Sensing as a Service (S^{2}aaS): Vision, System and Applications

Jian Tang
Department of Electrical Engineering and Computer Science
Syracuse University

Outline

- Vision
- System: Design and Implementation of A Unified Platform
- App 1: Objective Ranking based on Mobile Phone Sensing (SOR)
- App 2: Lifestyle Learning via Mobile Phone Sensing (LIPS)
- New Research Directions

Outline

- ***<u>Vision</u>***
- System: Design and Implementation of A Unified Platform
- App 1: Objective Ranking based on Mobile Phone Sensing (SOR)
- App 2: Lifestyle Learning via Mobile Phone Sensing (LIPS)
- New Research Directions

Our Vision

- **Our vision:** To offer every person a third eye to extend his/her sense to every corner in this world.

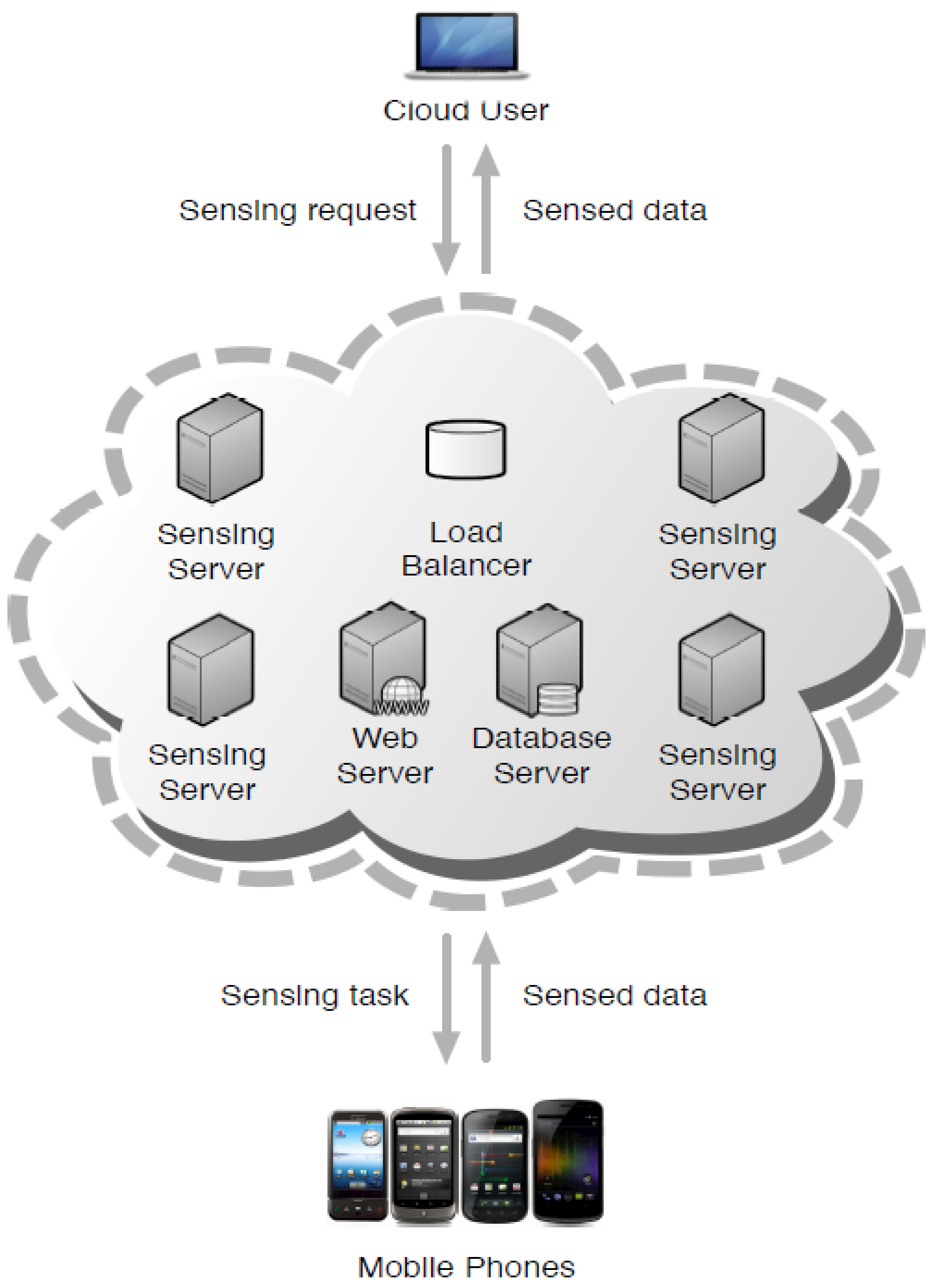
- **Our approach:** To develop a unified and green cloud computing platform for mobile phone sensing.

Mobile Phone Sensing

- Mobile phones (iPhone and Android phones) and their sensors: Camera, GPS, digital compass, accelerometer, 3G/4G/WiFi, thermometer, …
- External sensors (such as Google Glass, Fitbit, Sensordrone, etc) can be attached to a mobile phone via a network interface such as bluetooth.
- **Participatory sensing:** mobile users actively engage in sensing activities by manually determining how, when, what, and where to sense.
- **Opportunistic sensing:** sensing activities are fully automated without the involvement of mobile users.

