Providing Long-Term Participation Incentive in Participatory Sensing

Lin Gao[*], Fen Hou[†], and Jianwei Huang[*]

[†] **Network Communications and Economics Lab (NCEL)**
The Chinese University of Hong Kong (CUHK), Hong Kong

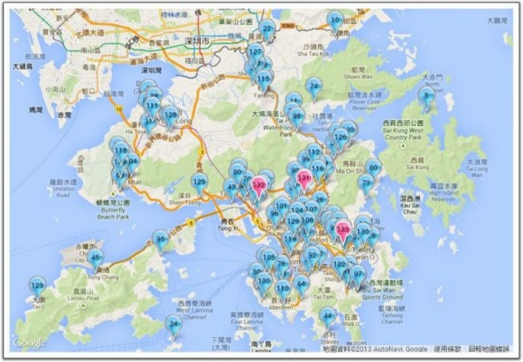
[†] **Electrical and Computer Engineering**
University of Macau, Macau

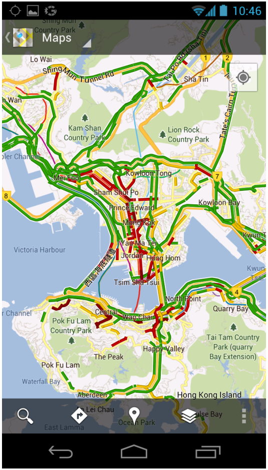
What is this about?

- **Mobile Crowdsensing**
  - Also known as Participatory Sensing;
  - A novel data collection and interpretation scheme, in which mobile users voluntarily participate in and actively contribute to the sensing system, by using their carrying smartphones or other custermized portable devices.

- **This work focuses on the incentive issue in mobile crowdsensing.**

Typical Examples





(a) Air Pollution in Hong Kong, (b) Road Traffic Congestion in Hong Kong

Incentive

- **Incentive** in Mobile Crowdsensing
  - Short-Term Incentive
    - Objective: Compensating the instantaneous sensing cost in a particular sensing action, e.g., energy consumption, transmission cost, etc;
    - Approaches: Pricing, Auction, Contract, etc;
    - Existing Works (Many): [T. Luo et al, INFOCOM 14], [I. Koutsopoulos, INFOCOM 13], [D. Yang et al, Mobicom 12], etc;
  - Long-Term Incentive
    - Objective: Encouraging the user participation in the long run, by guaranteeing an average Return-over-Investment (RoI);
    - Approaches: Dynamic Pricing;
    - Existing Works (Few): [J. S. Lee et al, PerCom 2010];

- *Most of the existing work focused on the short-term incentive; Only few works considered the long-term (user participation) incentive, but without mathematically rigorous analysis.*

Our Focus

- This work is to study a **mobile crowdsensing system** with the explicit consideration of **long-term participation incentive**;
  - Modeling
    - To model a location-aware, time-dependent crowdsensing system, and formulate the long-term user participation incentive expecitly;
  - Optimization
    - To optimize the sensor scheduling in the proposed crowdsensing system under different network information;
  - Incentive Mechanism
    - To incentivize mobile users to report their private information truthfully when information is asymmetric.

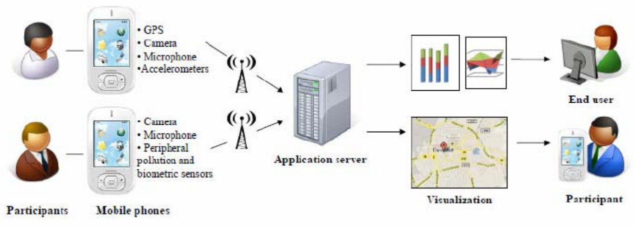
Outline

Background

- Mobile crowdsensing is enabled by the explosive increase of powerful mobile device (e.g., smartphones) with
  - Rich embedded sensors;
  - Advanced data process capability;
  - Programmable;
  - etc.
- Benefit: Low deploying cost, High sensing coverage;
- Application: Environment, infrastructure, and community monitor.
  - Real Examples:
    - Waze, https://www.waze.com/.
    - OpenSignal, http://opensignal.com/.
    - Sensorly, http://www.sensorly.com/.
    - NoiseTube, http://www.noisetube.net/.
    - Mobile Millennium, http://traffic.berkeley.edu/.
    - Intel Urban Atmosphere, http://www.urban-atmospheres.net/.
    - etc.

