regions.

Disaster Situation-specific Tools: Commercial systems such as Web EOC and E-Team are usually used by Emergency Management departments located in urban areas (WebEOC, 2002; E-Teams, 2004). Recently Ushahidi provides a platform to crowd source news stories and crisis information using multiple channels and prepares visualization and interactive maps (Ushahidi, 2012) and GeoVISTA monitors tweets to form situation alerts on a map-based user interface according to the geo-locations associated with the tweets (GeoVISTA, 2010). These situation-specific tools provide query interfaces, GIS and visualization capabilities to support user interaction and query (Zheng, Shen, Tang, Zeng, Li, Luis, and Chen, 2013). However, they do not generate textual storylines to improve the situation awareness.

3. PROBLEM DEFINITION

To summarize what is happening in the vicinity of a given disaster, we present a storyline of the disaster in the form of a two-layer graph of events.

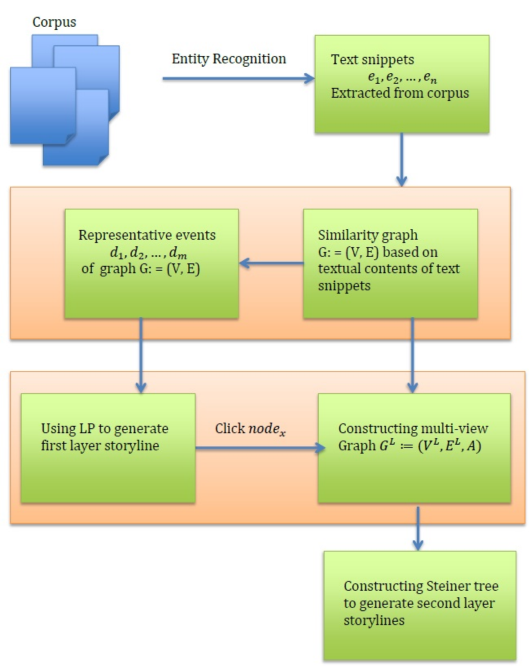
[1] http://projects.ldc.upenn.edu/TDT/

Definition 1. *An* event *is represented by a tuple* (t, l, s) *where* t *is the time that the event occurs,* l *is the location and* s *is the textual description about the event. For example, (08/27/2011, New York City, "The five main New York City-area airports will be closed to arriving flights") represents an event in Hurricane Sandy.*

The problem of generating a storyline can be defined as follows:
Input: A collection of documents related to a disaster.
Output: A two-layer storyline consists of the most representative events summarizing the evolution of a disaster-relevant topics. The **first layer** (or the upper layer) is a chain of events $(o_1, \ldots, o_n)$, as the global temporal and spatial evolution of a disaster, therefore also referred as the global storyline. An event of the upper layer o_i can be further expanded in the **second layer** (or the lower layer) to a connected tree of events as the temporal and topic evolution locally for a specific location of o_i.

A global storyline, which is a chain of events, describes how the disaster moves over time by the location attribute of the events and how the disaster affects different areas by the description attributes. The chain structure is used under the assumption that a disaster at any time should have only one geo-spatial center, which should move continuously over time. Such an assumption is valid for most of the natural disasters like hurricanes, storms, and blizzards, but not for the man-made disasters like cyber attacks, which will be explored in our future work. For local storyline generation, we follow previous work of storyline generation (Wang et al., 2012) to use a tree structure as the storyline to capture more topics in the topic evolution, allowing multiple topics to coexist at the same time.

4. SYSTEM FRAMEWORK



. High-level system overview

Figure 1 shows our system framework. Given a collection of documents related to a disaster, we first extract text snippets as sentences with time and location phrases, which are identified by Stanford NER (Finkel, Grenager, and Manning, 2005). Time phrases are normalized by

SUTime (Chang and Manning, 2012) to timestamps and location phrases are mapped to geocodes by Google API[2]. Together with its timestamp and geocode, a snippet approximately describes an event.

