SenseWeb: Sharing and Browsing Environmental Changes in Real-time

Liqian Luo, Aman Kansal, Suman Nath, and Feng Zhao

Microsoft Research

In 2008, the world was hit by some of the worst natural disasters. The Cyclone Nargis killed over 140,000 people and led to 10 billion financial losses. The Great Sichuan Earthquake caused over 60,000 deaths and 20 billion of estimated property damages. How do we get prepared for natural hazards? How do we enhance our resilience to natural hazards? Part of the answer lies in innovative techniques to monitor, model, and predict environmental changes that potentially lead to natural hazards.

SenseWeb addresses this need by providing a scientific investigation platform that enables global sharing of sensory streams and real-time exploration of environmental measurements. By allowing users to create heterogeneous sensors and register distributed data feeds, SenseWeb is able to federate individually deployed sensor networks into a large-scale, unified sensing system across the globe. SenseWeb provides a web-based frontend, *SensorMap*, for end-users to browse real-time measurements, to discover their spatial correlations, as well as to explore historic trends.

1. Application Scenarios

SenseWeb is targeted at facilitating scientists throughout the whole life cycle of environmental experiments, as detailed by the following scientific investigation scenarios. The scenarios are based on our actual experience in SwissEx, a collaboration of environmental science projects in Switzerland to study environmental hazards from floods, landslides, to earthquakes and avalanches. Figure 1 and Figure 2 are examples of the actual usage of SensorMap in SwissEx.

Deployment Planning: Marc, a renown hydrologist, is in his office and wishes to review existing datasets that have been captured in the past year at the Le Genepi rock glacier field deployment in order to plan the deployment for this year. He wants to better understand the interaction between the rock glacier and the atmosphere, in particular how the wind patterns are driving the ventilation of the rock glacier. To do this, he generates a visualization of the temperature difference on the SensorMap interface based on the existing datasets. He is surprised about the large deviations at some locations and decides to concentrate more stations there. When visualizing the measurements of the rain gauges he realizes that they gave mostly uniform measurements and decides to reduce the number of rain gauge sensors.

Experiment Monitoring: The sensor stations have been deployed as planned. Through SensorMap Marc and his group can at any time observe the current measurements. One evening Marc receives a warning mail generated by the underlying data stream processing middleware that the measurements of some sensors are out of the expected ranges. Inspecting time series charts of recent measurements Marc realizes that some sensors are seriously broken and decides to go to the field next day by helicopter. In the field his team discovers that some of the wind sensors have been frozen and fix the problem. The description of the problem and time of replacement are immediately updated and fed back to SensorMap so that later models are correctly computed.

Data Analysis: After the campaign, since more stations have been placed in critical regions Marc can refine the resolution of his energy balance model. The dataset collected during the campaign is huge. To get an intuitive view of the data, through SensorMap in seconds Marc creates time series charts to understand correlations among different metrics, overlays contour maps on top of satellite images to see spatial interactions, and animates contour maps for different time periods to understand long-term trends. Later on, Marc identifies an interesting subset of the data, and simulates the models that are implemented in Matlab based on the data in his large workstation. Several hours later, he obtains visualizations of the energy flows from Matlab that can be again overlaid on top of terrain maps through SensorMap.

2. Challenges and Contributions

The design of SenseWeb aims to address the unique requirements and constraints posed by its constituent sensors, data, and target applications. Below, we summarize the major challenges.

Heterogeneity: Unlike the sensors comprising an application specific system, components of a shared sensing infrastructure may be highly heterogeneous along several dimensions. Sensors and, consequently, their output data are heterogeneous. The types of shared sensors may range from wireless motes, mobile phones, network cameras, pollution sensors, weather stations, to even RSS feeds.

Scalability: A shared system becomes more useful as the number of participants grows, creating the community effect. This introduces significant challenges for scalability. As the number of sensors and applications grows, the demand for resources increases. First, the large data volume continuously generated by shared sensors imposes challenges in data collection and storage. It is not efficient to transfer all data to a central storage system, which may then become a bottleneck for communication and processing. Second, in typical applications, raw data streams traverse through multiple layers of processing (e.g., data cleaning, gap filling, aggregation and transformation). It is not scalable to exhaustively pre-compute the multiple layers due to diverse application and user requirements. On the other hand, if data processing is always on demand, end users may experience intolerable response delays. Hence, a balance between pre-computation and on-demand computation is needed.

SenseWeb tackles these challenges by the following approaches:

1. An open and scalable back-end architecture that enables sharing of heterogeneous sensors is used in SenseWeb. Remote sensor gateways host sensor data streams, creating a federated Internet-scale sensing system. Moreover, it provides a tree-based type system, enabling customization of sophisticated sensor types.
2. Interactive geocentric data exploration is provided in a map-based frontend, providing spatio-temporal views of sensor streams, overlaid on a 2D or 3D map. For scalability, it caches computationally expensive visualizations and efficiently reuses the relevant portions based on overlapping regions among queries. Our techniques leverage spatio-temporal localities in data exploration to enable rapidly changing spatial visualizations of numerous streams.

Our current prototype (available at http://atom.research.microsoft.com/sensewebv3/sensormap) already incorporates a large heterogeneous collection of sensors including wireless motes, weather stations, traffic sensors, rain meters, and web cameras. Several groups of environmental scientists from different institutions including EPFL Switzerland, NTHU Taiwan, and NTU Singapore, have been using the prototype system to share their sensor deployments.

