Author: Christian Lendl* and Peter Purgathofer**

Challenges in Interactive Time-Based Information Visualization

Keywords: historical data; online maps; mashup; timeline

Summary: This research paper tries to raise some key questions in the quest for interactive time-based information visualization. Since our ultimate goal is to specify and develop a web-based system for this task that will most likely be a mashup using existent geographical data, we first explain and discuss mashups. We then try to identify a number of use cases for such a system, followed by a state of the art overview. We then try to devise a categorization for the display of and interaction with time-based geospatial data, and close with a discussion of particular challenges identified in the context of the project described in this paper.

Introduction

Since there is cartography, artisans have tried to visualize time-based data through the means of static maps. The peculiarities of the medium have led to numerous approaches to overcome its shortcomings. And while we have learned to read these static representations of dynamic information quite well, recent advances in computer generated imaging push for new approaches to old questions. Also, current ways of interacting with online maps (e.g. Google Maps) and new technical standards (e.g. HTML5, GML) create a variety of chances especially for web-based map services. The increasing number of available data sources makes it possible to gain new insights from enriching plain geospatial data with (especially time-based) data from various other sources and displaying it on an interactive map.

Mashups

A mashup is a web application that combines data from various sources into a new service that fulfills specific needs. One of the first popular mashups was HousingMaps (www.housingmaps.com). Its creator Paul Rademacher combined real estate lists from Craigslist with Google's map service. Houses, apartments and rooms that are for sale or rent are displayed on a Google Map. Originally neither Craigslist nor Google had designated their services for this kind of use, HousingMaps simply combined the different services to a new one.

Current standards like Extensible Markup Language (XML) simplify the handling of data. In addition to these standards, there are numerous open source libraries for various programming languages (e.g. Java, PHP) that can be used by programmers. Moreover the number of web services providing access to their data via Application Programming Interface (API) is increasing. These APIs can be very well used for creating a new mashup.

* Graduand, Institute for Design & Assessment of Technology, Vienna University of Technology, Austria [christian.lendl@tuwien.ac.at]

** Professor, Institute for Design & Assessment of Technology, Vienna University of Technology, Austria [peter.purgathofer@tuwien.ac.at]

Complex mashups have to manage five steps according to (Tuchinda 2008): *Data Retrieval* deals with extracting the data from the source (e.g. HTML-Page or API). Structure and location of the original data influence the complexity of the data extraction. *Source Modeling* targets the assigning of attribute names for each column of the various data sources. The next step, *Data Cleaning*, is needed to transform the extracted data from all sources into one common format (e.g. transforming “120 participants” to “120”). *Data Integration* gives details how the data sources are combined. For Example, to create a list with all races won by members of a certain team a team member list and a list of all races have to be combined (with the member names as key values). The final step is *Data Visualization* which displays the combined data (e.g. on a map).

Normally the creation of a mashup needs certain programming skills - especially accessing the data sources and visualizing the combined data. In order to simplify these steps numerous providers offer services that help end users create their own mashups. For Example, Yahoo Pipes (pipes.yahoo.com) offers a variety of widgets that have different tasks (e.g. filtering data) and can be combined to solve a certain task.

Currently there are two topics in mashup-related research (Zang 2008). The first is *information systems*, where mashups are built to fulfill certain needs. The second topic is *end-user programming* where the goal is to allow end-users with limited programming skills to use professional programming techniques for their tasks (e.g. with tools that support mashup creation) (Tuchinda 2008).

The website ProgrammableWeb (www.programmableweb.com) provides an overview of various mashups and APIs. Currently (January 26, 2010) there are 1.643 available APIs and 4.609 existing mashups listed (ProgrammableWeb 2010). Nearly 50% (2.249) of all mashups use one of the main mapping APIs[2], while Google Maps as the most used API (1.899 mashups) is used nearly four times as much as second placed FlickR API (497 mashups). This fact shows the variety of mashups using maps.

Use Cases for the Interaction with Time-Based Geospatial Information

There are various use cases for a mashup that combines and displays spatial and time-based information. A main use case is the display of historical data (e.g. rise and fall of empires displayed in a historical atlas). An elementary overview of the most important conflicts and wars in history visualized on a map and a timeline can be found at ConflictHistory (www.conflicthistory.com).