In our framework, the extracted text snippets are first organized as a similarity graph, followed by two layers of processing, corresponding to the two layers of the output. In the first layer, a minimum dominating set algorithm is employed on the snippet graph to find several representative events, on top of which an integer linear programming method is then proposed to find a chain of events reflecting the overall spatial evolution of the disaster as the global storyline. We visualize the global storyline on a map using Google map APIs.

If a user is interested in certain area and click it on the map, the map will be zoomed-in the clicked area and display the local storyline of the area. To do this, a sub-graph of the overall similarity graph is first induced and augmented to a multi-view graph. The same minimum dominating set algorithm is first applied to the sub-graph for finding representative events, and then followed by a Steiner tree algorithm to make the selected events temporally smooth and coherent.

5. GLOBAL STORYLINE GENERATION

5.1 Text Snippet Graph Construction

Although each text snippet can be considered as an event, many of those are redundant. To remove the redundancy and obtain a set of representative events, we construct a graph $G = (V, E)$ with the given text snippets as the vertex set V, and add an edge between each pair of snippets which are likely to refer to the same event. Specifically, for two nodes $v_i, v_j \in V$, we first convert these two text snippets into two feature vectors as n-gram bags-of-words, then compute the cosine similarity between these two feature vectors. $e_{ij} = (v_i, v_j) \in E$ if and only if both the similarity of v_i and v_j is greater than a similarity threshold parameter α, and their distance calculated by their geocode is less than a distance threshold parameter $radius$. Note that it is this latter constraint that takes account of the spatial smoothness of events.

5.2 Identifying Events via Dominating Set

We identify the set of representative events in the original snippets with minimum redundancy by solving the minimum dominating set problem. 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 of 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 *Minimum Dominating Set* (MDS) problem is to find a dominating set with minimum size. MDS has been previously used to model multi-document summarization problem (Shen and Li, 2010). In our case, we use the MDS of text snippets to capture the representative events from the text snippets of disaster event descriptions. The MDS problem is known to be NP-hard but an efficient greedy algorithm by Johnson (Johnson, 1974) is known to achieve an approximation ratio of $H(d + 1)$, where d is the maximum degree of the graph and $H(n) = \sum_{i=1}^{n} \frac{1}{i}$ is the harmonic function.[3] The greedy algorithm is described in Algorithm 1 and was also used in (Shen and Li, 2010).

5.3 Storyline Generation by Connecting Dominating Objects via Linear Programming (LP)

Using Algorithm 1, we generate the dominating set of $G(V, E)$, m text snippets $d_1, \ldots, d_m$, as the representative events. Without loss of generality, the set of events are assumed to be in chronological order. To generate a global storyline capturing the major location change of the disaster, we select a sequence of nodes $o_1, o_2, \ldots, o_l$ from the representative events in chronological order. Intuitively, the generated storyline should also be in spatial coherence, reflecting the

[2] https://developers.google.com/maps/documentation/geocoding

[3] Johnson's greedy algorithm was initially designed for the SET COVER problem, but it is well-known that there is an L-reduction between MDS and SET COVER.

Algorithm 1 Greedy MDS Approximation Algorithm

INPUT: Graph $G = (V, E)$, MDS upper bound W
OUTPUT: dominating set S

$S \leftarrow \emptyset$
$T \leftarrow \emptyset$
while $|S| < W$ and $S \neq V(G)$ **do**
  for $v \in V(G) \setminus S$ **do**
    $s(v) \leftarrow |N(v) \setminus T|$
  end for
  $v^* \leftarrow \arg\max_v s(v)$
  $S \leftarrow S \cup \{v^*\}$
  $T \leftarrow T \cup N(v^*)$
end while

continuous location change of the disaster over time. Since a disaster is likely to affect adjacent areas in a similar fashion, the storyline should be coherent in content as well.