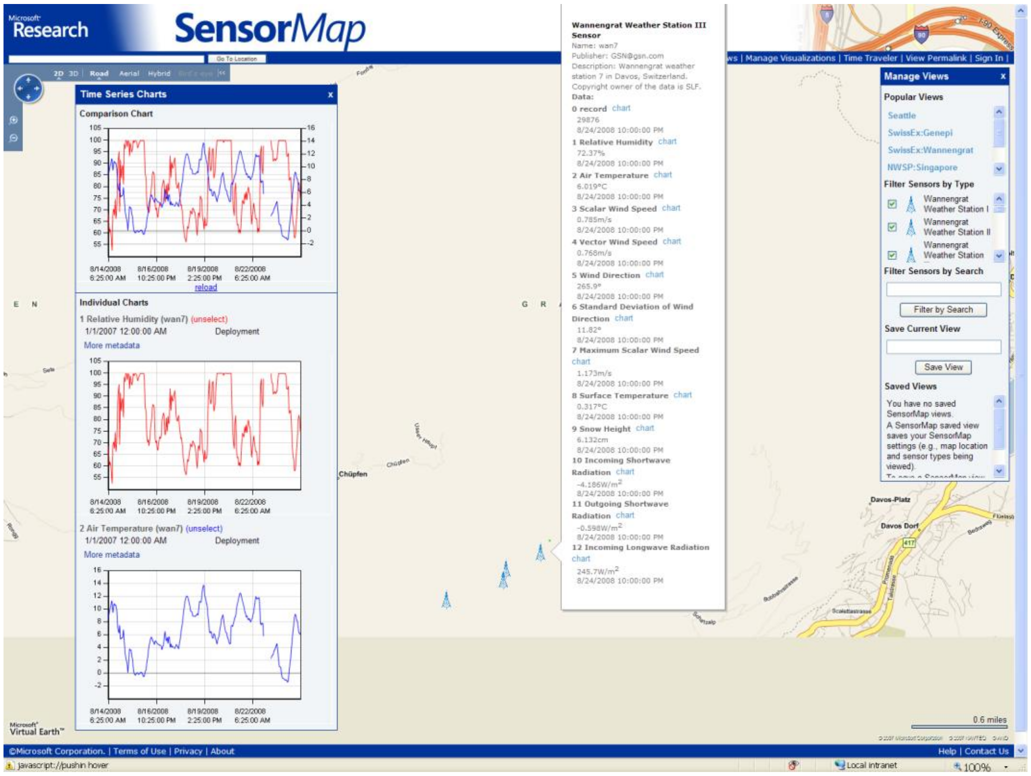

Figure 1 SensorMap presents most recent readings of a station as well as illustrates correlations between humidity (red) and temperature (blue) measurements.

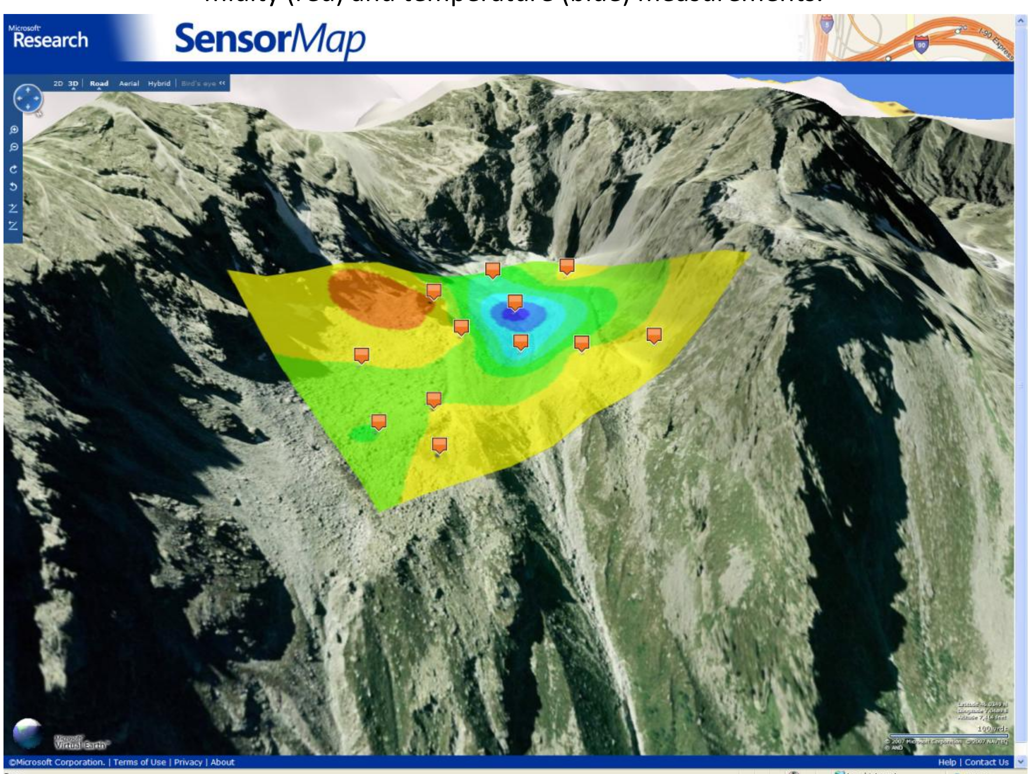

Figure 2 SensorMap visualizes spatial distribution of ambient temperature over 3D terrain maps.