

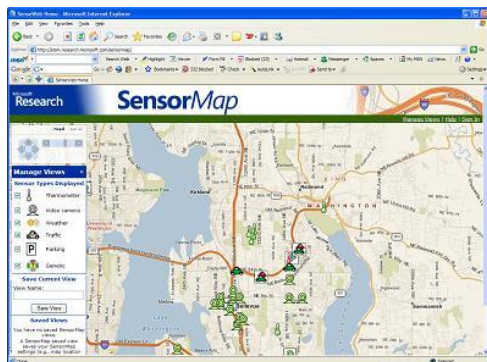
SenseWeb: a Shared Infrastructure for Internet-scale Sensing

Aman Kansal

(joint work with: Jie Liu, Suman Nath, and Feng Zhao)

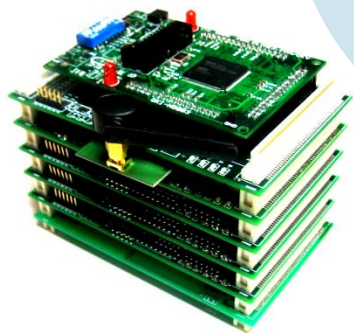
Networked Embedded Computing

Microsoft Research

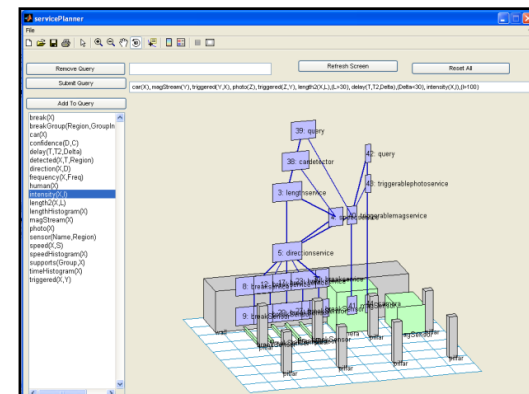
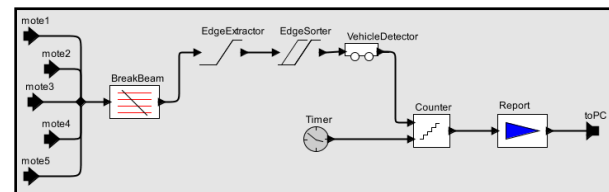


SenseWeb
Shared sensor
networks

mPlatform
reconfig
proc., comm,
storage



MSRSense
Composing
apps on
sensor
streams



Peer Produced Sensor Networks

- Multiple sensor networks shared to yield a single large system
 - Individual systems retain their identity
 - Resources shared for mutual advantage
- Peer production found useful in other domains



WIKIPEDIA



SecondLife
YouTube

...

Example: Rainstorm Management



Govt. deployed stream
Sensors
(Eg. USGS)



Road flooding aware
cab dispatch



Geo-stamped
flooded road image
from cellphones
(Eg. SensorPlanet)



City road route
management and
repairs



Scientist deployed
soil moisture sensor
network
(Eg. LifeUnderYourFeet)



Tourist day planner



Landscape
maintenance
scheduling
and inventory



Data from
home weather stations
(Eg. Wunderground)



Mountain bike router

Applications of Shared SN's

Community Fitness and Recreation

- Runners: Where are sidewalks broken? Construction finished on 24th St?
- Mountain Bikers: Average biker heart rate at Adams Pass on trail 320? [SlamXR]
- Surfer: What is the wave level and wind speed at Venice Beach now?

Real Time Information

- Public initiated instant news coverage

Science

- Continent scale phenomenon study using sensors deployed by multiple labs

Business

- What are people doing tonight? Restaurant waiting times in midtown?
- Mall visitor activity and parking usage across franchise outlets worldwide

Advantages

Standalone deployments

- Large scale deployment is challenging: cost, access rights, network/power infrastructure, maintenance overhead
- Sensors and data underused, replicated