[2] Google Maps, Microsoft Virtual Earth, Yahoo Maps, Google Earth, OpenLayers and OpenStreetMap

Figure 1: map and timeline displaying conflicts from 1940-1944 (source: www.conflicthistory.com)

The display of historical economic data is a special case (e.g. combination of map/timeline with diagrams). Also continuously measured data can be displayed on such a mashup. For example, flight altitude and flight paths of aircraft can be visualized on a map/timeline. A basic realization (without timeline) is GMaps Flight Tracker (www.gmapsflighttracker.com). The display of real-time events such as (breaking) news or social media messages (e.g. twitter) on a map can help getting a clearer understanding of the spatial coherences of these events. For example, the service twittermap (www.twittermap.de) displays tweets in real-time on a map. As more and more devices automatically ad a geolocation to the recorded media (e.g. photographs with embedded GPS-position), geolocated and time stamped media (e.g. pictures, video or audio material) can also be a use case for such a mashup. For example, two of the biggest photo-communities, Flickr (www.flickr.com) and Panoramio (www.panoramio.com) provide map services where all geolocated pictures can be viewed.

Related Work

TimeMap (www.timemap.net) is a research project by the University of Sydney with a focus on developing new methods of handling temporal data in a geographical information system (Johnson 2003). At the project start in 1997 three research goals were defined (Johnson 1997): “*(1) a methodology for recording time-based cultural features, (2) an interface for displaying time-based maps, and (3) the generation of map-based animations to display time depth.*” A major development is a Java-based software that allows users to create time-based maps with new or existing data (e.g. ECAI-databases) and display them in a web- or desktop-application.

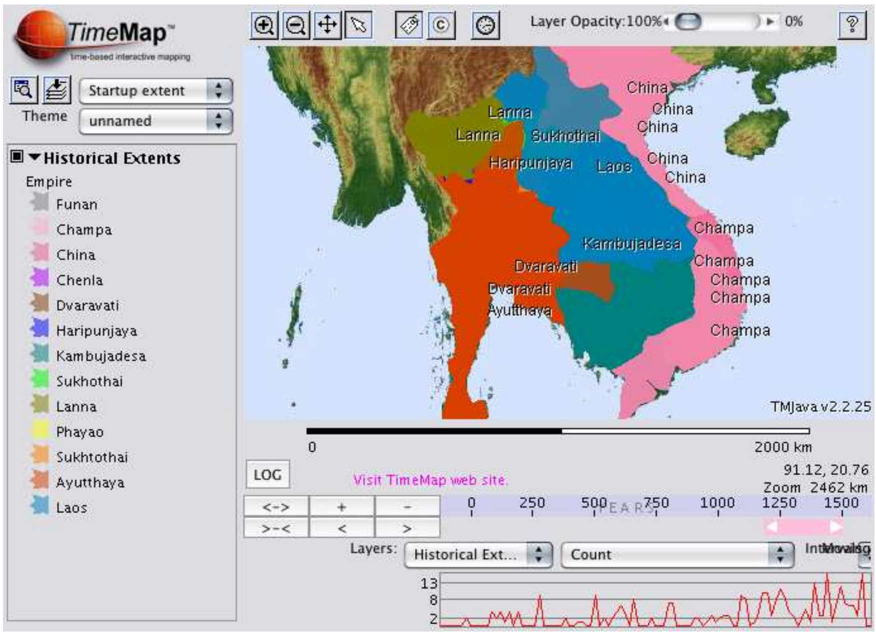

Figure 2: History of Southeast Asia (source: TimeMap Open Source Consortium 2010)

DIGMAP (www.digmap.eu) was a research project by the Lisbon Technical University from 2006 to 2008 that focused on digitized historical maps. Numerous web services were developed around an online-catalogue for historical maps. The maps were enriched with both extracted data (e.g. information from the digitized images) and external metadata from numerous European national libraries (Martins 2008).

Besides these two research projects, there is a number of online libraries that provide collections of digitized historical maps and/or rendered maps with historic data. For example, the Institute of European History (University of Mainz, Germany) runs a server for digital historical maps (www.ieg-maps.uni-mainz.de). This server provides a variety of rendered historical maps for various topics and eras. Some are animated (e.g. animated map of the development of railways in Germany from 1835-1885).

It is remarkable, that there is no common standard for storing geospatial time-based data. The Geography Markup Language (GML) is an encoding standard for expressing geographical features based on XML (OpenGIS 2010). It also includes attributes for points of time and time intervals, but these attributes may not cover all use cases and can be seen as a first step.

Complex maps include a variety of geospatial and time-based information. For example, Figure 3 shows a map of Christianity in Europe from 600 to 1500. It contains one-dimensional points (e.g. bishoprics, monasteries), two-dimensional lines (e.g. borders, paths) and two-dimensional surfaces (e.g. spreading of other religions). This spatial information is combined with time-based information in the form of numerous dates and time intervals. A considerable amount of time-based and spatial information is packed into this map leading to confusion if no additional help or descriptive text is provided.