Architecture

- Service Provider
  - A application server that launches a set of sensing tasks with different data requirements for different purposes;
- Participants
  - A set of mobile users who actively participate in and contribute to one or multiple sensing task(s), by using their smartphones;
- End-users
  - A set of data consumers who access and consult the collected data.



Key Problem

- Service Provider
  - A application server that launches a set of sensing tasks with different data requirements for different purposes;
- Participants
  - A set of mobile users who actively participate in and contribute to one or multiple sensing task(s), by using their smartphones;
- End-users
  - A set of data consumers who access and consult the collected data.

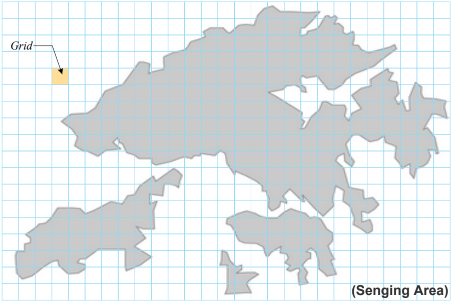
Key Problem — Sensor Scheduling

Who senses What at Where, and When?

Outline

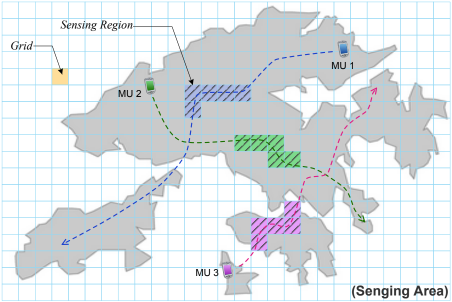
System Model

- One Service Provider (SP)
  - Launch time-dependent and location-aware sensing tasks;
    - ★ Require data in different locations periodically;
  - Divide the total sensing area into I grids: $i = 1, ..., I$;
  - Divide the total sensing time into T slots: $t = 1, ..., T$;
    - ★ $w_i[t]$: the value of data in grid i at time slot t.



System Model

- Mobile Users: $n = 1, ..., N$
  - Mobility: Move randomly according to certain mobility pattern;
  - Sensing Region: The locations that a user passes (hence can sense);
    - ★ $z_{n,i}[t] \in \{0, 1\}, i \in \mathcal{I}$: The sensing region of user n in time slot t;



System Model

- Mobile Users: $n = 1, ..., N$
  - Sensing Schedule: Choose a subset of users to perform sensing;
    - ★ $x_n[t] \in \{0, 1\}$: Scheduling indicator of user n in time slot t;
  - Sensing Cost: The total instantaneous cost of all scheduled users;

$$C[t] \triangleq \sum_{n \in \mathcal{N}} c_n[t] \cdot x_n[t]$$

    - ★ $c_n[t] \geq 0$: The sensing cost of user n in time slot t;
    - ★ E.g., the energy consumption and the transmission expense;
  - Sensing Value: The total data value generated by all scheduled users;

$$V[t] \triangleq \sum_{i \in \mathcal{I}} w_i[t] \cdot \underbrace{\left[\sum_{n \in \mathcal{N}} x_n[t] \cdot z_{n,i}[t] \right]_0^1}_{y_i[t]}$$

    - ★ $y_i[t] \in \{0, 1\}$: denote whether a grid i is sensed by at least one user;
  - Social Welfare

$$S \triangleq \frac{1}{T} \sum_{t \in \mathcal{T}} (V[t] - C[t])$$

System Model

- Mobile Users: $n = 1, ..., N$
  - **Long-term Participation Incentive**
    - ⋆ Depends on the user's Return on Investment (RoI);
    - ⋆ Estimated by the user's Scheduling Probability;
  - **Participatory Constraint (New)**

$$D_n \leq d_n(\mathbf{x}_n) \triangleq \frac{1}{T} \sum_{t \in \mathcal{T}} x_n[t]$$

