Lowering the Barriers to Large-Scale Mobile Crowdsensing

Presenter:
Sandeep
Department of Computer Science and Engineering
University of Minnesota, Twin Cities
Outline

- Example Case
- Introduction and Motivation
- Barriers
- System Architecture
- Workflow
- Benefits and Challenges
Jake, Mom and Dad visit a crowded parade

Curious Jake gets lost in the crowd

The Mom starts to panic

She is furious over the Dad who chooses to play with his smartphone rather than finding Jake
Is the Dad Crazy?

- No!!! He is a Computer Scientist
- He starts an application on his smartphone
- Sends the image of the child and current location to someone
- Receives a message that has a image of his child and his potential location.
No!!! Crowdsensing

The Dad starts a search with his Crowdsensing application by uploading the image of the child.

The application matches the child’s image with the many videos and pictures taken by various users at the given location.

All potential matches are forwarded to the Dad with the location.
Introduction and Motivation

Crowd
(Smartphone Users)

Sensing
(Sensors)

Camera/Mic.
Compass/Gyro
GPS/WPS
Proximity
Accelerometer
...
Crowdsensing: Individuals with sensing and computing devices collectively sharing information to measure and map phenomena of common interest.

Rich information about the smartphone user's activity and environment

Growth in richness and diversity of sensors on smartphones shifted focus from individual sensing to crowd sensing
Offline applications of Crowdsensing includes analysis of transportation activities in urban spaces, measurement of inter person similarity, health assessment of elder people.

Real-Time applications include traffic monitoring, public safety, collaborative searching.
Sample of Crowdsensing studies

- Crowd-size is often removed while reporting studies

- Numbers are generally small

- Data sources that are easy to collect allow larger crowds to be studied. (Twitter etc.)

---

<table>
<thead>
<tr>
<th>Reference</th>
<th>Mobile Platform</th>
<th>Application Category</th>
<th>Crowd Size</th>
<th>Input</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhou et al. [32]</td>
<td>Android</td>
<td>Transportation</td>
<td>unknown</td>
<td>cell tower ID, audio signal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>accelerometer</td>
</tr>
<tr>
<td>Tiramisu [33]</td>
<td>iOS</td>
<td>Transportation</td>
<td>28</td>
<td>GPS</td>
</tr>
<tr>
<td>SignalGuru [10]</td>
<td>iOS</td>
<td>Transportation</td>
<td>13</td>
<td>video frames</td>
</tr>
<tr>
<td>Balan et al. [2]</td>
<td>Car GPS</td>
<td>Transportation</td>
<td>15000</td>
<td>GPS</td>
</tr>
<tr>
<td>Mathur et al. [14]</td>
<td>Car GPS</td>
<td>Transportation</td>
<td>500</td>
<td>GPS</td>
</tr>
<tr>
<td>Niu et al. [17]</td>
<td>Blackberry</td>
<td>Transportation</td>
<td>unknown</td>
<td>GPS</td>
</tr>
<tr>
<td>Bao et al. [3]</td>
<td>Symbian, iPod</td>
<td>Social Application</td>
<td>unknown</td>
<td>video</td>
</tr>
<tr>
<td>Wirz et al. [27]</td>
<td>Android</td>
<td>Social Application</td>
<td>unknown</td>
<td>GPS</td>
</tr>
<tr>
<td>CrowdSearcher [28]</td>
<td>iOS</td>
<td>Search</td>
<td>unknown</td>
<td>image</td>
</tr>
<tr>
<td>GeoLife [31]</td>
<td>GPS phones</td>
<td>User Behavior Study</td>
<td>107</td>
<td>GPS</td>
</tr>
<tr>
<td>SoundSense [13]</td>
<td>iOS</td>
<td>User Behavior Study</td>
<td>unknown</td>
<td>audio stream</td>
</tr>
<tr>
<td>#EpicPlay [25]</td>
<td>Twitter</td>
<td>Social Application</td>
<td>unknown</td>
<td>tweets</td>
</tr>
<tr>
<td>Wakamiya et al. [26]</td>
<td>Twitter</td>
<td>User Behavior Study</td>
<td>39898</td>
<td>tweets</td>
</tr>
<tr>
<td>Fujisaka et al. [7]</td>
<td>Twitter</td>
<td>User Behavior Study</td>
<td>8139</td>
<td>tweets</td>
</tr>
<tr>
<td>CrowdSearcher [4]</td>
<td>Facebook</td>
<td>Search</td>
<td>137</td>
<td>text</td>
</tr>
</tbody>
</table>