A Sensing as a Service (S^2aaS) Cloud

- Multiple sensing servers can be deployed to handle sensing requests from different locations.
- When a cloud user initiates a sensing request, it will be forwarded to a sensing server which will then push the request to a subset of mobile phones that happen to be in the area of interest.
- The corresponding sensing task will be fulfilled by these mobile phones.
- The sensed data will then be collected by a sensing server, stored in the database and returned to the requester.
- **Mobile phones can be both service users and service providers**.



Research Challenges

- **A unified platform:** It must be able to support various mobile phone sensing applications with different embedded and external sensors.

- **Energy efficiency:** It must be energy-efficient.

- **Incentive mechanisms:** It must have effective incentive mechanisms that can be used to attract mobile users to participate in sensing activities.

Existing Systems and Their Shortcomings

- In [5], Das *et al.* presented a Platform for Remote Sensing using Smartphones (PRISM), which allows application writers to package their applications as executable binaries and push them automatically to an appropriate set of phones.

- The Bubble-Sensing [10] allows sensing tasks to be posted at specific physical locations of interest.

- Cornelius *et al.* introduced AnonySense in [3], which is a sensor tasking and reporting system designed for both participatory and opportunistic sensing.

- Microblogs [7] is another system for participatory sensing, where users upload blogs annotated with sensed information (e.g., photos) to a micro-blog server. Mobile devices also upload their locations to the server periodically.

Existing Systems and Their Shortcomings

- PRISM [5] uses executable binaries to deliver sensing tasks to mobile phones, which is platform-dependent (Windows Mobile only) and may cause security issues.
- AnonySense [3] uses a customized, yet very limitedly-used Lisp dialect for implementation.
- Important issues, such as energy-efficiency and user incentives, have not been addressed.

References

- **NSF Grants:**
  1. **A Green and Incentive Platform for Mobile Phone Sensing (#1218203)**
  2. **Enabling High-Quality Mobile Crowdsourcing with Lifestyle-aware and Energy-efficient Control (#1525920)**

- **[Sensors'13] X. Sheng, J. Tang, X. Xiao and G. Xue, Sensing as a Service: challenges, solutions and future directions, *IEEE Sensors Journal* (One of the 25 most downloaded papers) Vol. 3, No. 10, 2013, pp. 3733-3741.**
- **[ICDCS'14]X. Sheng, J. Tang, J. Wang and G. Xue, SOR: an objective ranking system based on mobile phone sensing, *IEEE ICDCS'2014*.**
- **[Globecom'16] X. Sheng , J. Tang, J. Wang, T. Li, G. Xue and D. Yang, LIPS: lifestyle learning via mobile phone sensing, *IEEE Globecom'2016*, Accepted.**

- [Globecom'15] J. Wang, J. Tang, X. Sheng, G. Xue and D. Yang, Enabling green mobile crowd sensing via optimized task scheduling on smartphones, ***IEEE Globecom'2015*** (**Best Paper Award).**
- [IoT13]X. Sheng, J. Tang, X. Xiao and G. Xue, GPS-less sensing scheduling for mobile crowd sensing, ***IEEE Internet of Things Journal***, Vol. 1, No. 4, 2014, pp. 328-336.
- [Infocom'12] X. Sheng, J. Tang and W. Zhang, Energy-efficient collaborative sensing with mobile phones, ***IEEE Infocom'2012***, pp. 1916-1924.

- [ICDCS'16]J. Wang, J. Tang, D. Yang, E. Wang and G. Xue, Quality-aware and fine-grained incentive mechanisms for mobile crowdsensing, ***IEEE ICDCS'2016***, pp. 354-363.
  [Mobicom'12] D. Yang, G. Xue, X. Fang and J. Tang, Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing, ***ACM MobiCom'2012***, pp. 173-184.

