An Algorithm to Coordinate Measurements Using Stochastic Human Mobility Patterns in Large-Scale Participatory Sensing Settings

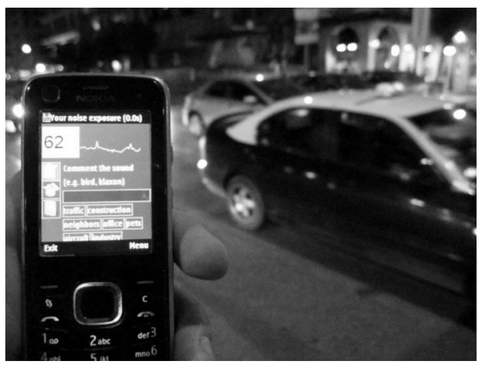
Alexandros Zenonos, Sebastian Stein and Nicholas R. Jennings
Agents, Interaction and Complexity Research Group
School of Electronics and Computer Science
University of Southampton
{az2g13, ss2, nrj}@ecs.soton.ac.uk

Introduction

Participatory sensing is becoming an effective and cheap tool for monitoring environmental phenomena.

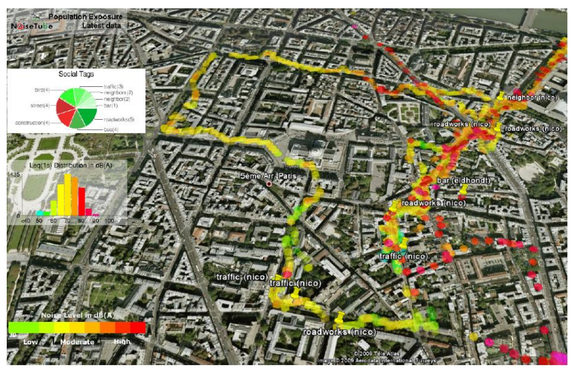
Participatory sensing is about crowdsourcing sensory information via sensors carried by ordinary people (i.e. non-experts).

Sensor information includes air quality, noise, radiation levels, light, humidity and gas concentration.



Problems in current participatory sensing applications:

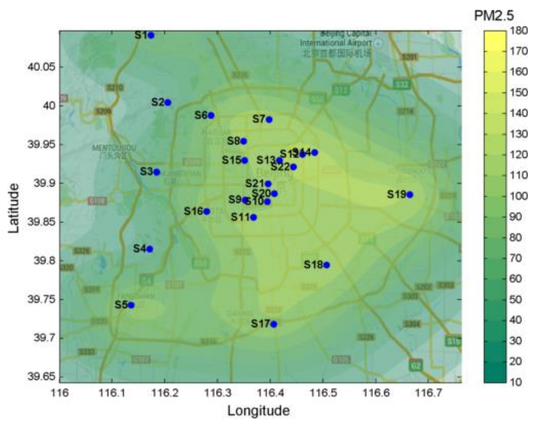
- Partial coverage of the areas of interest over time results in **limited knowledge** about the environment.
- Duplicate work, which results in energy loss.



Our Model

Aim

Provide algorithms for **coordinating measurements** in the participatory sensing setting for environmental monitoring in order to achieve better understanding of the phenomena and avoid duplicate work.



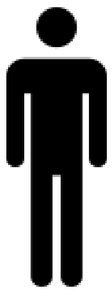
We use D-optimality Criterion to value the information at spatio-temporal locations.

$$I(X_B; X_A) = H(X_B) - H(X_B|X_A)$$

Use Gaussian Processes to model correlations over space and time.