Figure 3: map of Christianity in Europe (source: Barraclough 1994: 38)

Categories of Time-Based Geospatial Information

In the following section we propose a two-dimensional taxonomy of time-based geospatial information. It serves as a vehicle to identify domain-specific challenges. The following table shows a combination of spatial and time-based data. It is based on the following assumptions: (1) All spatial data is two-dimensional and (2) all data is available with adequate accuracy. [3]

	Data is 0D (point)	Data is 1D (line)	Data is 2D (area)
One point in time	modification of status (e.g. capture of a city)	modification of boundary line (e.g. national border after war)	territorial modification (e.g. dissolution of the German Democratic Republic)
Two points in time	e.g. population increase, economic data	e.g. troop movements	e.g. migration period
n points in time	special case of "two points of time"		
Interval	continuously measured data, e.g. from weather station	only possible if spatial resolution of data is higher than display resolution, e.g. sensor network	

Table 1: Combination of spatial and time-based data

[3] It is interesting to see that these assumptions were only identified after the categorization was developed. This shows how such a taxonomy can aide in the identification of domain-specific challenges.

Domain-specific Challenges

We are still in the process of identifying domain-specific challenges in the visualization of and interaction with time-based geospatial data. The following issues are among the problems we have already determined. One goal of this paper is to evoke an expert's discussion in order to understand better the particular challenges in the design of such a system.

(1) At the moment, our model categorizes data that is located on the surface of earth. Three-dimensional information such as data gathered from a weather balloon does not fit into this schema.

(2) Frequently, two events (such as the beginning and the end of a migration period) are connected through a nonlinear change or movement, where visualization cannot rely on linear interpolation. Still, a way has to be found to display data in between events, making this distinctive feature apparent.

(3) Historic data differs as far as their precision is concerned; some events can be pinned to the minute, but most historic information is much more blurred, and in many instances accuracy means a couple of decades or even a century. While for example the day of death of a famous historic person is usually reported precisely, his time of birth can often only be defined within a range of months or even years.

(4) Neither historic information is objectively reported, nor weather data from an observatory. Instead, such information relies on artifacts, such as books or reports, which are always subject to distortions. This source of uncertainty has to be regarded, and we have to think of ways to visualize the degree to which the available information has to be read as uncertain, subjective or tainted.
The core challenge of interaction design will be to find interaction methods and widgets to allow more than just the controlled display of predefined animations. While most of the challenges described above are highly domain specific and as of yet untackled, there is significant prior work in the field of designing the interaction with time-based data, e.g. the work on dynamic query interfaces by Ben Shneiderman (Shneiderman, 1994).

Conclusion

In order to design the interaction with time-based geospatial information, one has to move beyond the usual scope of interactive visualizations using maps due to the specific properties of knowledge in the domain of historic sciences. Uncertainty and inaccuracy, multiplicity, often inconsistent sources, and the very nature of the representation of multi-dimensional geospatial information are all challenges that make visualization difficult, and interacting with visual information harder than anticipated.

References

Barraclough, G. (Editor) (1994). *"Times" Concise Atlas of World History*. Times Books, London, UK, 38

Johnson, I. (1997). Mapping the fourth dimension: the TimeMap project. In *Proceedings of the 1997 Computer Applications in Archeology Conference.* BAR International Series. Birmingham, UK.

Johnson, I. and Wilson, A. (2003). The TimeMap Project: Developing Time-Based GIS-Display for Cultural Data. In *Journal of GIS in Archeology*. Volume I, April 2003. 125-135

Martins, B., Borbinha, J., Pedrosa, G., Gil, J., and Freire, N. (2007). Geographically-aware information retrieval for collections of digitized historical maps. In *Proceedings of the 4th ACM Workshop on Geographical information Retrieval* (Lisbon, Portugal, November 09 - 09, 2007). GIR '07. ACM, New York, NY, 39-42

Open Geospatial Consortium (2010). *Geography Markup Language*, http://www.opengeospatial.org/standards/gml, accessed January 29, 2010

ProgrammableWeb (2010). *API Directory*, http://www.programmableweb.com/apis/directory, accessed January 26, 2010

Shneiderman, B. (1994). Dynamic queries for visual information seeking. In *IEEE Software* 11, 6 (1994), 70-77

TimeMap Open Source Consortium (2010), *Map2: Customizing TMJava*, http://timemap.net/index.php?option=com_content&task=view&id=133&Itemid=157, accessed January 28, 2010

Tuchinda, R., Szekely, P., and Knoblock, C. A. (2008). Building Mashups by example. In *Proceedings of the 13th international Conference on intelligent User interfaces* (Gran Canaria, Spain, January 13 - 16, 2008). IUI '08. ACM, New York, NY, 139-148

Zang, N., Rosson, M., and Nasser, V. (2008). Mashups: who? what? why?. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems* (Florence, Italy, April 05 - 10, 2008). CHI '08. ACM, New York, NY, 3171-3176