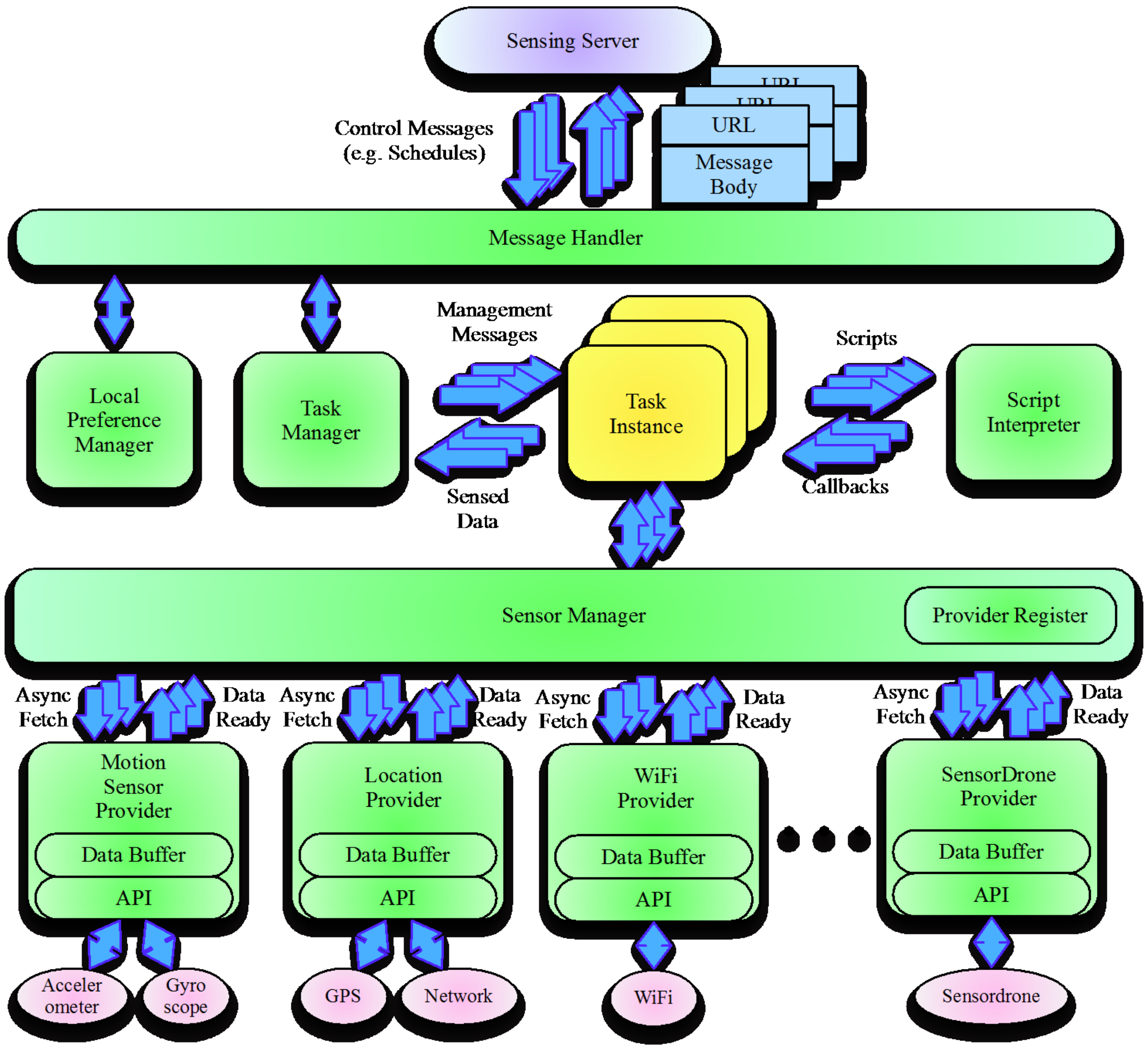
Outline

- Vision
- ***<u>System: Design and Implementation of A Unified Platform</u>***
- App 1: Objective Ranking based on Mobile Phone Sensing (SOR)
- App 2: Lifestyle Learning via Mobile Phone Sensing (LIPS)
- New Research Directions

Design and Implementation Considerations

- **Code portability, fast loading, security:** scripting language
- **Support for various applications**
- **Support for various sensors**
- **Re-configurability:** modular design and well-defined control messages.
- **Energy-efficiency:** sensing scheduling
- **Mobile user recruitment**: game-theoretic incentive mechanisms

Mobile Frontend: Mobile App



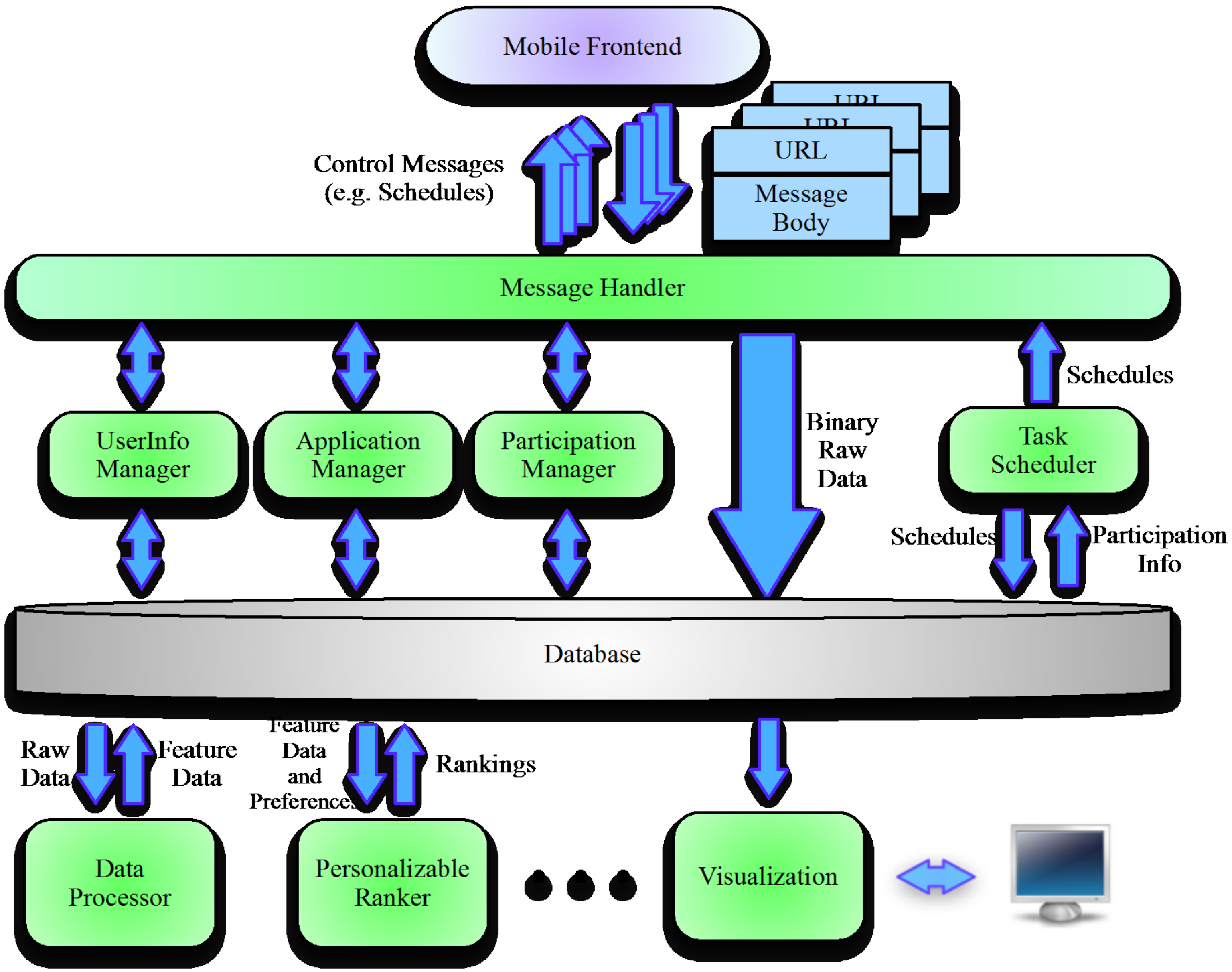

```
—[[ If Sensordrone is connected, get 5 light readings, with an interval of 2s]]—

if sensordrone_connected() then
      get_light_readings(5, 2)
end

—[[ If GPS query is permitted and the location data could be obtained,
       then, get 2 GPS location readings; otherwise, get 2 coarse locations. ]]—

if fine_location_permitted and gps_data_available() then
         get_location(2, 'fine')
else
         get_location(2, 'coarse')
end
```

