Search Results using Timeline Visualizations

Omar Alonso
Dept. of Computer Science
University of California, Davis
oralonso@ucdavis.edu

Michael Gertz
Dept. of Computer Science
University of California, Davis
gertz@cs.ucdavis.edu

Ricardo Baeza-Yates
Yahoo! Research
Barcelona, Spain
rbaeza@dcc.uchile.cl

1. OVERVIEW

As search applications keep gathering new and diverse information sources, presenting relevant information anchored in time becomes more important. Time can help in describing the context of a document or document collection or in recreating a particular historical period. Temporal information is available in every document either explicitly or implicitly. Recognizing such temporal information and exploiting it for document retrieval and presentation purposes are important features that can significantly improve the functionality of today's search applications.

A look at functionality of any of the current search engines shows that temporal aspects of documents are exclusively used to sort the hit list by date, which is primarily the date a Web page or a document has been created or last modified. In this demonstration, we present a new approach in which search results are arranged in a well-defined timeline using a few existing visualization metaphors.

To take advantage of time related information for IR purposes, we process each document in a document annotation pipeline that extracts temporal expressions and produces output in TimeML. Recently TimeML has emerged as a standard markup language for events and temporal expressions in natural language [1].

As the basis for anchoring documents in time, we assume a discrete representation of time based on the Gregorian Calendar, with a single day being an atomic time interval called *chronon*. Our *base timeline*, denoted T_d, is an interval of consecutive day chronons. For example, the sequence "March 12, 2002, March 13, 2002, March 14, 2002" is a contiguous subsequence of chronons in T_d.

For a query, the search engine returns an intermediate representation of the hit list grouped by year. This is later transformed to a particular visualization format. The timeline construction is based on a clustering algorithm that uses temporal expressions (chronons), which have been extracted from the documents during the above annotation process and anchored in a timeline.

For our system development, we use the TimeBank 1.2 corpus, which contains 183 news articles that have been annotated using TimeML with temporal expressions related to events, times and temporal links between events and times.

In the first visualization, we use Simile TimeLine [2] to construct a timeline that consists of two *bands* representing different time scales: decade and year. Both bands are synchronized such that panning one band also scrolls the other. In the second visualization, we use Inxight's TimeWall [3] to define *cards* with attributes that allow their placement in the wall along the timeline. In both metaphors the user can click a particular document and the content of the document is shown in a separate frame.

Both approaches not only demonstrate the high utility of temporal information associated with search results and documents, but they also indicate an important direction in the development of future search engine features.

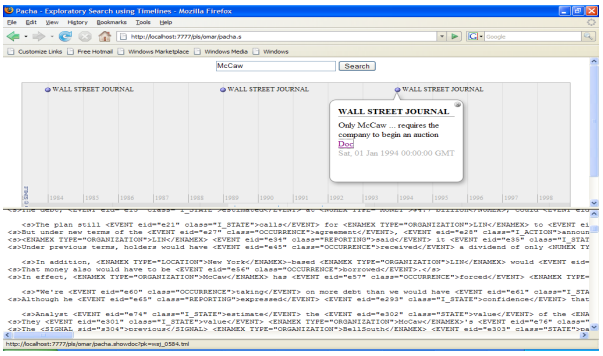

Figure 1: Search Results with SIMILE

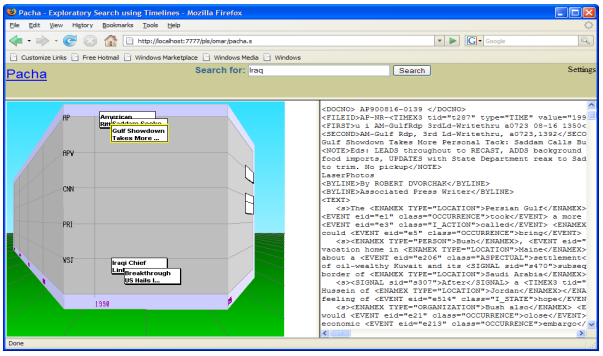

Figure 2: Search Results with TimeWall

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Copyright 200X ACM X-XXXXX-XX-X/XX/XX ...$5.00.

2. REFERENCES

[1] TimeML 1.2.1 Specification: `http://www.timeml.org`
[2] TimeLine: `http://simile.mit.edu/timeline/`
[3] TimeWall: `http://www.inxight.com/products/sdks/tw/`