Based on the above discussions, we model the storyline generation problem by an integer linear program. To select a chain of nodes from $d_1, \ldots, d_m$, we use variables $\textit{node-active}_i \in \{0,1\}, i = 1 \ldots m$ to indicate whether d_i is included in the selected chain, and $\textit{next-node}_{ij} \in \{0,1\}, i,j = 1 \ldots m$ to indicate that d_i and d_j are two successive nodes (i.e., a transition) in the chain. The objective function aims to maximize the storyline's content coherence which is defined as the minimal similarity between two successive nodes along the storyline as shown below:

$$Coherence(o_1, o_2, \ldots, o_n) = \min_{i=1,2,\ldots,n-1} similarity(o_i, o_{i+1}).$$

We further impose the following set of constraints to model storyline's spatial coherence.

—**Chain Constraints:** The variables $\textit{node-active}_i$ and $\textit{next-node}_{ij}$ should be consistent, and the selected nodes should form a chain in chronological order.

// A node has at most one incoming edge and at most one
// outgoing edge:

$$\forall_j : \textstyle\sum_i \textit{next-node}_{i,j} \leq \textit{node-active}_j, \tag{1}$$

$$\forall_i : \textstyle\sum_j \textit{next-node}_{i,j} \leq \textit{node-active}_i. \tag{2}$$

// The number of active transitions is equal to the
// number of active nodes minus one:

$$\textstyle\sum_i \textit{node-active}_i - \sum_{i,j} \textit{next-node}_{i,j} = 1. \tag{3}$$

// The chain is ordered chronologically:

$$\forall_{i>j} : \textit{next-node}_{i,j} = 0. \tag{4}$$

// A transition of two node can not be active if
// there exists an active node between them:

$$\forall_{i<k<j} : \textit{next-node}_{i,j} \leq 1 - \textit{node-active}_k. \tag{5}$$

—**Length Constraints:** The selected chain should be in a reasonable length ranged between pre-defined minimum length threshold $\mathcal{L}_{min}$ and maximum length threshold $\mathcal{L}_{max}$.

$$\mathcal{L}_{min} \leq \textstyle\sum_i \textit{node-active}_i \leq \mathcal{L}_{max}. \tag{6}$$

—**Location Smoothness Constraints:** We require both pairwise and triple-wise smoothness of location change on the selected chain. Let $\mathcal{D}_{i,j}, i,j = 1, \ldots, m$ be the distance based pairwise location relationship between d_i and d_j, and $\mathcal{D}_{i,j} = 1$ if distance between d_i and d_j is less than a pre-defined distance parameter, $\mathcal{D}_{i,j} = 0$ otherwise. For triple-wise smoothness, let $\mathcal{A}_{i,j,k}$ be the angle based triple-wise location relationship, and $\mathcal{A}_{i,j,k} = 1$ indicates the angle constructed by three successive nodes d_i, d_j and event k is not an acute one, otherwise $\mathcal{A}_{i,j,k} = 0$. By not including in the chain three successive nodes of which the angle is acute, we excludes the back-and-forth events from the storyline and smooth the location change.

$$\begin{gathered} \text{// Distance of two successive nodes should be} \\ \text{// within some range} \\ \forall_i : \textstyle\sum_j (1 - \mathcal{D}_{i,j}) \cdot \mathit{next\text{-}node}_{i,j} \le 0. \end{gathered} \tag{7}$$

$$\begin{gathered} \text{// Three successive nodes can not construct} \\ \text{// an acute angle} \\ \forall_{i,j,k} : \mathit{next\text{-}node}_{i,j} + \mathit{next\text{-}node}_{j,k} \le 1 + \mathcal{A}_{i,j,k}. \end{gathered} \tag{8}$$

—**Minimal Similarity Constraints:** Let $\mathcal{S}_{ij}, i, j = 1\ldots, m$ be the cosine similarity between d_i and d_j. we can use the following constraints to find the similarity of the minimum similar transition *min-edge* among active transitions.

$$\forall_{i,j} : \mathit{min\text{-}edge} \le 1 - (1 - \mathcal{S}_{i,j}) \cdot \mathit{next\text{-}node}_{i,j} \tag{9}$$

—**The Objective Function:** Besides to maximize minimal similarity between two successive nodes along the storyline, we also try to make storyline as long as possible, so the objective function has the following form

$$\text{Maximize: } \mathit{min\text{-}edge} + \lambda \cdot l, \tag{10}$$

where λ is a coefficient parameter.