Sensing Server



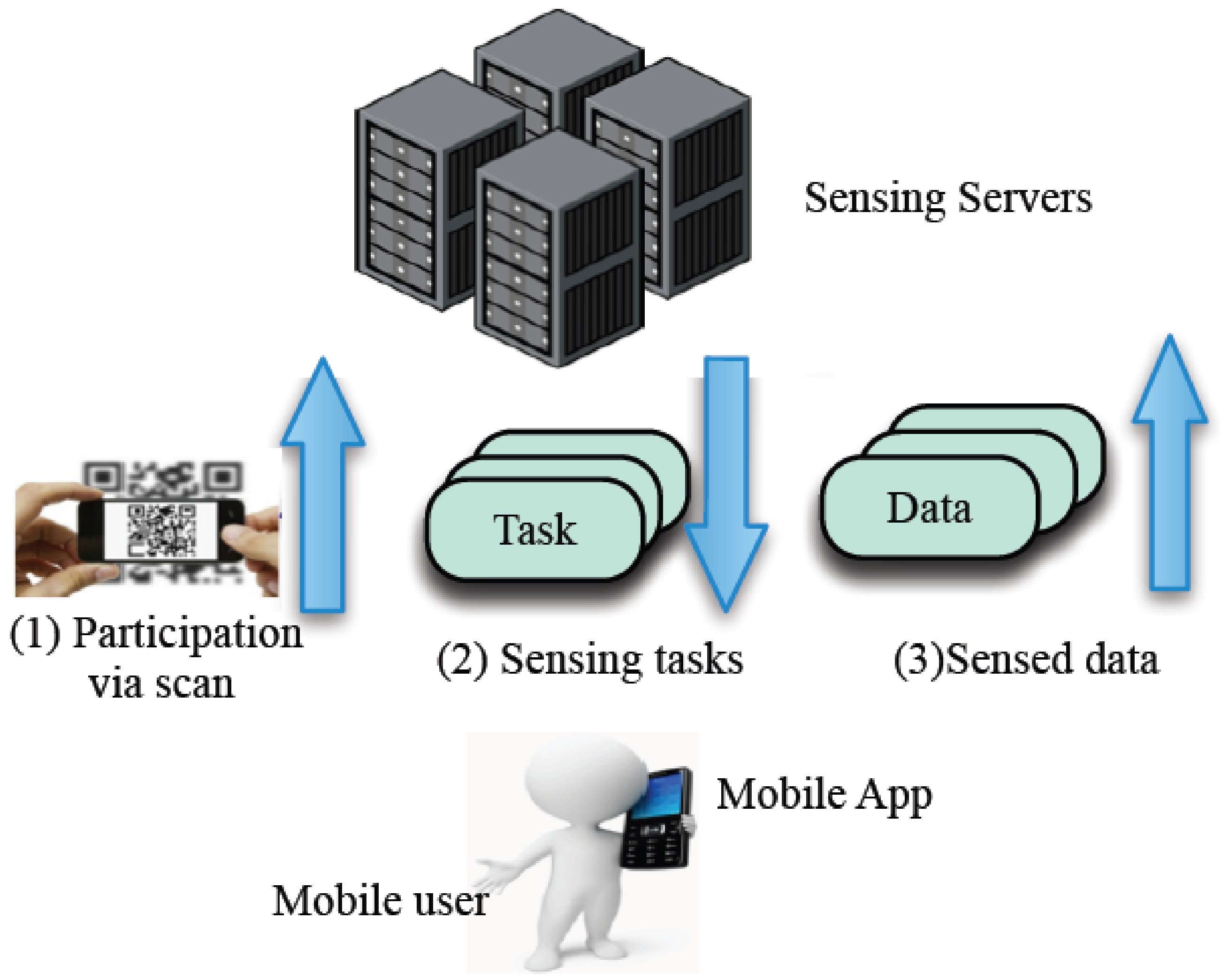
Outline

- Vision
- System: Design and Implementation of A Unified Platform
- ***App 1: Objective Ranking based on Mobile Phone Sensing (SOR)***
- App 2: Lifestyle Learning via Mobile Phone Sensing (LIPS)
- New Research Directions

SOR

- Ranking target places based on objective data (collected via mobile phone sensing).
- Energy-efficient sensing scheduling
- Personablizable ranking

Application Logic



Ranking: Data Pre-processing

- SOR’s ranking is conducted based on values of a set of features, such as temperature, WiFi signal strength, roughness of road surface.