- ⋆ $d_n(\mathbf{x}_n)$: the time average scheduling probability of user n;
- ⋆ D_n: the dropping threshold of user n, that is, the minimum scheduling probability with which user n is willing to stay in the sensing system;
- ⋆ *The scheduling probability captures the user RoI in the long run;*
- ⋆ *This constraint captures the long-term user participation incentive.*

System Model

- The objective is to study *the optimal scheduling of users that maximizes the social welfare*, considering the user participatory constraint (long-term participation incentive).
- The formulation and solution depend on network information;
  - With complete information: offline optimization;
  - With incomplete information: online optimization;

Network Information

The network information in each time slot t consists of the location data value, user sensing region and sensing cost, i.e.,

$$\theta[t] \triangleq \{w_i[t], \mathbf{z}_n[t], c_n[t], \ \forall i \in \mathcal{I}, n \in \mathcal{N}\} .$$

System Model

- **Information Scenairo**
  - Regarding future information,
    - ★ Complete future information: $\theta[t], \forall t = 1, ..., T$;
    - ★ Stochastic future information: $f(\theta)$;
    - ★ No future information: nothing;
  - Regarding current information (realization),
    - ★ Symmetric information: the SP observes the all the information realized in the current time slot;
    - ★ Asymmetric information: the SP cannot observe the private information of users realized in the current time slot;

Outline

Benchmark Solution

- Complete future information (Symmetric current information)

$$
\begin{aligned}
\max_{\mathbf{x}} \quad & \frac{1}{T} \sum_{t \in \mathcal{T}} \big(V[t] - C[t] \big) \\
\text{s.t.} \quad & \text{(a) } x_n[t] \in \{0, 1\}, \quad \forall n \in \mathcal{N}, \forall t \in \mathcal{T}, \\
& \text{(b) } D_n \le d_n(\mathbf{x}_n), \quad \forall n \in \mathcal{N}.
\end{aligned} \tag{1}
$$

  - The above problem is an off-line allocation problem, and the solution presents *the explicit allocation of each user in each time slot*.
  - Formulating and solving the above problem requires the stochastic future network information.

Benchmark Solution

- Stochastic future information (Symmetric current information)

$$
\begin{aligned}
\max_{\mathbf{X}} \quad & \int_{\theta \in \Theta} \big(V(\theta) - C(\theta) \big) \cdot f(\theta) \mathrm{d}\theta \\
\text{s.t.} \quad & \text{(a) } x_n(\theta) \in \{0, 1\}, \quad \forall n \in \mathcal{N}, \forall \theta \in \Theta, \\
& \text{(b) } D_n \le d_n(\mathbf{x}_n), \quad \forall n \in \mathcal{N},
\end{aligned} \tag{2}
$$

  - The above problem is an off-line allocation problem, and the solution defines *a contingency plan that specifies the allocation of each user under each possible information realization θ*.
  - Formulating and solving the above problem requires the stochastic future network information.

Benchmark Solution

- Equivalence between two benchmarks
  - S°: maximum social welfare with complete future information;
  - S^*: maximum social welfare with stochastic future information;

Lemma

If $T \to \infty$, then $S^* \to S^\circ$.

Solution

- No future information (Symmetric current information)
  - **Lyapunov-based Optimization**
    - ★ A widely-used technique for solving stochastic optimization problems with time-average constraints, without future information;
    - ★ Key idea: Queue stability $\Leftrightarrow$ Time-average constraint
  - In our problem, there is a time-average constraint:

$$D_n \leq d_n(\mathbf{x}_n) \triangleq \frac{1}{T} \sum_{t \in \mathcal{T}} x_n[t]$$

  - Hence, we solve the problem using Lyapunov optimization framework.