SenseWeb: shared sensor networks

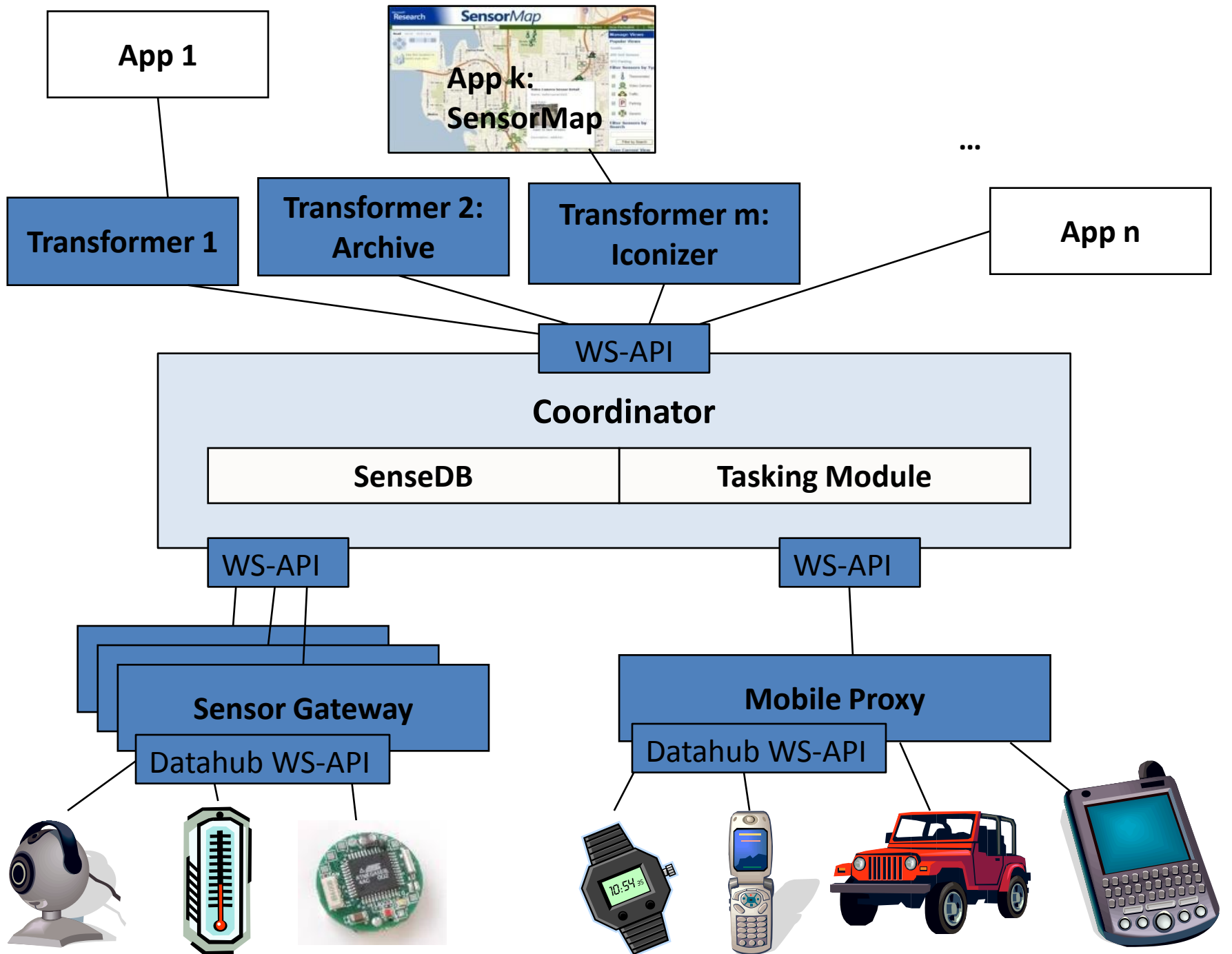
- Many sensor owners deploy at small scale
- Shared system yields large coverage
- Common sensors used by multiple apps

Challenges

- Heterogeneity
 - Resource capability: bandwidth, power, computation
 - Willingness to share
 - Measurement accuracy
- Scalability
 - Streaming all raw data from all sensors to all applications not feasible
- Security and Privacy
- Data Verifiability, Trust

SenseWeb

- A shared software infrastructure for sharing multiple sensor networks among multiple concurrent sensing applications
- Components
 - Tools for sensor contributors to share data
 - Interfaces for apps to collect relevant data
 - Efficient sharing of data and sensing tasks
 - SensorMap: a GUI to visualize available data