- Raw data need to be pre-processed to calculate a numerical value for each feature.

- SOR calculates feature values and stores them into the database as input for ranking.

Ranking: Problem

- Input:
  - Feature matrix: $\mathbf{H} =< h_{ij} >, i \in \{1, \cdots, N\}, j \in \{1, \cdots, M\}$, where N and M are the numbers of target places and features respectively.
  - $\mathbf{U} =< u_j >, j \in \{1, \cdots, M\}$, where u_j is the value preferred by the user on feature j.
  - $\mathbf{W} =< w_j >, j \in \{1, \cdots, M\}$, where w_j is the weight given by the user on feature j.

- Output: a ranking **R** of target places for the user

Ranking: Algorithm

- Calculate the distances between input feature values and preferred values.
- Produce an individual ranking on each feature by sorting all the target places in the ascending order of the corresponding distances.
- Aggregate individual rankings produced (based on a single feature) in the second step to generate the final ranking.

Kemeney Distance

- Kemeney distance: the number of pariwise violations between two rankings.

- Example: R_1: A, B, C; R_2: B, C, A.

$$d_K(\mathbf{R}_1, \mathbf{R}_2) = 2$$

Kemeney Distance

- Weighted K-ranking distance: the quality of a ranking based on user preferences.

  Let $\mathbf{\Omega}$ be a collection of M individual rankings on all features

  $$\mathbf{\Omega} = \{\mathbf{R}_j : j \in \{1, \cdots, M\}\}.$$

  The weighted K-ranking distance from a ranking $\mathbf{R}$ to a collection of individual rankings $\mathbf{\Omega}$ is:

  $$\kappa_K(\mathbf{R}, \mathbf{\Omega}) = \sum_{j=1}^{M} w_j * d_K(\mathbf{R}, \mathbf{R}_j),$$

  where w_j is the weight assigned to feature j by the user.

- The ranking problem is to find a ranking such that its weighted ranking distance is minimized among all rankings.

Footrule Distance

- Bad news: computing optimal aggregated ranking based on K-distance is NP-hard.
- Good news: Kemeney distance can be approximated by Footrule distance:
  - Definition: $d_f(\mathbf{R}_1, \mathbf{R}_2) = \sum_{i=1}^{N} |\pi(i, \mathbf{R}_1) - \pi(i, \mathbf{R}_2)|$
  - Bound: $d_K(\mathbf{R}_1, \mathbf{R}_2) \leq d_f(\mathbf{R}_1, \mathbf{R}_2) \leq 2d_K(\mathbf{R}_1, \mathbf{R}_2)$
- Weighted f-ranking distance is: $\kappa_f(\mathbf{R}, \mathbf{\Omega}) = \sum_{j=1}^{M} w_j * d_f(\mathbf{R}, \mathbf{R}_j)$
- Footrule distance based ranking problem: $\mathbf{R}^* = \underset{\mathbf{R}}{\operatorname{argmin}}\, \kappa_f(\mathbf{R}, \mathbf{\Omega})$

Aggregation

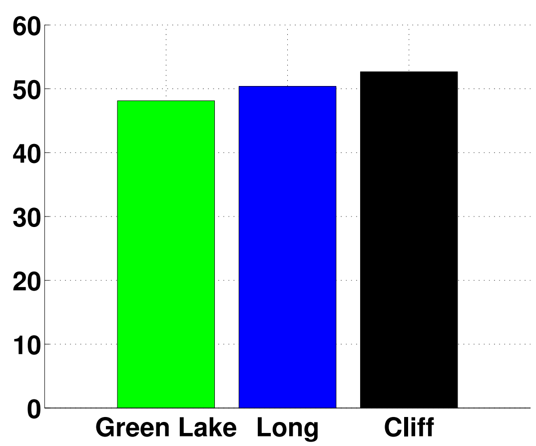
- Unweighted of the footrule distance based ranking problem could be transferred to a minimum cost perfect matching problem, which can be solved efficiently in polynomial time.

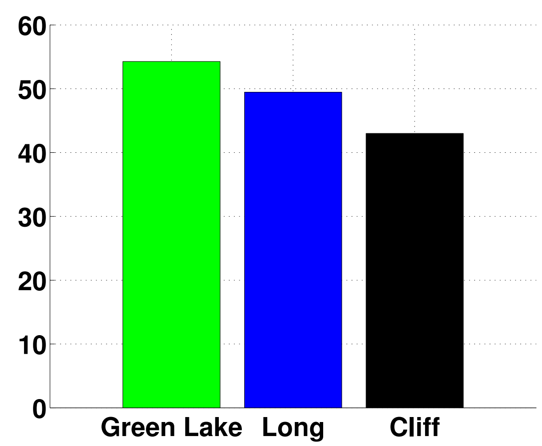
- Our *weighted* version can be efficiently solved by constructing an auxiliary flow graph and using a min-cost flow based algorithm.