Solution

- No future information (Symmetric current information)
  - Time-Average Constraint in our problem:

$$D_n \leq d_n(\mathbf{x}_n) \triangleq \frac{1}{T} \sum_{t \in \mathcal{T}} x_n[t]$$

  - Virual Queue Definition — User Virtual Request
    - ★ One virtual request represents that "*to satisfy the user participatory constraint, the user* ***should*** *be scheduled as sensor one more time*";
    - ★ Arrival: D_n (constant) in each time slot;
    - ★ Departure: $x_n(t)$ (schedule) in time slot t;

$$q_n^{t+1} = \left[q_n^t - x_n[t]\right]^+ + D_n,$$

Virtual Queue

Solution

- **Lyapunov-based Policy 1 (Information Symmetry)**
  - Initialization: $\mathbf{q} = \mathbf{q}^0$;
  - For each time slot $t = 0, 1, ..., T$
    - Allocation Rule:

$$\mathbf{x}^{\dagger}[t] = \arg\max_{\mathbf{x}[t]} \left(V[t] - C[t] + \sum_{n \in \mathcal{N}} \frac{q_n^t}{\phi} \cdot x_n[t] \right)$$

    - Queue Update Rule:

$$q_n^{t+1} = \left[q^t - x_n^{\dagger}[t] \right]^+ + D_n, \quad \forall n \in \mathcal{N}$$

Solution

- Optimality of Policy 1
  - $S^{\dagger}[t]$: the social welfare generated in each slot t
  - S^{*}: the maximum social welfare benchmark;

Theorem (Optimality of Policy 1 (Information Symmetry))

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t\in\mathcal{T}} \mathbf{E}(S^{\dagger}[t]) \geq S^{*} - \frac{B}{\phi}.$$

That is, Policy 1 converges to the maximum social welfare benchmark asymptotically, with a controllable approximation error bound $O(1/\phi)$.

Solution

- No future information (Asymmtric current information)
  - *The allocation rule in Policy 1 requires all of the realized information in each time slot*;
  - Under asymmetric information, however, the SP cannot observe the realized private information of users (i.e., sensing costs);
  - Incentive compatible mechanism is necessary for eliciting the realized private information of users in each time slot
    $\rightarrow$ **VCG Auciton**

Solution

- **Lyapunov-based VCG Policy 2 (Information Asymmetry)**
  - Initialization: $\boldsymbol{\mu} = \boldsymbol{\mu}^0$;
  - Denote $c'_n[t]$ as the bid of each user n;
  - For each time slot $t = 0, 1, ..., T$
    - ★ Allocation Rule:

$$\mathbf{x}^{\ddagger}[t] = \arg\max_{\mathbf{x}[t]} \ V[t] - \sum_{n \in \mathcal{N}} x_n[t] \cdot (c'_n[t] - \mu_n^t)$$

    - ★ Payment Rule:

$$p_n[t] = x_n^{\ddagger}[t] \cdot \left(V^{\ddagger}[t] - C^{\ddagger}_{-n}[t] - \widetilde{S}^{\sharp}_{-n}[t] + \mu_n^t \right)$$

    - ★ Queue Update Rule:

$$\mu_n^{t+1} \cdot \phi = \left(\left[\mu_n^t \cdot \phi - x_n^{\ddagger}[t] \right]^+ + D_n \right), \forall n \in \mathcal{N}$$

Solution

- Truthfulness and Optimality of Policy 2

Theorem (Truthfulness of Policy 2 (Asymmetry))

The auction in Policy 2 is truthful.

Theorem (Optimality of Policy 2 (Asymmetry))

The auction in Policy 2 achieves the same asymptotically optimal social welfare as in Policy 1.

