SensorBase.org
A Centralized Repository to Slog Sensor Network Data

Kevin Chang, Nathan Yau, Mark Hansen, Deborah Estrin
University of California, Los Angeles
Center for Embedded Network Sensing
Department of Computer Science
Department of Statistics
Outline

• SensorBase.org overview
• User interface and implementation
• Applications of SensorBase.org repository data
• Future of SensorBase.org
SensorBase.org overview

• Similar to blogging sites, SensorBase.org is a *slogging* site that allows users to log sensor network data.

• Blog user interfaces (e.g. Blogger, WordPress) are very user-friendly; users can publish, delete, and set permissions on entries with little effort. We wish to create a similar interface and backend for slogging.

• Blog entries can be categorized and RSS is generated on the fly for easy notification and syndication.

• This is much the aim of SensorBase.org-- to be a user-friendly slogging tool for publication of data that makes it easy to share, collaborate, and notify via a variant of RSS, Numerical RSS (NRSS, nrss.org).
User interface and implementation

• To publish data, a user first creates a sensor network project which is a bit like categories in blogs.

• Data can be uploaded via XML and published to a project.

• Published data can be searched for with the SensorBase search engine; the search engine accepts queries that are a cross between natural language and SQL.

• Data can be retrieved as easy-to-parse XML or as a user-favorite, comma-delimited text.
Create a new sensor network project

creating a new proj -> create a new table -> create key(index) -> create relationships -> publish data

Short project name (to be used in search):

Project Description:

USPS address:

North/west bound coordinate ("top left" in lat,long or y,x):

South/east bound coordinate ("bottom right" in lat,long or y,x):

Min/max true MSL altitude in meters: _______ to _______

Data read permission for any user: ○ Allow ○ Deny
Data write permission for any user: ○ Allow ○ Deny
Data read permission for all users in the group: ○ Allow ○ Deny
Data write permission for all users in the group: ○ Allow ○ Deny

Submit
Publish sensor data using XML

1. Local Data File: Choose File no file selected
2. Overwrite existing data with the same primary key: Yes No, skip
3. Output verbose warnings: Yes No
4. Return status format: Text HTML
5. Publish data

(click here to upload binary directly)

Sample data file format:

```xml
<sensorbase_xml>
  <project project_id="5">
    <!-- data datetime="fufu" location_id="203" -->
    <data datetime="2003-01-01" location_id="202" celsius></celsius></data>

    <data datetime="1136945702" location_id="202">
      <relative_humidity>17</relative_humidity>
      <raw_humidity>38254</raw_humidity>
    </data>
    <data datetime="1136945702" location_id="202"></celsius>18</celsius></data>

    <data datetime="1136945702" location_id="202">
      <raw_voltage>2323</raw_voltage></data>

    <!-- unparsed human readable comments here -->
    <location location_id="3">x=33.232<y>-118.2325</location>
    <location location_id="203">x=33.235<y>-118.2388</location>
  </project>
</sensorbase_xml>
```
<table>
<thead>
<tr>
<th>Project Name (project_id)</th>
<th>Details</th>
</tr>
</thead>
</table>
| **the greatest project ever** (8) | Project tables (2): data, location  
Unique attribute types (12): celsius, datetimo, description, elevation, humidity, location_id, raw_humidity, raw_temperature, raw_voltage, voltage, x, y  
Address: 1603 Granville ave, 90025 |
| **second greatest project** (9) | Project tables (4): another_test, data, location, testing  
Unique attribute types (15): blah, celsius, datetimo, description, elevation, humidity, id, location_id, raw_humidity, raw_temperature, raw_voltage, voltage, whatever, x, y  
Address: ucla, 90095 |
the greatest project ever (8)

this is the absolute greatest project ever created

[Tags]

This project is centered at 1603 Granville ave, 90025.

Project members can read and can write project data.
Those who are not project members can read but cannot write project data.

<table>
<thead>
<tr>
<th>Table name</th>
<th>Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>data</strong></td>
<td><strong>datetime</strong> (datetime) This is the <em>measurement</em> date and time <strong>location_id</strong> (int) This <em>links</em> back to location.location_id</td>
</tr>
<tr>
<td></td>
<td><strong>raw_temperature</strong> (int) Raw temperature</td>
</tr>
<tr>
<td></td>
<td><strong>celsius</strong> (float) Actual celsius value</td>
</tr>
<tr>
<td></td>
<td><strong>raw_humidity</strong> (int) Raw uncalibrated humidity</td>
</tr>
<tr>
<td></td>
<td><strong>humidity</strong> (float) Calibrated humidity</td>
</tr>
<tr>
<td></td>
<td><strong>raw_voltage</strong> (int) Raw uncalibrated voltage</td>
</tr>
<tr>
<td></td>
<td><strong>voltage</strong> (float) Calibrated voltage</td>
</tr>
<tr>
<td></td>
<td>PRIMARY KEY (<code>datetime</code>, <code>location_id</code>)</td>
</tr>
<tr>
<td><strong>location</strong></td>
<td><strong>location_id</strong> (int) Unique identifier of the location</td>
</tr>
<tr>
<td></td>
<td><strong>description</strong> (text) my apartment</td>
</tr>
<tr>
<td></td>
<td><strong>x</strong> (float) GPS x</td>
</tr>
<tr>
<td></td>
<td><strong>y</strong> (float) GPS y</td>
</tr>
<tr>
<td></td>
<td><strong>elevation</strong> (float) Elevation in meters</td>
</tr>
<tr>
<td></td>
<td>PRIMARY KEY (<code>location_id</code>)</td>
</tr>
</tbody>
</table>
SensorBase.org
User interface and implementation