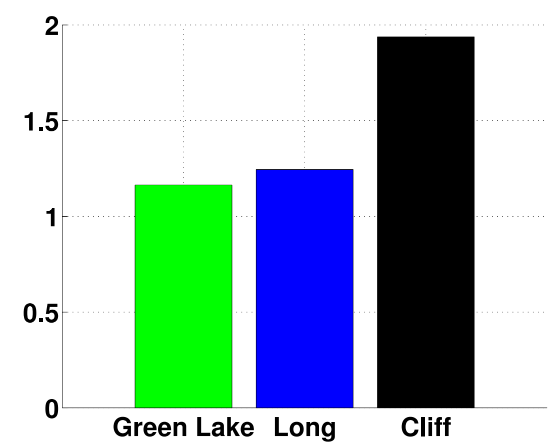
Field Testing Results

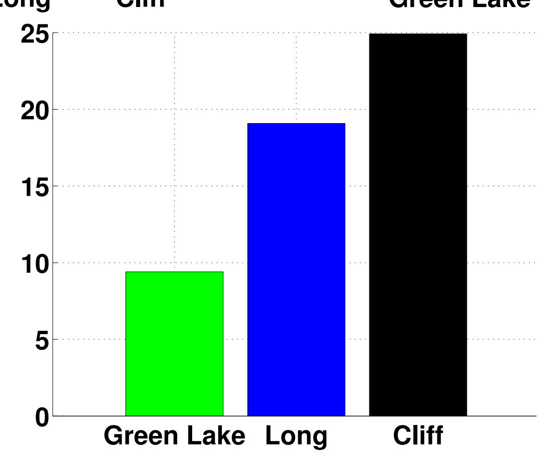
- **Target places:** three hiking trails in or around Syracuse
  - Green Lake Trail (in the Green Lake State Park), the Long Trail and the Cliff Trail (both of them are in the Clark Reservation)
- **Features:**
  - Temperature: it is an average of all temperature sensor readings;
  - Humidity: it is an average of all humidity sensor readings;
  - Roughness of road surface: it is an average of the standard deviations of all accelerometer's readings within Δt
  - Curvature: it is calculated based on GPS trajectories
  - Altitude change: it is the standard deviation of averages of all altitude sensor readings within Δt.

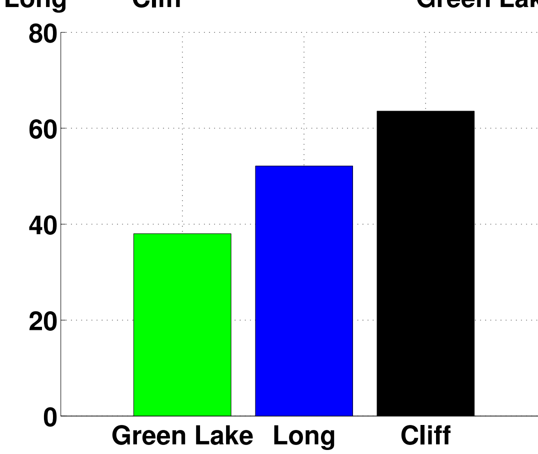
Field Testing Results











Ground Truth

	Real Users' Comments	Key Opinions
Green Lake	"…trails around the lake to run…" "…this is a great running trail…" "…trails… are wide and mostly flat …quite suitable for beginner…" "…The trail is almost entirely level…"	very flat, easy.
Long	"...wood land and meadow…" "... which runs through a meadow""...parts are easy…"	flat, easy.
Cliff	"…rocky outcrops…""...There are a couple of steep billy climbs…""This trail is extremely difficult,..."	rocky, difficult.

Field Testing Results

Alice

An experienced hiker who prefers difficult trails

Feature	Pre. Val.	Weight
Temp.	73 (Default)	3
Humi.	30%(Default)	3
Roug.	Max	5
Curv.	Max	5
Alti.	Max	5

Bob

A beginner who likes dry and even trails.

Feature	Pre. Val.	Weight
Temp.	73 (Default)	3
Humi.	Min	5
Roug.	Min	5
Curv.	Min	0 (does not care)
Alti.	Min	3

Chris

A beginner who likes jogging near a lake/sea/river.

Feature	Preferences	Weight
Temp.	73 (Default)	0 (does not care)
Humi.	55%	5
Roug.	Min	5
Curv.	Min	5
Alti.	Min	5

Rankings of Hiking Trails Computed by SOR

User	No. 1	No. 2	No. 3
Alice	Cliff Trail	Long Trail	Green Lake Trail
Bob	Long Trail	Cliff Trail	Green Lake Trail
Chris	Green Lake Trail	Long Trail	Cliff Trail

Sensing Scheduling: Model

- A sensing scheduling period $[t^S, t^E]$ is divided in to a set $\mathbf{T}$ of N time intervals with equal time durations.
- $p(t_i, t_j)$ means that if a measurement is taken at t_i , the reading at t_j stays the same with a probability of $p(t_i, t_j)$.
- A sensing schedule is a set Φ of time instants which specify when to sense.
- The probability of time instant t_j is covered by a given sensing schedule Φ is:

$$p(t_j, \mathbf{\Phi}) = 1 - \prod_{t_i \in \mathbf{\Phi}} (1 - p(t_i, t_j))$$

Sensing Scheduling: Problem

- Input:
  - A set $\mathbf{T}$ of equally spaced instants within a scheduling period $[t^S, t^E]$
  - The duration each mobile user k participating in sensing activities
- Output:
  - Sensing schedule set $\mathbf{\Phi} = \{\mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots, \mathbf{\Phi}_K\}$, each $\mathbf{\Phi}_k \in \mathbf{\Phi}$ is a set of time instants falling into $[t_k^S, t_k^E]$.
- Problem formulation:

$$\max_{\{\mathbf{\Phi}_1, \cdots, \mathbf{\Phi}_K\}} \sum_{t_j \in \mathbf{T}} \sum_{k=1}^{K} p(t_j, \mathbf{\Phi}_k)$$