Summary of Solutions

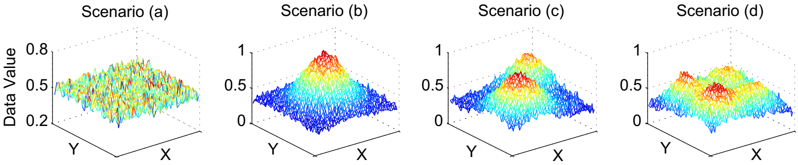
Table: A Summary of Solutions

Future Information	Current Information	Solution	Performance	Section
Complete / Stochastic	Symmetric	Off-line Solution	Optimal (Benchmark)	III
No	Symmetric	On-line Policy 1	Asymptotic Optimal	IV
No	Asymmetric	On-line Policy 2	Truthful, Asymptotic Optimal	V

Outline

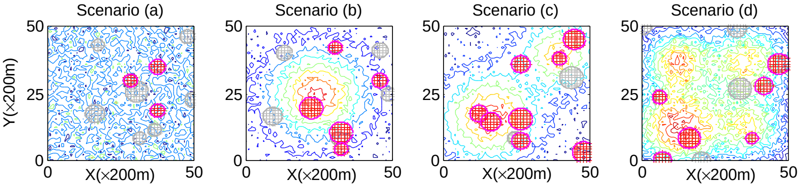
Simulations

- Simulation Scenarios
  - A square of 10km×10km, divided into 2500 grids;
  - 4 scenarios (in term of data value): (a) no hotspot, (b) one hotspot, (c) two hotspots, and (d) four hotspots



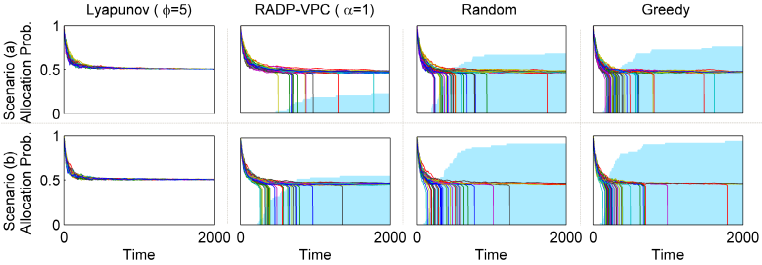
Simulations

- A Snapshot of Sensor Selection
  - Red circle: selected; Grey circle: not selected.



Simulations

- User Dropping Probability
  - Allocation Probability Dynamics and User Dropping in Scenario (a) (the first row) and Scenario (b) (the second row);

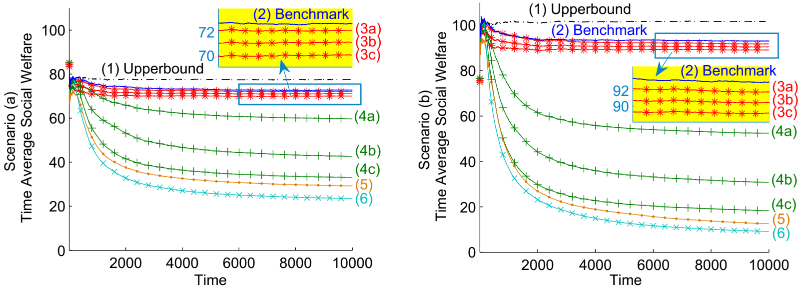


★ The dropping of a user is illustrated by the sudden decrease of its allocation;

★ The percentage of dropping users is denoted by the blue shadow area;

Simulations

- Achieved Social Welfare
  - Average Social Welfare under Different Policies in Scenario (a) (left) and Scenario (b) (right);



- ★ (1) Upperbound (Without Participatory Constraint);
- ★ (2) Benchmark (with complete/stochastic information);
- ★ (3a)-(3c) *Lyapunov-based policy* ($\phi = \{20, 10, 5\}$) proposed in this work;
- ★ (4a)-(4c) *RADP-VPC policy* ($\alpha = \{1, 0.5, 0.2\}$) proposed in [Lee et al. 2010];
- ★ (5) *Random policy*; (6) *Greedy policy*.

Conclusion

- First work analyzing the long-term participation incentive strictly;
- Formulate & solve problem under different information scenarios.

- Future Extension
  - More specific way to formulate long-term participation incentive;
  - Study the truthful mechanism design when users are not myopic.

Thank You

LGAO@IE.CUHK.EDU.HK

Network Communications and Economics Lab (NCEL)

The Chinese University of Hong Kong (CUHK)