Although integer linear programming is NP-hard problem, there are efficient approximation algorithms and implementations such as IBM CPLEX[4], which is used for optimization in this paper.

6. LOCAL STORYLINE GENERATION

A global storyline presents a general high-level picture of how a disaster affects different areas when it hits these areas. To show how the disaster affects a specific area locally for a longer time period during preparation and recovery, we allow users to zoom-in to a node $node_x$ of the global storyline. Once a user clicks the node $node_x$, a new graph $G^L(V^L, E^L)$ will be constructed, which is an induced sub-graph of $G(V, E)$, where V^L includes all text snippet nodes which are close to $node_x$ according to their associated geocodes. For the graph $G^L(V^L, E^L)$, we employ the storyline generation method proposed in (Wang et al., 2012) to generate a storyline for the selected area.

6.1 Augmented Multi-view Graph Construction

Definition 2. *A* multi-view graph *is a triple $G = (V, E, A)$, where V is a set of vertices, E is a set of undirected edges, and A is a set of directed edges.*

Different from the global storyline generation where the temporal and spatial information of text snippets are modeled by integer linear programming, here we incorporate temporal information in an augmented multi-view graph $G^L = (V^L, E^L, A)$ from $G^L = (V^L, E^L)$, where A is a set of directed edges for temporal relationship between events. To define edges in A, we introduce two additional parameters $\tau_1, \tau_2, 0 < \tau_1 < \tau_2$. For every pair of nodes o_i, o_j in V, we draw an arc from o_i to o_j if $\tau_1 < t_j - t_i < \tau_2$, where t_i, t_j are the timestamps of o_i and o_j, respectively.

[4] `http://www.ibm.com/software/commerce/optimization/cplex-optimizer/`

6.2 Generating Storylines via Directed Steiner Tree

Similar to generating global storylines, after extracting a dominating set of $G^L = (V^L, E^L)$ which represent the main content topics, we need to generate a storyline capturing the temporal and structural information of the local event descriptions. To tackle this problem, we use the concept of Steiner Tree. A *Steiner tree* of a graph G with respect to a vertex subset X is the edge-induced subtree of G that contains all the vertices in X with minimum cost, where the cost is often measured by the size of the tree.

PROBLEM 1. *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 of G rooted at v_0 containing X with the smallest total vertex weight.*

This problem is known to be NP-hard since the undirected version is already NP-hard. While the undirected version has been well studied, much less work has been done on directed version (Charikar, Chekuri, Cheung, Dai, Goel, Guha, and Li, 1999). An intuitive solution for this problem is to find the shortest path from the root to each of the terminal and then merge the paths. Of course, this does not guarantee the optimal solution.

We make use of an algorithm due to Charika et al. (Charikar et al., 1999). The algorithm takes a level parameter $i \geq 1$. In addition, it takes as input the target terminal set Y, the root r, and the required number of nodes to cover, k. When $i = 1$, it leads to the intuitive solution: i.e., selecting the top k shortest path from the root to k nodes and return the union of those paths. Let the length of every arc $(u, v) \in A$ is 1. We will make initial call of $A_i(k, v_0, X)$ with X is the dominating set calculated by Algorithm 1 based on graph G, v_0 is the event among X with the earliest timestamp, and k is $|X|$, the size of X. We interpret the output tree as a local storyline evolving from the root event to all the other dominating events. For a constant i, the algorithm is known to run in polynomial time and produces an $O(k^{\frac{1}{i}})$-approximate solution (Charikar et al., 1999).