Subject to:

$$|\mathbf{\Phi}_k| \leq N_k^B, k \in \{1, \cdots, K\}.$$

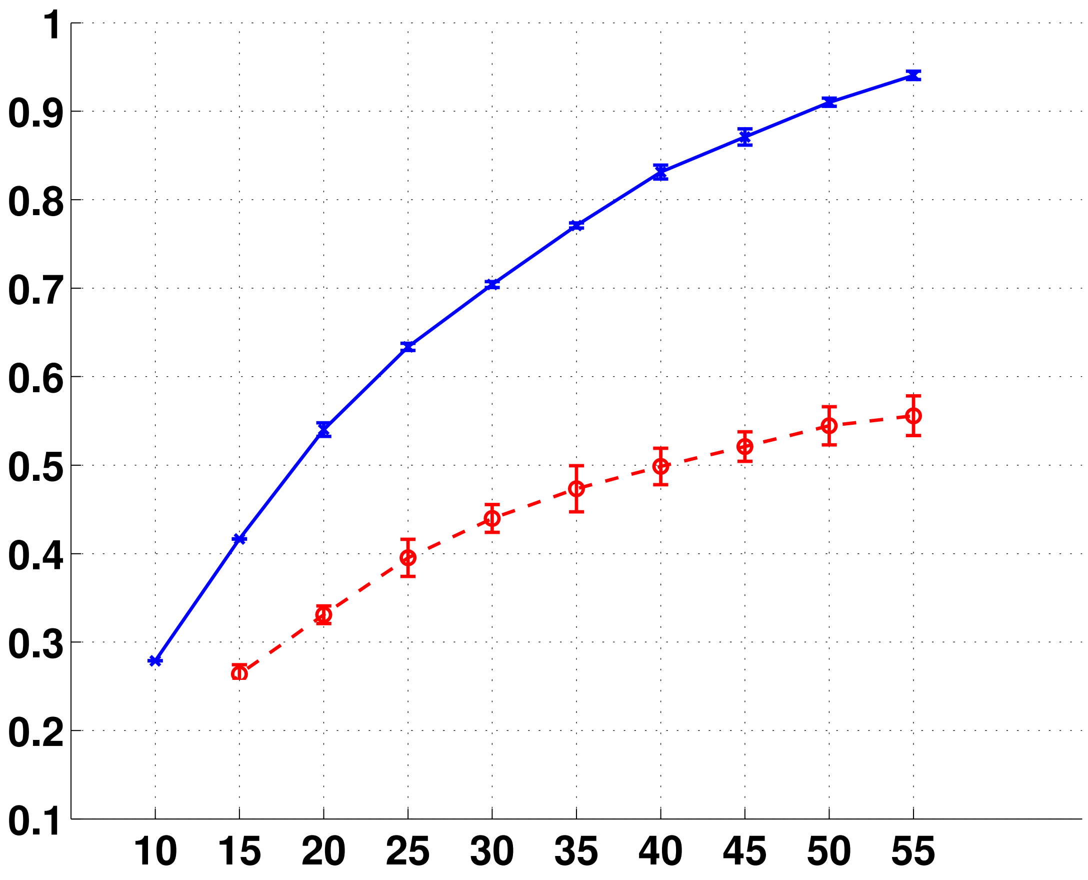
Sensing Scheduling: Algorithm

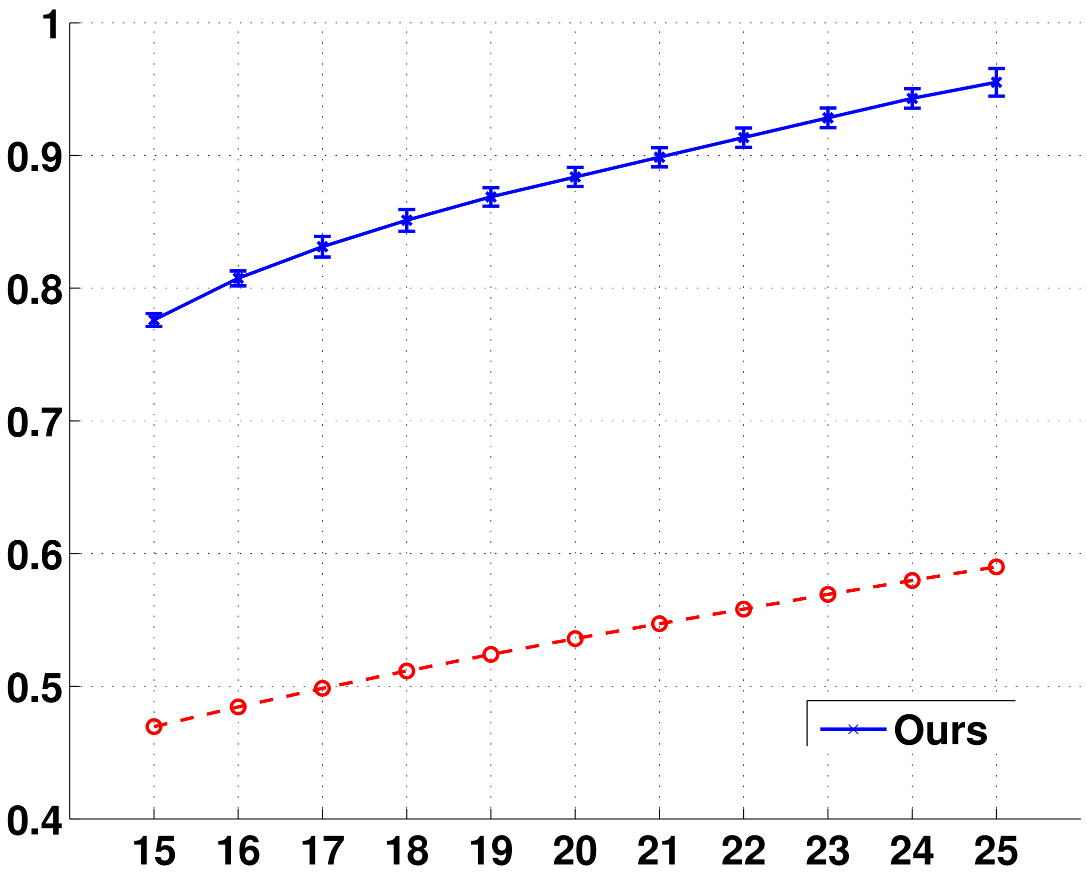
- **Sensing Scheduling Algorithm:** Keep adding into the schedule the time instant that can result in the maximum incremental coverage until no mobile users can be scheduled to sense more without violating their budget constraints.