Coordinator



Accepts
application
sensing
demands

Determines
sensing tasks
and data
overlap

Distributes
sensing tasks to
selected
sensors

Data Re-use

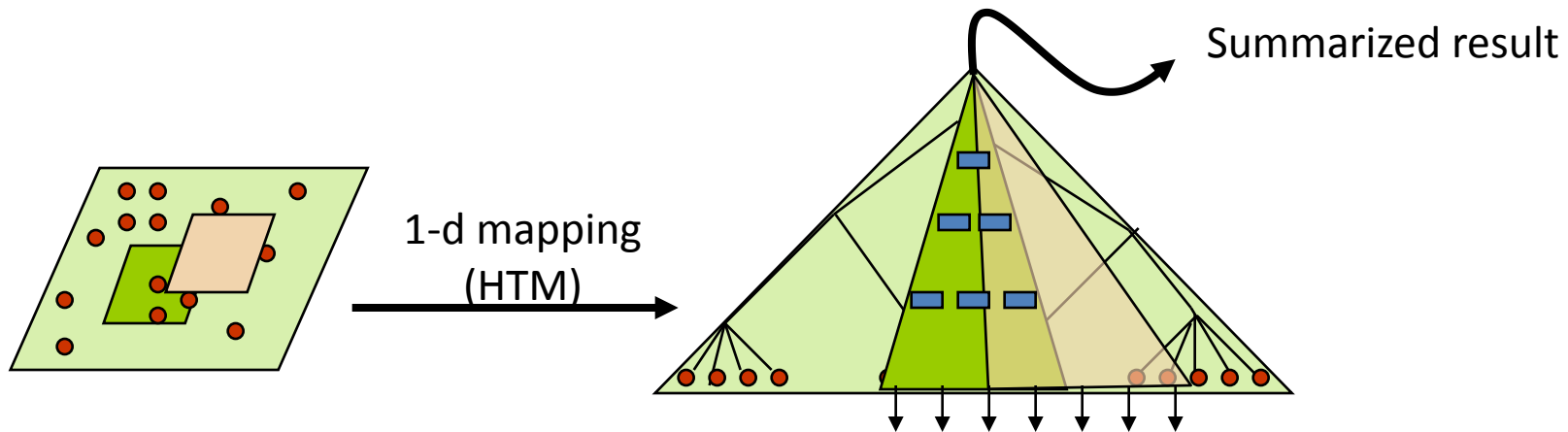
- Many applications may need similar data
 - Within a tolerable latency of each other
 - From overlapping region
- Can cache data and aggregates to reduce load on sensors and network
 - Overlap may be partial: computed aggregates may need partial new data

Query Model

```
SELECT Count(*)  
FROM Sensor  
WHERE sensor.location in Polygon(A) AND  
      sensor.time BETWEEN now()-10 and now()+10  
REPORTRATE 10 min  
SAMPLESIZE 50  
EVENT EventSpec(T>25)
```

COLR-Tree (COLlection R-Tree)

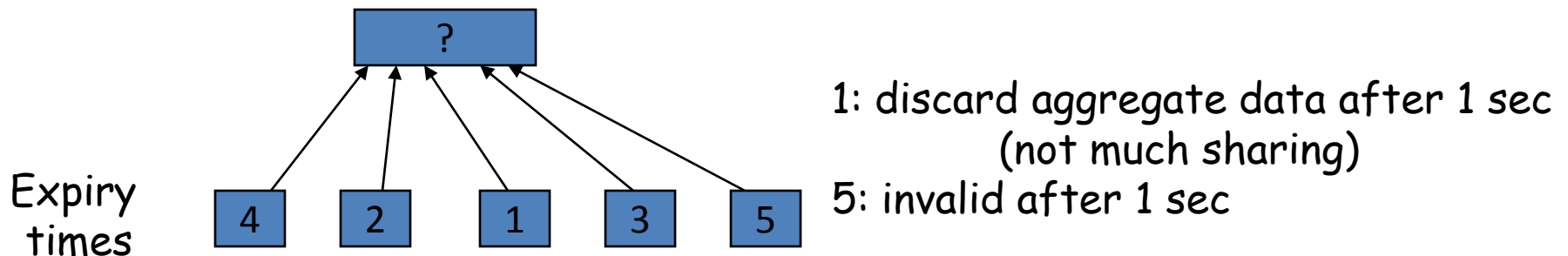
Index 2-D data with aggregates



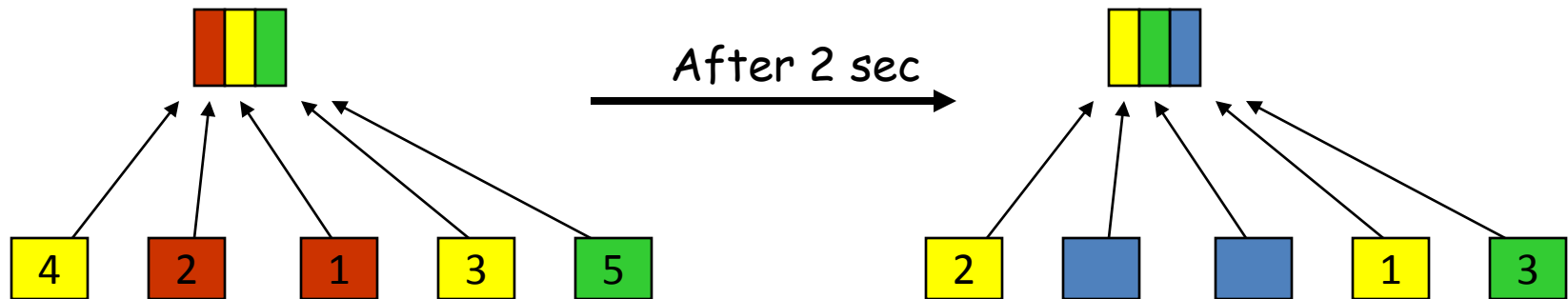
- Minimizing sensor access
 - Cached data may have skewed distribution
 - Sample more from non-cached region
- Implemented on MS-SQL Server: usable with all SQL server capabilities

