Dynamic Data Driven Crowd Sensing Task Assignment

Layla Pournajaf, Li Xiong, and Vaidy Sunderam

Department of Mathematics and Computer Science
Emory University

DDDAS Workshop, June 11, 2014
Crowd Sensing

- Sensing paradigm using mobile devices
- Syndromic surveillance, traffic monitoring, news coverage, intelligence data collection, crowd/congestion monitoring
Spatial Crowd Sensing Task Assignment

- Sensing tasks: collect data about specific targets at particular locations.
- Task assignment
- Maximize sensing coverage
- Minimize sensing cost
Challenges

- **Uncertain and dynamic trajectories of participants:**

  1) Error in location-detection devices, noisy transmissions, or explicit location perturbation due to privacy concerns

  2) Participants may be constantly moving (e.g. commuters) and

- **Dynamic tasks:** Sensing tasks may be updated (e.g. spread of disease outbreak, moving crowd in an event)
Contributions

● A dynamic data driven framework for spatial task assignment in mobile crowd sensing with dynamic and uncertain participant locations

● Utilize DDDAS to steer and assign the sensing tasks in targeted ways, adapting dynamically to

1) sensing/application needs
2) uncertain and moving locations of participants
Adaptive dynamic data driven framework
Adaptive dynamic data driven framework

Participants, based on their assigned tasks, report collected data to the application.
Adaptive dynamic data driven framework

The collected data are dynamically integrated into an executing simulation to augment or complement the application model, which updates the sensing targets for future data collection.
Adaptive dynamic data driven framework
Adaptive dynamic data driven framework

A mobility model built offline by mining public trajectories or historic trajectories.
Adaptive dynamic data driven framework

- Prediction: computes a participant's current location based on her historic trajectory and the mobility model (i.e. prior estimates).
- Correction: integrates the reported (uncertain) location and predicted location into a more accurate location (i.e. posterior estimate).
Bayesian Filtering

- State space model:
  \[ X_t \sim p(X_t | X_{t-1}) \]
  \[ Z_t \sim p(Z_t | X_t) \]

- Measurement sequence \( Z_{1:t} = Z_1, \ldots, Z_t \).

Prediction:
\[
p(X_t | Z_{1:t-1}) = \sum p(X_t | X_{t-1}) p(X_{t-1} | Z_{1:t-1})
\]

Update:
\[
p(X_t | Z_{1:t}) = \frac{p(Y_t | X_t) p(X_t | Z_{1:t-1})}{\sum p(Y_t | X_t) p(X_t | Z_{1:t-1})}
\]
Adaptive dynamic data driven framework

Given the posterior location estimate of participants, assign tasks to participants in real time while globally optimizing sensing coverage and cost.
Global Task Assignment

- Given a set of N participants with uncertain locations and sensing range, a set of M sensing targets, a desired task coverage goal, assign targets to qualified participants with minimum cost.

- Task Cost:
  \[ TC = \sum_{j \in M} \sum_{i \in N} x_{i,j} d_{i,j} \]

- Task Coverage:
  \[ TU = \sum_{j \in M} \frac{\sum_{i \in N} x_{i,j}}{k_j} \]
Global Task Assignment

\[
\min_x \sum_{i \in N} \sum_{j \in M} d_{i,j} x_{i,j}
\]

s.t.
\[
\sum_{j \in M} \frac{\sum_{i \in N} x_{i,j}}{k_j} \geq gm
\]
\[
\sum_{j \in M} x_{i,j} d_{i,j} \leq b_i
\]

- Matrix \(d\) is uncertain, hence distances should be estimated.
- NP-Hard Problem (Minimum Set Cover Problem reduction)
- Probabilistic greedy algorithms
Adaptive dynamic data driven framework

if participant has access to her exact location, a local task refinement step can be used to further optimize her set of tasks.
Local Task Refinement

- Minimize distance to travel
- To avoid over/under coverage, the difference between refined assignments \( y \) and global assignments \( x \) should be small
- NP-Hard Problem (Minimum Set Cover Problem reduction)
- Probabilistic greedy algorithms

\[
\begin{align*}
\min_y & \sum_{j \in M} d_{i,j} y_{i,j} \\
\text{s.t.} & \quad |y_i - x_i| < \epsilon \\
& \quad \sum_{j \in M} \frac{y_{i,j}}{k_j} \geq \sum_{j \in M} \frac{x_{i,j}}{k_j} \\
& \quad \sum_{j \in M} y_{i,j} d_{i,j} \leq b_i
\end{align*}
\]
Adaptive dynamic data driven framework
Experiments : Data Sets

- **Gowalla**: A real-world data set which contains check-in information of users in a location-based social network.

- **Geolife GPS Trajectories**: Collected in (Microsoft Research Asia) Geolife project by 182 users in a period of over three years (from April 2007 to August 2012).
Results

Relative task coverage for different coverage goal, $n = 200$, and $m = 200$, using datasets Gowalla
Penalized cost for different timepoints, $n = 100$, and $m = 100$, using dataset Geolife
Ongoing and Future Work

- Enhanced mobility modeling and trajectory prediction algorithms for more accurate participant tracking
- Implementing mobile app and server software for crowd sensing with dynamic task assignment
- Formal location perturbation mechanisms
- Trustworthiness of participants in terms of the quality of the data contributed by them