

***Integration of Geospatial and Online Data for Crisis Management
(USC's TerraWorld: An Integrated Virtual Repository of Online and
Geospatial Information Sources)***

**Cyrus Shahabi
Dept of Computer Science
University of Southern California
<http://infolab.usc.edu>**

With the huge amount of geospatial data now available (at USC we have our own digital library with image, map, vector, elevation, and point data for the entire world) and the huge amount of data available on the Web, there is a terrific opportunity to exploit the integration of these two types of information sources. The challenge is that the dynamic integration of online data and geospatial data is beyond the state of the art of existing integration systems. We are embarking on an effort to develop a general framework for integrating online and geospatial data sources.

The proposed dynamic integration of online and geospatial data sources will directly support the mission of developing revolutionary information technology to transform the nation's ability to respond to unexpected events. Consider, for example, the information needs of emergency personnel that respond to the crash of a large aircraft in a heavily populated area. Today, these needs are addressed in ad hoc fashion, with lack of information often hampering emergency response efforts. In the recent crash of an aircraft on Long Island, the news reported that despite knowing the precise location where the plane went down, emergency workers did not know what the plane had hit until they arrived on the scene. This information is directly available in the New York State property tax assessment site, but today there is no way to dynamically query this source by geocoordinates or to place this information on a map. In practice, there is a great deal of information that emergency response crews need to immediately access in order to properly respond to such an event. They would like to know what structures were hit. Are they businesses or residences? If they are businesses, what type and do they have any toxic substances on site? Emergency response crews would like to know the locations and distances to hospitals, fire stations, police stations, schools, or other structures that could be used to set up first aid stations if a large number of people are hurt. Emergency response crews would want to find out if there are any trains or bus lines that run through the crash site and if so, they would like to obtain the relevant service schedules to facilitate rerouting the trains and buses. These are just a few examples of the types of information that might be required.

In this talk, I survey some of our current efforts in accurately fusing spatial datasets and online sources so that we can support a rich set of queries that could not have been answered given any of these datasets in isolation. I focus on one of the main challenges in accurately integrating geo-spatial data from different data sources. The challenge is that spatial data obtained from various data sources may have different projections and different accuracy levels. We have proposed techniques to automatically integrate diverse geospatial data by exploiting what is known or can be inferred about each of the data

sources. For the imagery, we can exploit image metadata (such as ground resolution and geo-coordinates) as well as the color of imagery. For vector data of roads, we can exploit metadata about the vectors, such as address ranges, road names, or even the number of lanes and type of road surface. For raster maps, we can exploit map scales, map geo-coordinates or perform image processing techniques (such as edge-detection) at the pixel level to extract primitive entities, such as line intersections. I conclude the talk by a brief survey of the related work to the focused topics.

Bio:

Cyrus Shahabi is currently an Associate Professor and the Director of the Information Laboratory (InfoLAB) at the Computer Science Department and also a Research Area Director at the Integrated Media Systems Center (IMSC) at the University of Southern California. He received his Ph.D. degree in Computer Science from the University of Southern California in August 1996. He has two books and more than eighty articles, book chapters, and conference papers in the areas of databases and multimedia. Dr. Shahabi's current research interests include Streaming Architectures and Multidimensional Databases. He was the program committee chair of ACM WIDM'99 workshop and the local chair of ACM SIGMETRICS'2002. He is currently on the editorial board of ACM Computers in Entertainment magazine. He is also serving on many conference program committees such as ACM GIS 2004, ACM CIKM 2004, IEEE ICDE 2004 and ACM SIGMOD 2004.

Dr. Shahabi is the recipient of the 2003 National Science Foundation CAREER Award for his research on Immersive Sensor Data Streams. In 2001, he also received an award from the Okawa Foundations.