COLR-Tree: Aggregates

- Challenge: temporal aggregation



Solution: slotted aggregation



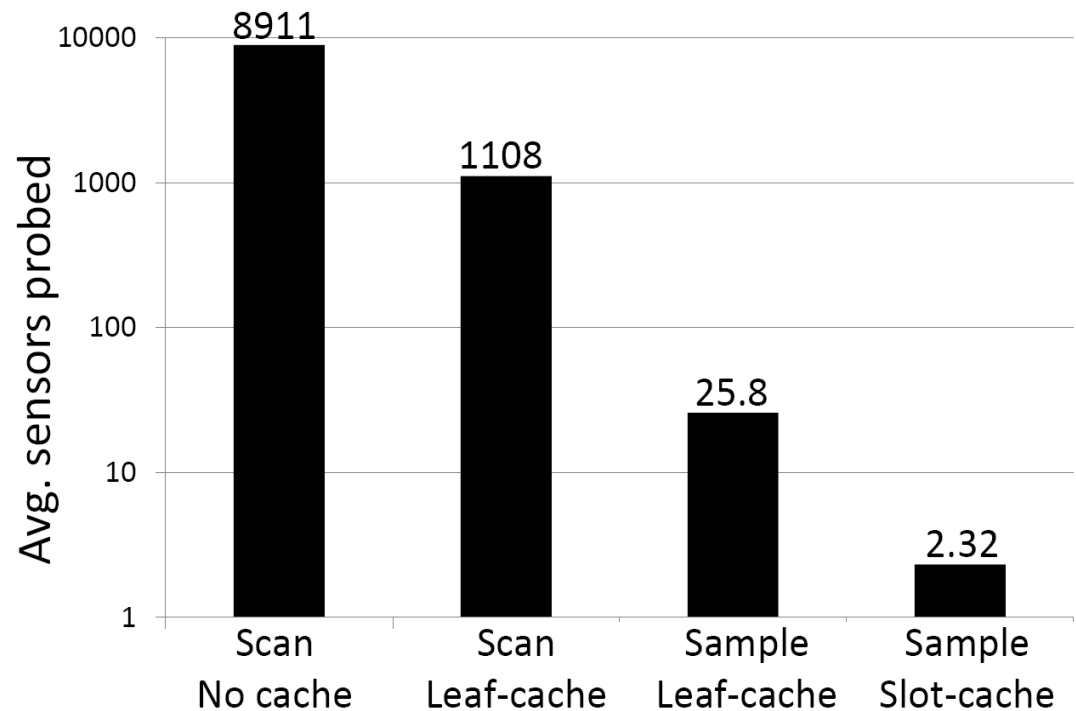
Spatial Sub-sampling

- Suppose sample size of R needed
- Layered sub-sampling along COLR-tree levels
- Partition R to achieve spatially uniform sample
 - $BB(i)$: area covered by i -th child, $c(i)$: data cached for i -th child, $w(i)$: sensors under i -th child, q : query region
 - For each child I at next level:

$$R(i) = R * \frac{w(i) - c(i)}{\sum_i w(i) * |overlap(BB(i), q)|}$$

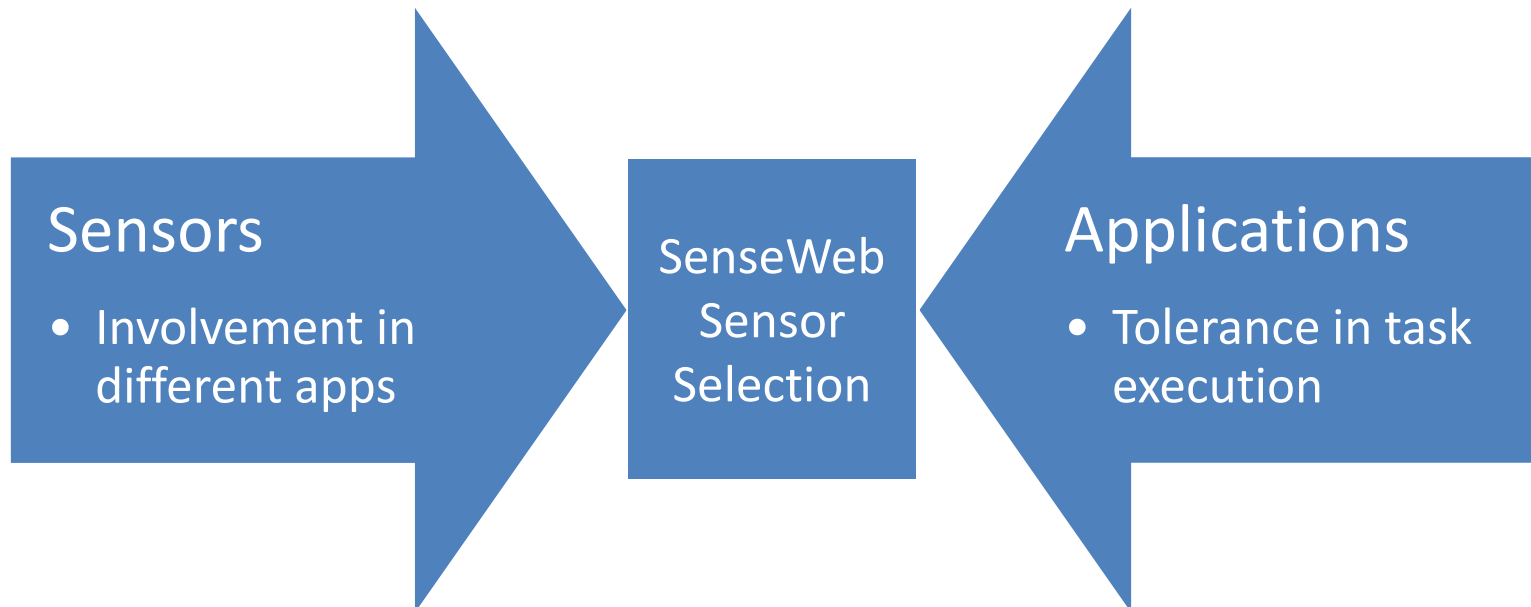
COLR-Tree Evaluation

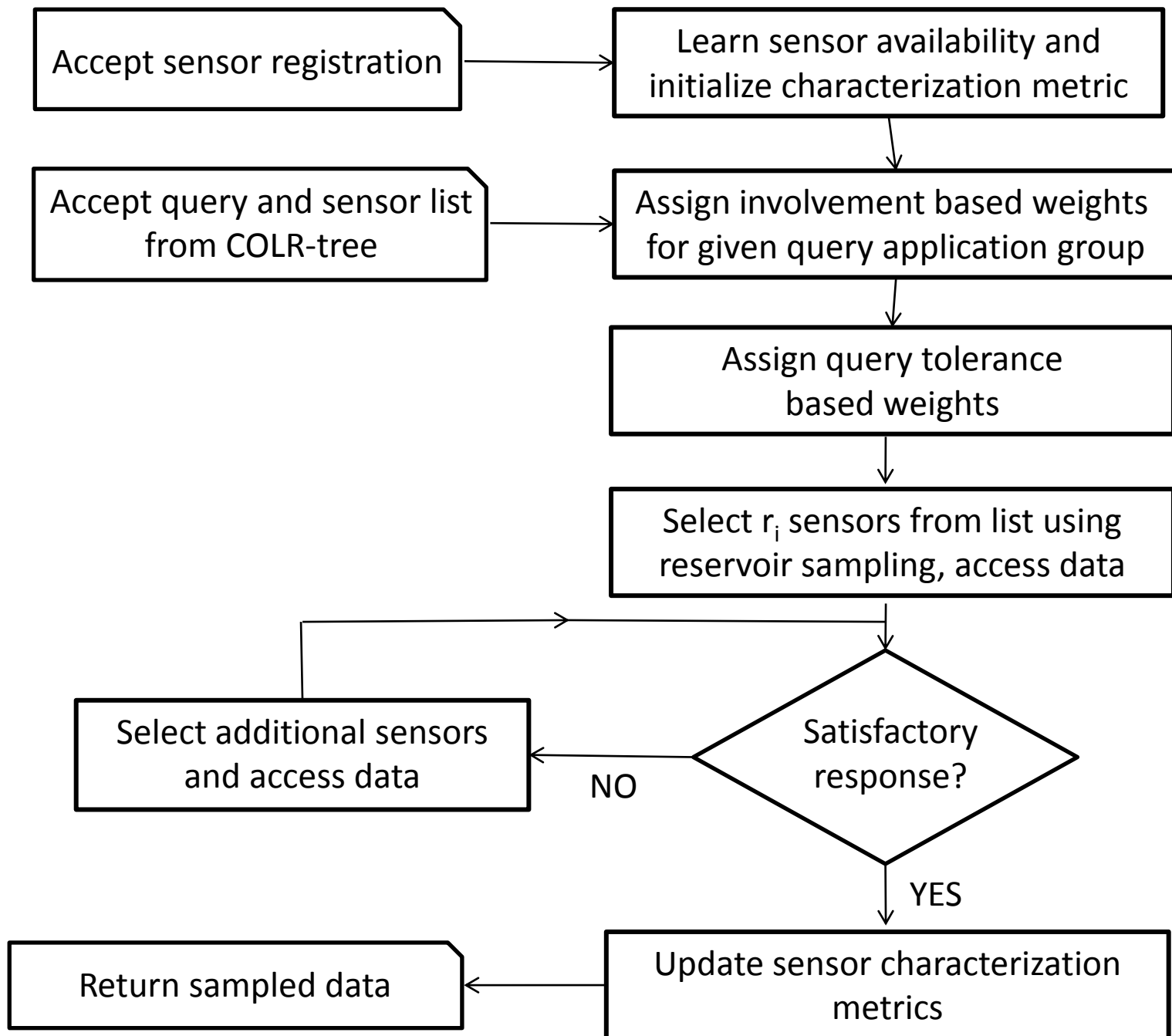
- Test data
 - 400K points from VE Yellow Pages
 - Regions queried: Virtual Earth usage trace