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

INPUT: $G = (V, A)$: directed multi-view graph
X : target vertex set X
$r \in X$: the root X
$k \geq 1$: the target size X
OUTPUT: T: a Steiner tree rooted at r covering at least k vertices in X

$T = \emptyset$
while $k > 0$ **do**
  $T_{best} \leftarrow \emptyset$
  $cost(T_{best}) \leftarrow \infty$
  for each vertex $v, (v_0, v) \in A$, and $k', 1 \leq k' \leq k$ **do**
    $T' \leftarrow A_{i-1}(k', v, X) \cup \{(v_0, v)\}$
    if $cost(T_{best}) > cost(T')$ **then**
      $T_{best} \leftarrow T'$
    end if
    $T \leftarrow T \cup T_{best}$
    $k \leftarrow k - |X \cap V(T_{best})|$
    $X \leftarrow X \setminus V(T_{best})$
  end for
end while
return T

7. EVALUATION

7.1 Datasets

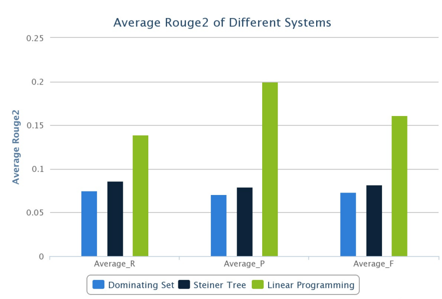
Table I. Statistics of the datasets.

keyword	#documents	#text snippets
Hurricane Katrina	800	1572
Hurricane Sandy	795	2253
Hurricane Irene	691	2186

Table II. Example on events extracted from document using entity recognition

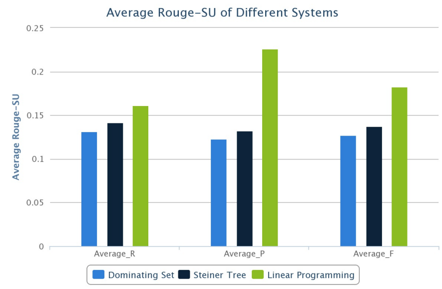
content	time	location
This photo made available by the New Jersey governor's office shows flooding and damage in Seaside Heights, N.J. on Oct. 30, 2012 after super-storm Sandy made landfall in the state.	2012-10-30	New Jersey — Seaside Heights N.J.
October 22, 2012 - Sandy develops into a tropical storm in the Caribbean Sea.	2012-10-22	Caribbean Sea
October 24, 2012 - Hurricane Sandy makes landfall near Kingston, Jamaica, with winds of 80 mph.	2012-10-24	Kingston Jamaica
By Patrick Clark September 26, 2013 Business owners pile muddy furniture outside their building off Canon Avenue in Manitou Springs, Colo.	2013-09-26	Manitou Springs Colo.

We collect datasets from Bing News Search[5] using keywords about three major hurricanes in the last ten years, Hurricane Katrina, Hurricane Irene, and Hurricane Sandy, to evaluate our storyline generation system. For the search results returned from Bing News Search, we extract the text content from the corresponding web pages. Basic statistics about the datasets are shown in Table I, and some examples of further extracted text snippets are shown in Table II.



. Average Recall, Precision, F-1 of ROUGE-2.

[5]http://news.bing.com



. Average Recall, Precision, F-1 of ROUGE-SU4.

7.2 Summarization Performance of Grobal Storylines

To evaluate the quality of global storylines generated by our proposed framework, one of the authors manually composed global storylines for the three hurricane disasters, which are compared with system-generated storylines using ROUGE (Lin and Hovy, 2003) toolkit (version 1.5.5). ROUGE is widely used by DUC for summarization performance evaluation. It measures the quality of a summary by counting the unit overlaps between the candidate summary and a set of reference summaries. Several automatic evaluation methods are implemented in ROUGE, such as ROUGE-N, ROUGE-L, ROUGE-W and ROUGE-SU. ROUGE-N is an n-gram recall computed as follows:

$$\text{ROUGE-N} = \frac{\sum_{S \in \text{ref}} \sum_{\text{gram}_n \in S} \text{Count}_{\text{match}}(\text{gram}_n)}{\sum_{S \in \text{ref}} \sum_{\text{gram}_n \in S} \text{Count}(\text{gram}_n)}, \tag{11}$$

where n is the length of the n-gram, and ref stands for the reference summaries. $\text{Count}_{\text{match}}(\text{gram}_n)$ is the maximum number of n-grams co-occurring in a candidate summary and the reference summaries, and $\text{Count}(\text{gram}_n)$ is the number of n-grams in the reference summaries. ROUGE-SU4 is based on skip-bigram plus unigram, where skip length is 4.