$$f \sim GP(0, K(\boldsymbol{x}, \boldsymbol{x}'))$$

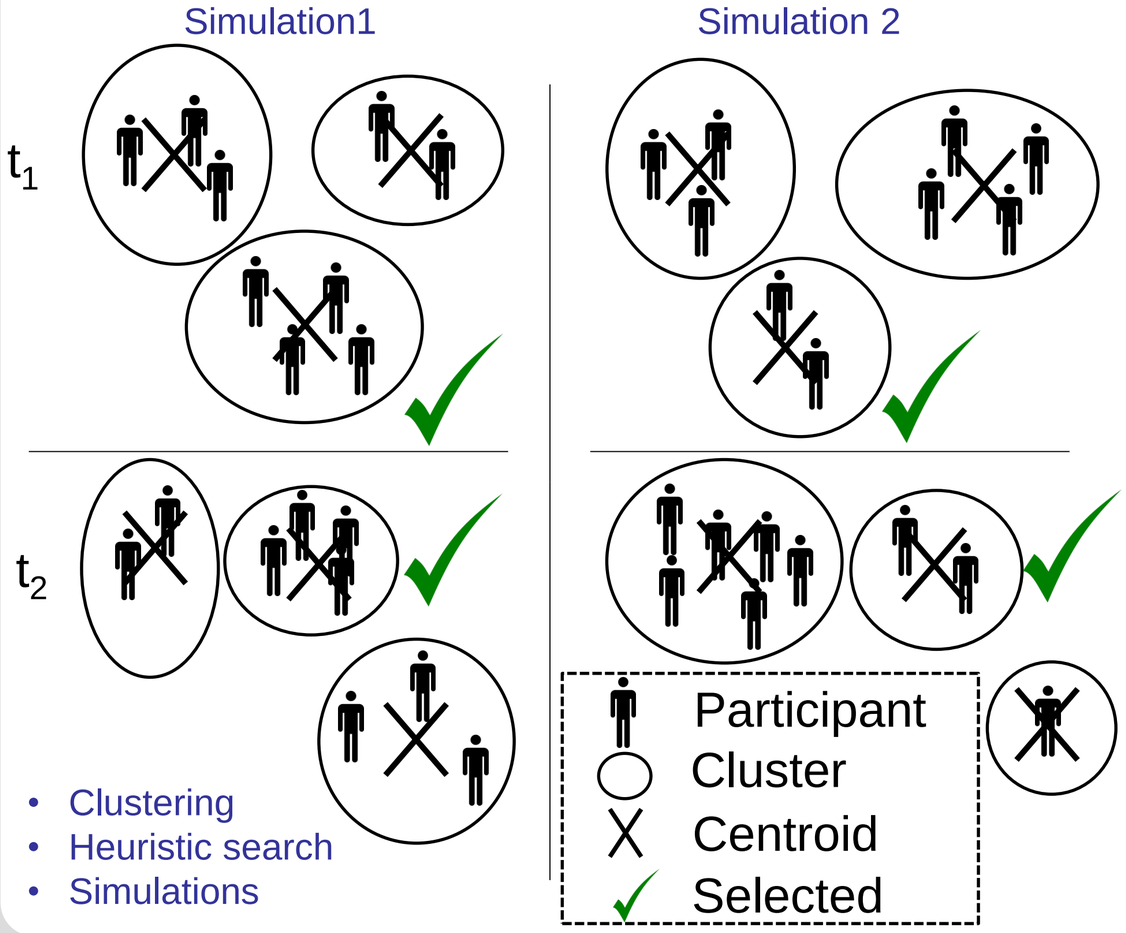
Participants have a budget



$$B_i \in \mathbb{N}, \qquad U(O_E) = \sum_{t=1}^{E} u(O_t)$$

Coordination Algorithm

Best-Match Algorithm



- Clustering
- Heuristic search
- Simulations

Experiments

Benchmarks

- Greedy algorithm: Greedily select the best observations at each timestep
- Patrol: Take all measurements at every timestep.
- Random: Take measurements randomly throughout time.
- Proximity-driven: Take measurements when in area of high uncertainty.
- Upper Bound: People have unlimited budget and their patterns are known.

Simulation

- Real sensor data used (U-air dataset).
- Real mobility patterns (Geolife dataset).
- 250-1000 participants

Key Results

Best-Match is up to 75% **better** than the state-of-the-art greedy algorithm and significantly faster.