Table 1: Representative Sample of Crowd-sensing Applications
Heterogeneity of Mobile devices

- There are multiple platforms in the smartphone market that require dedicated applications

- Different versions of same platform raise compatibility issues

- Popularity of standalone applications discouraging the development of external libraries

- Application written in HTML5 doesn’t have access to native hardware
Burden placed on Users

- Users must install a separate application for every Crowdsensing study
- Can take weeks or months to reach critical masses
- Users must be tolerant of the processing, memory and battery
- Deployment of application dependent at the rate at which users adopt to install it.
Barriers (3)

- Increasing network bandwidth demands
  - Researches related to augmented reality and multimedia sensing demand a lot of network bandwidth
  - With many users an application can easily overwhelm link capacity in regional networks and into data centers
  - Applications where each device sends data to centralized servers cannot scale to support data-rich sensors
  - Ensuring scalability requires rethinking architectures
## Proposed Solution

<table>
<thead>
<tr>
<th>Barrier</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterogeneity of sensing hardware and mobile platforms.</td>
<td>Separation of data collection and sharing from application-specific logic</td>
</tr>
<tr>
<td>Burden, Crowdsensing applications place on users</td>
<td>Removal of app installation on smartphones from the critical path of deployment</td>
</tr>
<tr>
<td>Increasing network bandwidth demands</td>
<td>Decentralization of processing and data aggregation near the source of data</td>
</tr>
</tbody>
</table>
System Architecture

1\textsuperscript{st} Layer – Mobile devices
2\textsuperscript{nd} Layer – Distributed Cloud Infrastructure
3\textsuperscript{rd} Layer – Coordinating Application Server
Mobile devices forwards captured data to proxy VMs in the second layer

Each proxy VM is an extension of mobile phone into the cloud

Proxy VMs are kept close to the user through VM Migration

Data is transferred to one or more application VMs through Proxy VMs
Application VMs performs tasks related to each Crowdsensing application

Application VMs are deployed by the highest layer coordinating entity.

The application server can initiate a master VM when many application VMs are run on the cloudlet

MAVM coordinates, clones, configures application VM and aggregates data before forwarding results
1. Task Generator –
   - Constructs and sends a task description to the application server
   - Format of the task-description can be application specific
   - Critical Information like type of search, location, sample image

Example Task Description:
<location> Fifth Avenue, New York </location>
<time> 13:00-13:30 EST 30.10.2011 </time>
<action> face detection </action>
<output> GPS </output>
<attachments>
  <image> photo_1.jpg </image>
  <image> photo_2.jpg </image>
</attachments>
2. Parsing of Task Description

- Application Server parses the task description
- Consults global registry to get a list of cloudlets near the target area

Example Task Description:

```
<location> Fifth Avenue, New York </location>
<time> 13.00-13.30 EST 30.10.2011 </time>
<action> face detection </action>
<output> GPS </output>
<attachments> 
  <image> photo_1.png </image>
  <image> photo_2.png </image>
</attachments>
```
3. Creation of MAVM

- Application server contacts each cloudlet in the target area
- Requests creation of MAVM to the cloudlet daemon
Workflow

4. Discover proxy VMs

- MAVM uses cloudlet registry discover proxy VMs
- VMs connected to devices can provide desired sensor data
### Workflow

5. Connecting to actual device

- MAVM requests participation from the mobile users through proxy VM
- Proxy may require explicit permissions or can automatically join crowds on behalf of the user depending on the user-defined policies
6. Create Application VMs

- Once permission is granted, the MAVM will request the cloudlet daemon to create application VMs.
- In reality, they can simply be the clones of the MAVM operating in different mode.
7. Configuration

- MAVM configures the networking setup
- Proxy VM adds the application VM to the subscriber list
Workflow

- After the seven steps, the proxy VMs starts forwarding images and videos to application VM

- AVMs apply face detection and through MAVM forwards potential matches to application server and ultimately to parent’s phone
Benefits:

- Scalable, efficient data sharing between multiple applications or users and participation of multiple applications at a global scale.

- Separation of data collection and sharing from application specific processing makes it easier for developers to focus on the latter part.

- Rapid deployment of Crowdsensing applications as users no longer need to install individual apps.

- Closer to the data sources thus reduces traffic on wide area networks and reduces latency as well.
Challenges:

- Virtualization overheads
- Migration induced reconfiguration
- Standardization of sensing interfaces
Thank you 😊