We compare the global storylines generated by our proposed method considering geo-spatial information with the results from the following methods: 1) Steiner tree based storyline generation (Wang et al., 2012), which does not consider geo-spatial information, and 2) dominating set based summarization method (Shen and Li, 2010), which is a standard multi-document summarization method. Figure 2 and Figure 3 show the performance comparison of the three methods using ROUGE-2 and ROUGE-SU4, respectively.

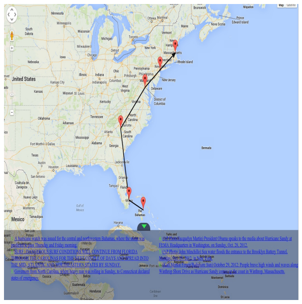
One can see that the Streiner tree based storyline generation method outperforms the pure multi-document summarization method that does not incorporate the temporal information. Our proposed storyline generation method, which considers not only the temporal information but also the spatial information, performs the best among all three methods.

7.3 A Case Study

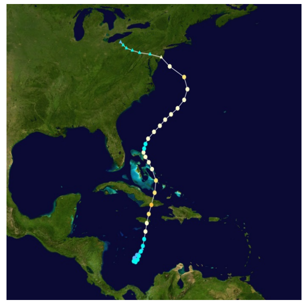
A case study is conducted to show the effect of the storylines generated using our proposed method. We draw the global storyline generated by our proposed method using Google Map API (on the left in Figure 4) and compare it to the storm paths downloaded from Wikipedia (on the right in Figure 4).

One can see that the paths in our generated storylines are similar to those of the ground truth. The differences are 1) In addition to showing the real paths, our generated storylines can provide more information about how the hurricanes affect different areas, 2) the generated storylines not only show how hurricanes move but also present text descriptions about their statuses and the damages they cause along their paths. With the geo-temporal storyline, users can easily obtain the overall evolution of a disaster.

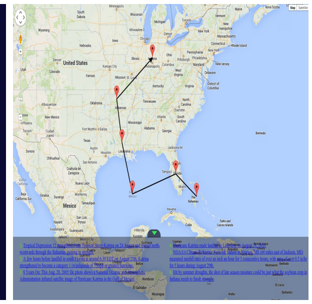
Figure 5 shows an example of a local storyline when the user is interested in a specific area like