Tasking Heterogeneous Sensors

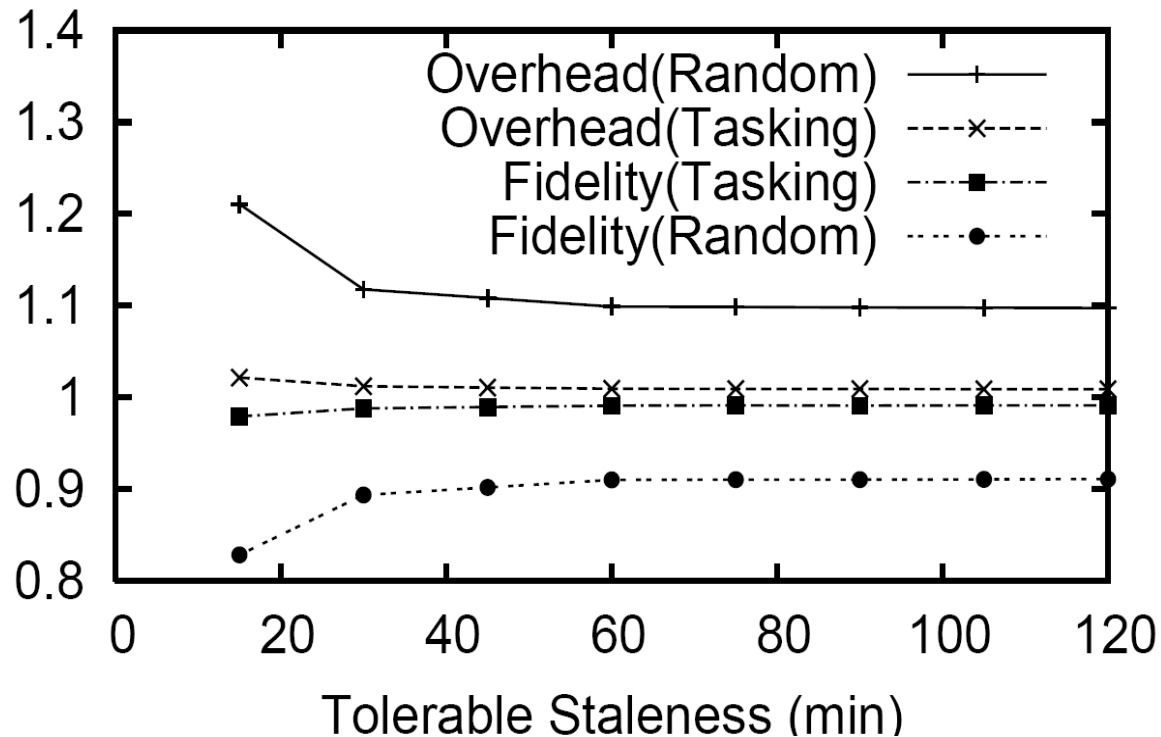
- Select uniformly rather than overloading the best sensors
- Leverage lower capability sensors when usable for a query
- Learn and adapt to sensor characteristics: availability, bandwidth
- Weighted reservoir sampling
 - Weighted random selection, with desired number of sensors





