Strategies for Crowdsourcing for Disaster Situation Information

Edward T.-H. Chu  Phd
Assistant Professor, CSIE department
National Yunlin University of Science and Technology, Taiwan

2nd International Conference on Disaster Management and Human Health: Reducing Risk, Improving Outcomes, Orlando, USA, 2011
Coauthors

- **Prof. J. W. S. Liu (IEEE Fellow)**
  - Academia Sinica, Taiwan
- **Prof. J. K. Zao**
  - National Chiao Tung University, Taiwan
- **Y.-L. Chen**
  - National Yunlin University of Science and Technology, Taiwan
YunTech

- **National Yunlin University of Science and Technology, Taiwan**
- **YunTech (1991~)**
  - One of enterprises' 10 favorite universities in 2011
  - 9,700 students (30% graduate, 70% undergraduate)
  - 334 faculties (86% phd degree, 80% industry experience)
Outline

- Introduction
- Crowd Model
- Simulation Results
- Summary and Future Work
Disaster Surveillance System

- **Functionalities of disaster surveillance system**
  - Collect real-time information
  - Estimate boundaries of threatened areas
  - Assess the threat potential

- **Sensor networks** play an important role in collecting data
  - Camera surveillance network
  - Wireless sensor networks
Limitation of Sensor Networks

- In a major disaster
  - Sensors can be easily damaged
  - Sensors may not be able to cover all areas

- People with wireless devices (human sensors):
  - Complement physical sensors
Crowdsourcing for Disaster Management

- **Mobile human sensors**
  - Use wireless devices (ex: smart phone) to collect data
  - Use social media (ex: facebook, Twitter) to report data

- **Crowdsourcing**
  - The act of outsourcing tasks to an undefined large group of people (a “crowd”) through an open call
Crowdsourcing for Disaster Management

- Crowdsourcing data collection process
  - Start: Broadcast a data collection request to a crowd
  - Stop: Collect enough data to construct a sufficient view of the threaten area

Japan Earthquake 2011

Hadi earthquake 2010
Crowdsourcing Strategy Types

- **Random**
  - After broadcasting a request, the system does nothing other than collecting the reports

- **Crowd-driven**
  - The system updates the collected information and the crowd guides themselves in their exploration

- **System-driven**
  - The system directs all individuals
Contributions

- A crowd model for characterizing each individual within a crowd
- A mobility model of crowd movements
- A general methodology for evaluating strategies for crowdsourcing sensor data collection
Outline

- Introduction
- Crowd Model
- Simulation Results
- Summary and Future Work
Crowd model

- **Graph model of threatened area**
  - Represent a neighborhood of missing sensors

- **Participant types & models**
  - Characterize the quality and speed of each individual participant

- **Mobility models**
  - Characterize the movement of a participant from sensor location to sensor location
Graph Model of Threatened Area

- **A directed graph** is used to characterize the threatened area

  - $S_i$: Missing sensor, $T_{i,j}$: The directed edge of $(S_i, S_j)$

Figure 1 Oil spill disaster scenario

Figure 2 Wildfire surveillance scenario
Participant Types

- The quality of collected data depends on crowd quality

- **I-type: ideal**
  - Police officers, fire fighters, and soldiers
  - Move to the right location promptly, and make a right observation

- **M-type: highly motivated**
  - Registered volunteers
  - The participant is known to the system

- **U-type: unknown**
  - May take longer time to respond to the request
Participant Models

- **Response Time:** $R_k(i, j) = \Delta_k + \Omega_k T_{i,j}$
  - $R_k(i, j)$: The amount of time required by the participant $k$ to travel from location $S_i$ to $S_j$
  - $\Delta_k$: The delay of the participant $k$
  - $\Omega_k$: The speed of the participant $k$

- **Sample Errors:** $\Theta_k = (\Theta_{k,1}, \Theta_{k,2}, ..., \Theta_{k,\eta})$
  - $\Theta_{k,i}$: the error of $i$th sample take by the participant $k$

<table>
<thead>
<tr>
<th>Type</th>
<th>$\Delta_k$ (Delay)</th>
<th>$\Omega_k$ (Speed)</th>
<th>$\Theta_k$ (Sample error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type-I</td>
<td>No delay</td>
<td>Fast</td>
<td>Zero</td>
</tr>
<tr>
<td>Type-M</td>
<td>Short delay</td>
<td>Mid</td>
<td>Small</td>
</tr>
<tr>
<td>Type-U</td>
<td>Long delay</td>
<td>Slow</td>
<td>Large</td>
</tr>
</tbody>
</table>
Mobility Models

- Characterize the movement of participant from sensor node to sensor node

- Random Walk (RM)
  - Choose an outgoing edge randomly
Mobility Models (cont.)

- Random Walk Forward-Only (RMFO)
  - Discard all visited edges and select one unvisited edge
Mobility Models (cont.)

- **Random-Least-Visited-First (RLVF)**
  - Visit the least visited link first

Tour: $S_1$ to $S_7$

RLVF: $\square$
Outline

- Introduction
- Crowd Model
- Simulation Results
- Summary and Future Work
Performance Index of Crowdsourcing

- **Response time**
  - How fast a crowdsourcing can be when compared to an official rescue action?

- **Spatial resolution**
  - How many locations can be visited by the crowd within a given time?

- **Is it possible to use crowdsourcing to eliminate official efforts on exploration?**
## Evaluation of the Crowd-driven Strategy

### Mobility models

<table>
<thead>
<tr>
<th>Type</th>
<th>Mobility models</th>
<th>Sample errors</th>
<th>Minimum time per sample taken</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type-I</td>
<td>RLVF (Rand Least Visited First)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Type-M</td>
<td>RLVF</td>
<td>[-20%,20%]</td>
<td>[1,2]</td>
</tr>
<tr>
<td>Type-U</td>
<td>RW (Random Walk)</td>
<td>[-25%,-25%]</td>
<td>[1,10]</td>
</tr>
</tbody>
</table>

**Figure 1** Oil spill disaster scenario

**Figure 2** Wildfire surveillance scenario
Response Time

- Response time: the time to visit all nodes
- The number of participants becomes large, the difference among crowds becomes small
- The response time does not improve much when the number of type-I participants increases

Figure 5 Oil spill disaster scenario
Figure 6 Wildfire surveillance scenario
Spatial Resolution

- $h$: the time required by a type-I participant to visit all locations
- **Spatial resolution**: the number of visited location to the number of total locations within $h$
- Result: It may not be necessary to use any I type participant

Figure 3 Oil spill disaster scenario  Figure 4 Wildfire surveillance scenario
Crowd Composition

- Mix different types of participants in an experiment: (type-I, type-M, type-U)
- A crowd with type-I participants always performs better than a crowd without type-I

<table>
<thead>
<tr>
<th>Table 1: Response time (in minutes) of different crowd model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Oil spill disaster</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>Wildfire surveillance</td>
</tr>
<tr>
<td>30</td>
</tr>
</tbody>
</table>
Outline

● Introduction
● Crowd Model
● Experimental Results
● Summary and Future Work
Summary

- Several models are proposed
  - represent different disaster scenarios and participants
- For a crowd without type-I, the number of type-M participants have a significant impact on the response time
- We may not need to use any type-I participant
- Crowd-driven crowdsourcing may be able to eliminate the efforts of official rescue
Future Work

- A comprehensive evaluation of the crowd-driven crowdsourcing strategy
- We are developing a system-driven strategy
  - 2011 International Conference on Environmental Emergency Response and Homeland Security, Taiwan
Thank you

http://sites.google.com/site/edwardchutw/