Title
Recruitment Services for Participatory Sensing Applications

Authors
Reddy, Sasank
Maldonado, Olmo
Burke, Jeff
et al.

Publication Date
2009-05-12

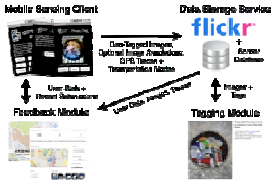
Recruitment Services for Participatory Sensing Applications

Sasank Reddy, Olmo Maldonado, Jeff Burke, Deborah Estrin, Mark Hansen, and Mani Srivastava
Urban Sensing | CENS - http://urban.cens.ucla.edu

Introduction: Use mobile devices to form interactive participatory networks where individual and institutional users gather, analyze, and share local data.

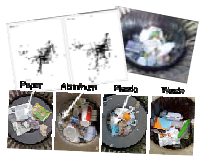
Leverages Large Installed Base of Mobile Phones and Open Web 2.0 Services

- Mobile phones offer media capture, signal processing, localization, network-connectivity
  - Web 2.0 tools and light-weight framework help support data collection campaigns



Enables Data Collection Campaigns

- GarbageWatch
Recycling Practices on Campus



- HabWatch
Harmful Algal Blooms

- WhatsBloomin
Blooming Flora on Campus

Problem Description: Based on the requirements of a campaign, determine the appropriate user base that will gather the data requested.

Organizer Provides Specifications for Campaign

- Budget, Life Time, Sensor Capabilities,
- Participation and Performance Standards
- Spatial / Temporal Coverage

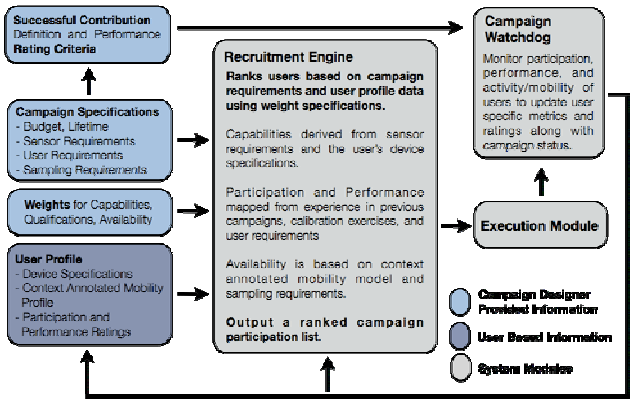
Participants Have Multi-Part Profiles

- Capabilities, Participation and Performance, and Availability

- **Capabilities:** Model of phone the users have available, what sensors they are willing to use with the phone, what specifications do these sensors have.
- **Availability:** Context annotated mobility models that help assess whether a user is appropriate for a particular campaign based on the spatial, temporal, and activity constraints.
- **Participation and Performance:** Metrics that keep track of participation and performance levels in the system as a whole and on a campaign basis.

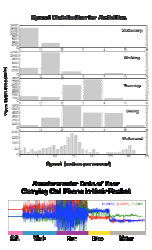
Proposed Solution: Transport Mode Classifier, Reputation Metrics, Mobility Modeling

Recruitment System Architecture



GPS and Accelerometer Data to Transport Mode

- Activities: Still, Walk, Run, Bike, Motorized Transport
- Features: GPS Speed, Acc. Variance and FFT (1-3 Hz)
- Classifier: Decision Tree + HMM
- Accuracy: > 90% using dataset of 16 users



GPS Data to Destinations and Routes

- Use Density Based Clustering (DBScan) to infer significant destinations
- Perform Hierarchical Clustering to group routes

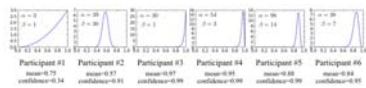
Reputation Metrics for Data Collectors

Concentrate on Campaign-Specific Measures

- Model reputation for each participant as a Beta distribution
- Beta is compact, efficiently updatable, handles stochastic and epistemic uncertainty
- Factors: timeliness, quality, relevance, participation likelihood
- Calculated based on calibration "exercises" where ground truth is obtained by an expert or through on-going data collection

$$f(p|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

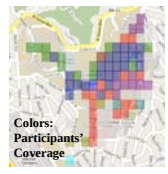
$$E(\alpha,\beta) = \frac{\alpha}{\alpha+\beta} \quad \begin{matrix} \alpha = \text{successful entries} + 1 \\ \beta = \text{unsuccessful entries} + 1 \end{matrix}$$



Transport Mode Mobility Profiles for Coverage

Max Cover Problem to Choose Participants

- Build an "association matrix" that captures the amount of time spent in a particular context during a time period
- Maximize coverage utility (U) provided by participants (p) while adhering to a budget (B) based on participant costs (C)



$$argmax\ U(P^*)\ subject\ to\ C(P^*) \leq B$$

Comparing Profiles for Similarity

- Singular Value Decomposition to obtain the eigenbehaviors (column signatures in association matrix)
- Compare periods by calculating th for different time periods.