Bits and Pieces of Lessons Learned on Making Sensor Data Available to The World

Jie Liu
Microsoft Research
Sensor Networks: 5 Years Ago
Sensor Networks: 2 Years Ago

- Internet
- Front end/gateway to internet
- Micro servers
- Sensors
One Candidate Consumer

Meta Info
GeoDBs (indexing & discovery)
Sensor lookup
Customized views
Aggregator (IconD)
DataHub
Live data
User query

[Diagram with various components and connections: Meta Info, GeoDBs, Sensor lookup, Customized views, Aggregator (IconD), DataHub, User query, Live data]
What’s between the clouds?

Sensor Net

Internet

web server

Business Logic

DB
We are not the only kids...

- We have to rely on Web infrastructure
- We can rely on Web infrastructure
  - Better programming support
  - Large pool of developers

- **Web Services**
  - Where to put the services?
    - On publisher sites
    - On portal/indexing sites
  - What to put in the API?
    - Sensor registration (metadata)
    - Real-time sensor data
SensorMap DataHub Service

DataHub

The following operations are supported. For a formal definition, please review the Service Description.

- **GetLatestBinarySensorData**
  Returns the latest image data reported by a sensor

- **GetLatestScalarSensorData**
  Returns the latest scalar data reported by a sensor

- **GetLatestSensingTime**
  Gets the latest timestamp of a sensor data

- **IsDataYoungerThan**
  Returns true if the sensor data is younger than specified milliseconds

- **RefreshSensorRegistration**
  Refresh a sensor registration.

- **RegisterSensor**
  Register a sensor.

- **RemoveSensor**
  Removes a sensor.

- **Sanity**
  Sanity check

- **SendBinaryData**
  Send binary sensor data such as images, sound or video. Data are treated as a binary file. Time parameter is the time stamp of the first data.

- **SendScalarData**
  Archive scalar sensor data. Array data are encoded as comma separated values. Time parameter is the time stamp of the first data.

- **SendXmlData**
  Send a XML document representing the data. The time stamp is considered as when the data is generate.

- **SensorList**
  Gets the list of sensors
<?xml version="1.0" encoding="utf-8"?>
    <Sensor xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
       xmlns:xsd="http://www.w3.org/2001/XMLSchema">
        <publisher>com.microsoft.research</publisher>
        <sensorName>MSR_Thermometer_112_3333</sensorName>
        <longitude>-122.14111761865615</longitude>
        <latitude>47.640932125477129</latitude>
        <altitude>0</altitude>
        <unit>Celsius</unit>
        <sensorType>
            http://research.microsoft.com/nec/sensor/type/SensorType.owl#Thermometer
        </sensorType>
        <address />
        <url>http://atom.research.microsoft.com/sensordatahub/datahub.asmx</url>
        <shortDesc>Temperature on Microsoft Campus 112/3333</shortDesc>
        <description>Temperature in Bldg 112/Room 3333. This is an indoor test sensor.</description>
        <dataType>double</dataType>
        <samplingPeriod>1</samplingPeriod>
        <reportPeriod>1</reportPeriod>
        <entryTime>2006-06-02T11:01:21.5572096-07:00</entryTime>
    </Sensor>
How are data published now?

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
<th>Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web camera</td>
<td>DOT traffic cam</td>
<td>URL</td>
</tr>
<tr>
<td>“Generic sensor”</td>
<td>JHU soil sensors</td>
<td>URL</td>
</tr>
<tr>
<td>Scalar data</td>
<td>MSR in building sensors</td>
<td>MSRSense-DataHub</td>
</tr>
<tr>
<td>Domain specific data</td>
<td>Weather, traffic</td>
<td>Web agent-DataHub</td>
</tr>
<tr>
<td>Domain specific</td>
<td>San Francisco Parking</td>
<td>Data query agent-DataHub</td>
</tr>
</tbody>
</table>
What Are Sensors After All?

<table>
<thead>
<tr>
<th>Continuously available</th>
<th>Sporadic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low semantic content, generic</td>
<td>High semantic content, specific</td>
</tr>
<tr>
<td>Machine readable</td>
<td>Human understandable</td>
</tr>
<tr>
<td>Data have to be cleaned, filtered, processed, screened before making available.</td>
<td>Already processed, edited, and distilled.</td>
</tr>
</tbody>
</table>

Why not support RSS as a sensor observation/notification format?
Technical Challenges

• Metadata and ontology
  – Corporate boundary and domain-specific formats
  – Standardization vs de-facto protocol/API

• User Management
  – Authorization, Authentication, Accounting
  – User information collection, privacy and auditing

• Sensor Management and Scalability
  – Naming and registration
  – Streaming and/or archiving
  – Event notification

• Active sensing and control
Non-Technical Challenges

- Data Ownership
SensorMap Data Publisher’s Agreements

- **You** have control over **Your Stuff**
  - **We** don’t claim ownership
  - Your Stuff is legal
  - You publish as you wish

- **You understand SensorMap is a research project**
  - We may reject/ remove Your Stuff
  - No commercial use
  - No liability

- **No Endorsement/ No Publicity**