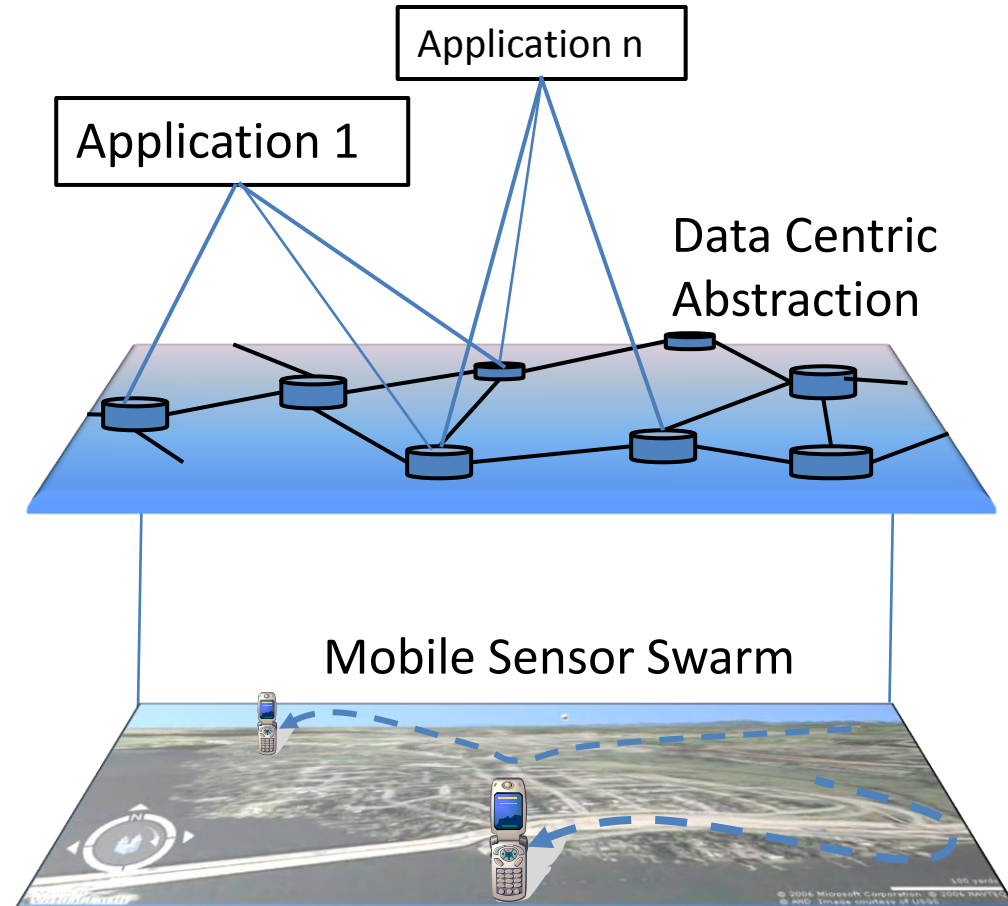
Tasking Algorithm Performance

- Test on USGS stream water sensors
 - Random selection vs. Weighted reservoir sampling

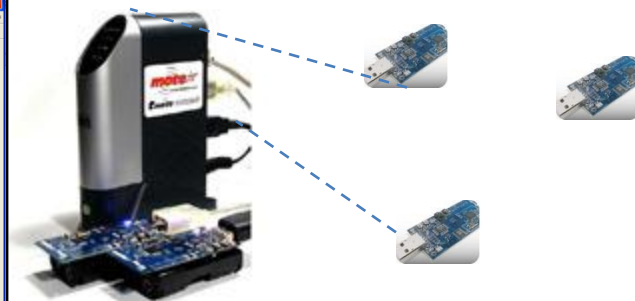
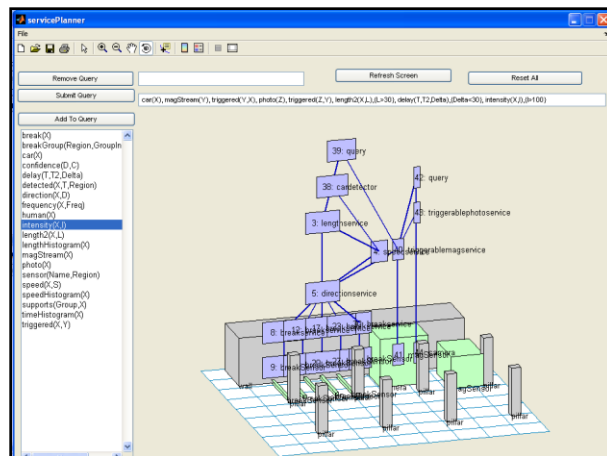