(a) Hurricane Sandy experiment

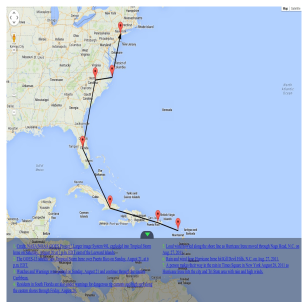


(b) Hurricane Sandy from wikipedia



(c) Hurricane Katrina experiment

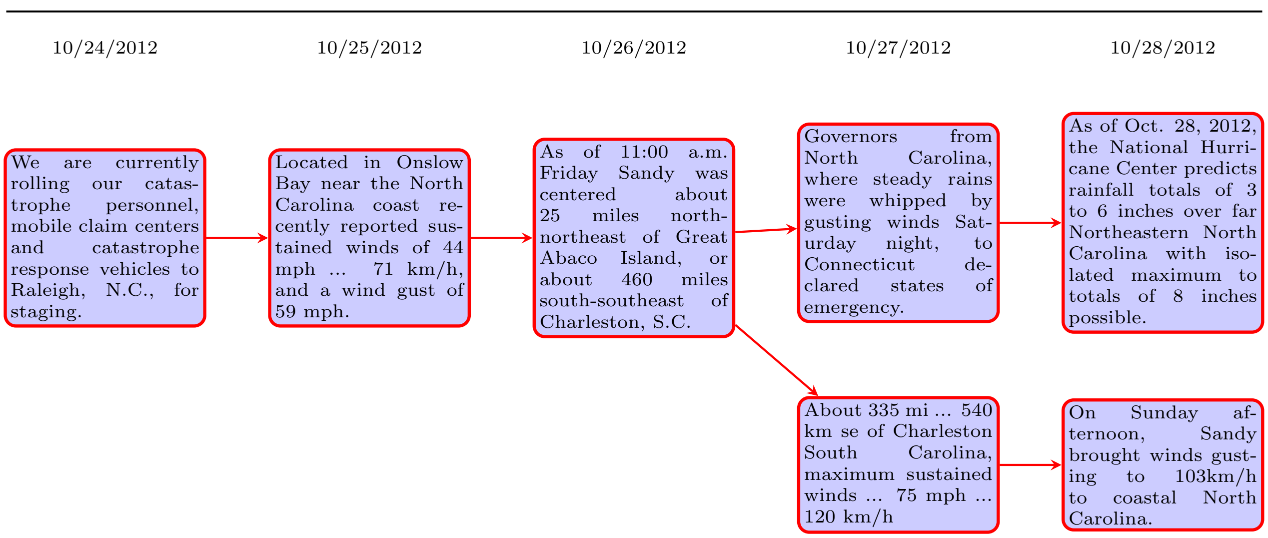
(d) Hurricane Katrina from wikipedia



(e) Hurricane Irene experiment

(f) Hurricane Irene from wikipedia

. Experimental result of Hurricane Sandy, Katrina and Irene compared to Wikipedia.



. An illustrative example of the local storyline for the area of the Carolinas during the Hurricane Sandy.

Carolina during Hurricane Sandy. The user can find out how Hurricane Sandy affects the area during a period of time and learn about different topics such as wind and rain.

8. CONCLUSION

In this paper, we present a novel storyline framework for summarizing multiple disaster-related documents to generate a hierarchical storyline which improves situation awareness during or after a disaster. We organize the storyline in a two-layer structure that naturally describes the situations of a large-scale disaster. One special feature of our framework is that not only temporal but also spatial factor are considered when generating the global storyline, thus capturing the evolution of a disaster better.

For future work, we will extend our framework to deal with more natural disaster types such as earthquakes and even man-made disasters. To make our system more practical in a real-time disaster environment, we plan to include Twitter streams as another data source.

Acknowledgment

The work was supported in part by the National Science Foundation under grants HRD-0833093, CNS-1126619, and IIS-1213026, the U.S. Department of Homeland Security under grant Award Number 2010-ST-062000039, and Army Research Ofce under grant number W911NF-1010366 and W911NF-12-1-0431.

References

Allan, J. 2012. *Topic detection and tracking: event-based information organization.* Vol. 12. Springer Science & Business Media.

Alonso, O., Baeza-Yates, R., and Gertz, M. 2009. Effectiveness of temporal snippets. In *WSSP Workshop at the World Wide Web ConferenceWWW*. Vol. 9. 13.

Chang, A. X. and Manning, C. D. 2012. Sutime: A library for recognizing and normalizing time expressions. In *Lrec*. Vol. 2012. 3735–3740.

Charikar, M., Chekuri, C., Cheung, T.-y., Dai, Z., Goel, A., Guha, S., and Li, M. 1999. Approximation algorithms for directed steiner problems. *Journal of Algorithms 33,* 1, 73–91.

E-Teams. 2004. `http://www.nc4.us/ETeam.php`.

Erkan, G. and Radev, D. R. 2004. Lexpagerank: Prestige in multi-document text summarization. In *EMNLP*. Vol. 4. 365–371.

Finkel, J. R., Grenager, T., and Manning, C. 2005. Incorporating non-local information into information extraction systems by gibbs sampling. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics.* Association for Computational Linguistics, 363–370.

GeoVISTA. 2010. http://www.geovista.psu.edu.