- Not getting into too many details, SensorBase.org is powered by PHP and a MySQL database.

- A new table(s) is created in the database each time a new sensor network project is created.

- User-defined tables removes the problem of attempting to create a universal data table that could hold all types of data.

- “Natural language” queries are parsed using regular expressions and corresponding SQL statements are sent to the database to find data sets.
<fieldList>
  <field><name>datetime</name><type>datetime</type></field>
  <field><name>user_id</name><type>int</type></field>
  <field><name>location_id</name><type>int</type></field>
  <field><name>celsius</name><type>real</type></field>
  <field><name>raw_temperature</name><type>int</type></field>
  <field><name>raw_humidity</name><type>int</type></field>
  <field><name>relative_humidity</name><type>real</type></field>
  <field><name>raw_voltage</name><type>int</type></field>
  <field><name>voltage</name><type>real</type></field>
</fieldList>

<result>2006-06-16 16:30:55,1000,2006-05-10 00:01:13,78,2629</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:02:49,68,2445</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:03:40,9,2590</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:03:43,158,258</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:03:54,10,2572</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:06:14,78,2630</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:08:40,9,2595</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:08:43,158,258</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:08:54,10,2576</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:11:14,78,2635</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:13:40,9,2591</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:13:43,158,258</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:13:54,10,2573</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:16:14,78,2636</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:18:41,9,2595</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:18:43,158,259</result>
<result>2006-06-16 16:30:55,1000,2006-05-10 00:21:14,78,2636</result>
Applications of repository data

- Google Earth
- Mapping data with ArcGIS
- Images as data
- Analysis and visualization with R
SensorBase.org
Google Earth and Roboducks

- Roboduck (robotic air boat) networks were deployed with a goal to track brown-tide algal blooms and follow their migration over time.

- Additional features which govern their abundance and survival in an area include temperature, nutrient concentrations, etc. which were monitored over time (Gaurav Sukhatme, USC).
Mapping data with ArcGIS

- ArcMap is a powerful mapping tool by the Environmental Systems Research Institute (ESRI) with 3D, spatial, and tracking analysis extensions.
<table>
<thead>
<tr>
<th>Average celsius/relative humidity/voltage from JR CAD</th>
<th>22.15°C</th>
<th>21.77°C</th>
<th>15.47°C</th>
<th>8.94°C</th>
<th>7.22°C</th>
<th>6.37°C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>34.31%</td>
<td>30.68%</td>
<td>66.18%</td>
<td>74.83%</td>
<td>72.04%</td>
<td>67.63%</td>
</tr>
<tr>
<td></td>
<td>2.83v</td>
<td>2.81v</td>
<td>2.78v</td>
<td>2.78v</td>
<td>2.79v</td>
<td>2.78v</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>location_id = 2</th>
<th>2006-06-14 12:04:00</th>
<th>2006-06-14 15:00:00</th>
<th>2006-06-14 18:12:00</th>
<th>2006-06-14 21:08:00</th>
<th>2006-06-15 00:04:00</th>
<th>2006-06-15 03:00:00</th>
</tr>
</thead>
<tbody>
<tr>
<td>location_id = 4</td>
<td>2006-06-14 12:04:00</td>
<td>2006-06-14 15:00:00</td>
<td>2006-06-14 18:13:00</td>
<td>2006-06-14 21:09:00</td>
<td>2006-06-15 00:05:00</td>
<td>2006-06-15 03:01:00</td>
</tr>
<tr>
<td>location_id = 10</td>
<td>2006-06-14 12:14:00</td>
<td>2006-06-14 15:16:00</td>
<td>2006-06-14 18:01:00</td>
<td>2006-06-14 21:02:00</td>
<td>2006-06-15 00:04:00</td>
<td>2006-06-15 03:05:00</td>
</tr>
</tbody>
</table>
Analysis and visualization with R

- R is a language and environment for statistical computing and graphics. It is a GNU project which is similar to the S language which was developed at Bell Labs by John Chambers (R-project.org).
Future of SensorBase.org

• Rule-based NRSS notifications using the R language.

• Refined query-handling for a more powerful data search engine.

• Applications built on top of SensorBase.org using such tools as Google Maps, ArcIMS.

• More extensive and useful visualization of environmental data.
Thank you.