Mobile Contributors in SenseWeb

- + More coverage
- Hard for application to track relevant devices
- **Solution:** data centric abstraction
 - Location based indexing
 - using GPS, cell-tower triangulation, content based location



Tools for Sensor Contributors

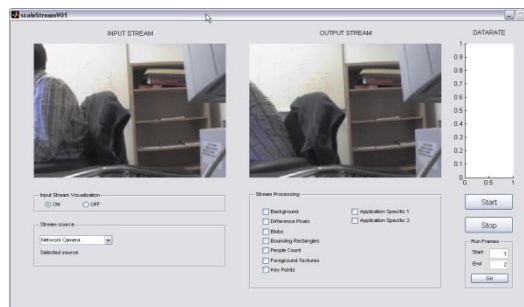


For mote networks

- Automatic data collection and sharing
- Simplified processing and application composition

Client for cell-phones

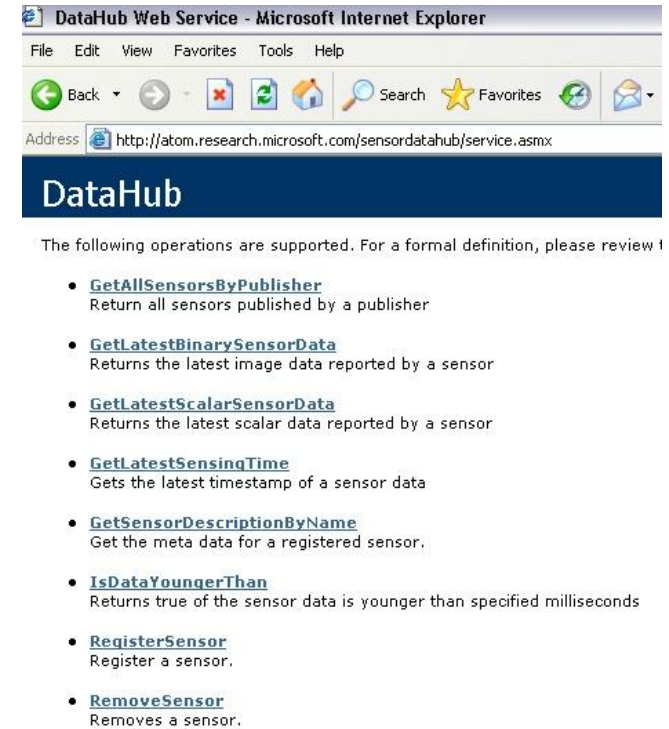
- Allows users to take pictures
- Automatically uploads data to server
- Location stamps using inbuilt/Bluetooth GPS



Webcam data processing and sharing tool

Tools for Sensor Contributors

- Gateway for sensor contributors
 - Web service API: Datahub
 - Supports several sensor types via semantic hierarchy
 - Also archives sensor data
- Tools available for download
 - Tutorials available online



The screenshot shows a Microsoft Internet Explorer browser window titled "DataHub Web Service - Microsoft Internet Explorer". The address bar displays the URL "http://atom.research.microsoft.com/sensordatahub/service.asmx". Below the address bar, there is a dark blue header with the text "DataHub". The main content area lists supported operations with their descriptions:

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

- [GetAllSensorsByPublisher](#)
Return all sensors published by a publisher
- [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
- [GetSensorDescriptionByName](#)
Get the meta data for a registered sensor.
- [IsDataYoungerThan](#)
Returns true if the sensor data is younger than specified milliseconds
- [RegisterSensor](#)
Register a sensor.
- [RemoveSensor](#)
Removes a sensor.

Application Access to Sensors

- Through web based GUI
 - <http://atom.research.microsoft.com/sensormap/>



- Through web services
 - Public access coming soon!

Summary

- Peer production can reduce the resource costs of sensor networks for many apps.
- SenseWeb: shared system to work with
 - Highly heterogeneous sensing resources
 - Dynamically varying availability
 - Multiple concurrent applications

See Your Sensor on SenseWeb

- SenseWeb is online
 - 69 unique users, from 10 countries, 2900 sensors
 - Eg. ETH, EPFL, Johns Hopkins, Ohio State, MIT, ...
- Share your sensor networks
- <http://research.microsoft.com/nec/senseweb>
 - See Overview, Tutorial, Forum