Jiang, Y., Perng, C.-S., and Li, T. 2011. Natural event summarization. In *Proceedings of the 20th ACM international conference on Information and knowledge management.* ACM, 765–774.

Johnson, D. S. 1974. Approximation algorithms for combinatorial problems. *Journal of computer and system sciences 9,* 3, 256–278.

Lavrenko, V., Allan, J., DeGuzman, E., LaFlamme, D., Pollard, V., and Thomas, S. 2002. Relevance models for topic detection and tracking. In *Proceedings of the second international conference on Human Language Technology Research.* Morgan Kaufmann Publishers Inc., 115–121.

Li, J., Li, L., and Li, T. 2012. Multi-document summarization via submodularity. *Applied Intelligence 37,* 3, 420–430.

Li, L. and Li, T. 2014. An empirical study of ontology-based multi-document summarization in disaster management. *Systems, Man, and Cybernetics: Systems, IEEE Transactions on 44,* 2.

LI, L., WANG, D., SHEN, C., AND LI, T. 2010. Ontology-enriched multi-document summarization in disaster management. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 819–820.

LIN, C.-Y. AND HOVY, E. 2003. Automatic evaluation of summaries using n-gram co-occurrence statistics. In *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*. Association for Computational Linguistics, 71–78.

MAKKONEN, J., AHONEN-MYKA, H., AND SALMENKIVI, M. 2004. Simple semantics in topic detection and tracking. *Information retrieval 7,* 3-4, 347–368.

MANI, I. AND MAYBURY, M. T. 2001. *Automatic summarization*. J. Benjamins Publishing Company New York, USA.

RADEV, D. R., HOVY, E., AND MCKEOWN, K. 2002. Introduction to the special issue on summarization. *Computational linguistics 28,* 4, 399–408.

RADEV, D. R., JING, H., AND BUDZIKOWSKA, M. 2000. Centroid-based summarization of multiple documents: sentence extraction, utility-based evaluation, and user studies. In *Proceedings of the 2000 NAACL-ANLP Workshop on Automatic Summarization*. Association for Computational Linguistics, 21–30.

RADEV, D. R., JING, H., STYŚ, M., AND TAM, D. 2004. Centroid-based summarization of multiple documents. *Information Processing & Management 40,* 6, 919–938.

SAGGION, H., BONTCHEVA, K., AND CUNNINGHAM, H. 2003. Robust generic and query-based summarisation. In *Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics-Volume 2*. Association for Computational Linguistics, 235–238.

SHAHAF, D., GUESTRIN, C., AND HORVITZ, E. 2012. Trains of thought: Generating information maps. In *Proceedings of the 21st international conference on World Wide Web*. ACM, 899–908.

SHEN, C. AND LI, T. 2010. Multi-document summarization via the minimum dominating set. In *Proceedings of the 23rd International Conference on Computational Linguistics*. Association for Computational Linguistics, 984–992.

SHEN, C., LI, T., AND DING, C. H. 2011. Integrating clustering and multi-document summarization by bi-mixture probabilistic latent semantic analysis (plsa) with sentence bases. In *AAAI*.

USHAHIDI. 2012. http://www.ushahidi.com/.

WANG, D., LI, T., OGIHARA, M., ET AL. 2012. Generating pictorial storylines via minimum-weight connected dominating set approximation in multi-view graphs. In *AAAI*.

WANG, D., LI, T., ZHU, S., AND DING, C. 2008. Multi-document summarization via sentence-level semantic analysis and symmetric matrix factorization. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 307–314.

WEBEOC. 2002. `http://www.esi911.com/home`.

WEI, F., LI, W., LU, Q., AND HE, Y. 2008. Query-sensitive mutual reinforcement chain and its application in query-oriented multi-document summarization. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 283–290.

ZHENG, L., SHEN, C., TANG, L., ZENG, C., LI, T., LUIS, S., AND CHEN, S.-C. 2013. Data mining meets the needs of disaster information management. *IEEE Transactions on Human-Machine Systems 43,* 5, 451–464.