- It is a 1/2-approximation algorithm.

Sensing Scheduling: Simulation Results

- Our scheduling algorithm outperforms the baseline algorithm by 65% in terms of average coverage probability.
- The variance of the coverage probability given by our scheduling algorithm is always less than that given by the baseline algorithm, which means our algorithm is more stable and is suitable for various situations.





Outline

- Vision
- System: Design and Implementation of A Unified Platform
- App 1: Objective Ranking based on Mobile Phone Sensing (SOR)
- ***<u>App 2: Lifestyle Learning via Mobile Phone Sensing (LIPS)</u>***
- New Research Directions

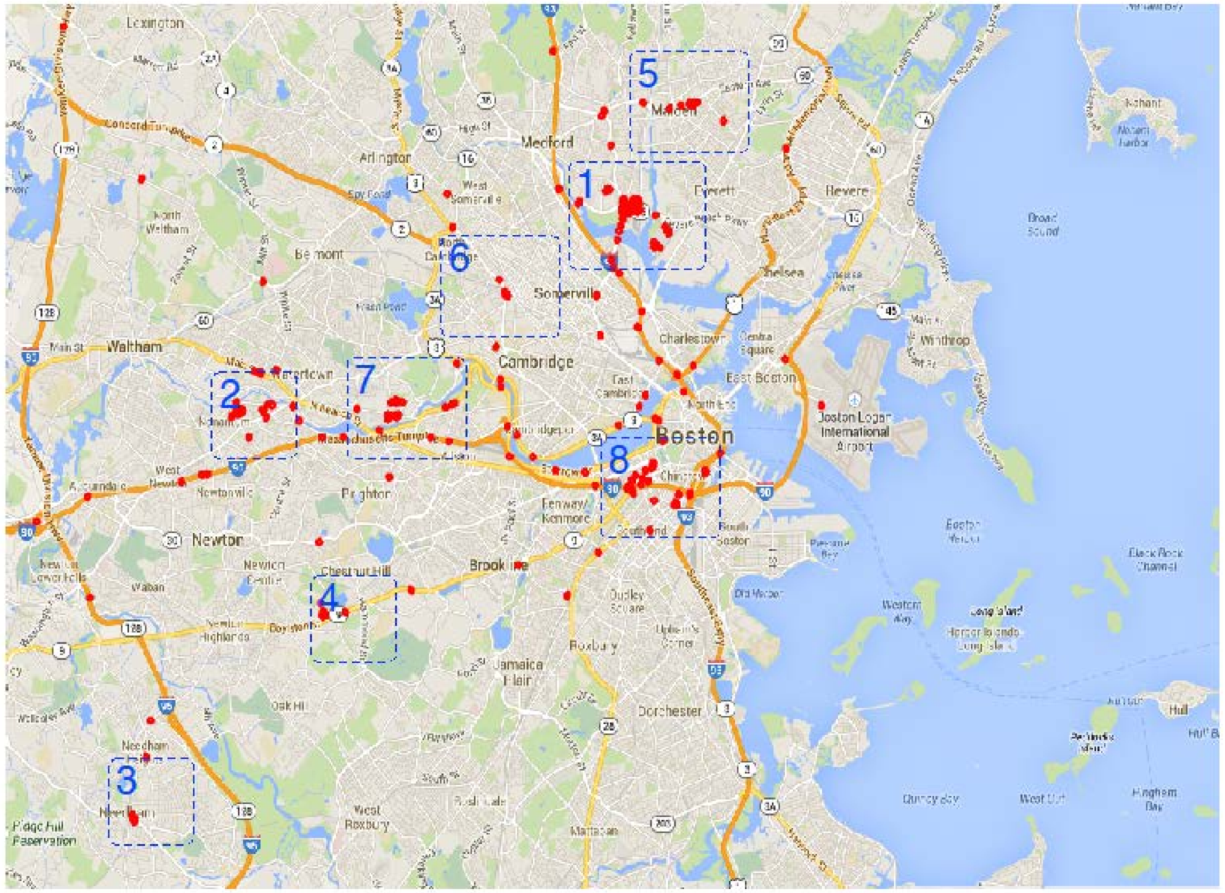
LIPS: Lifestyle Learning via Mobile Phone Sensing

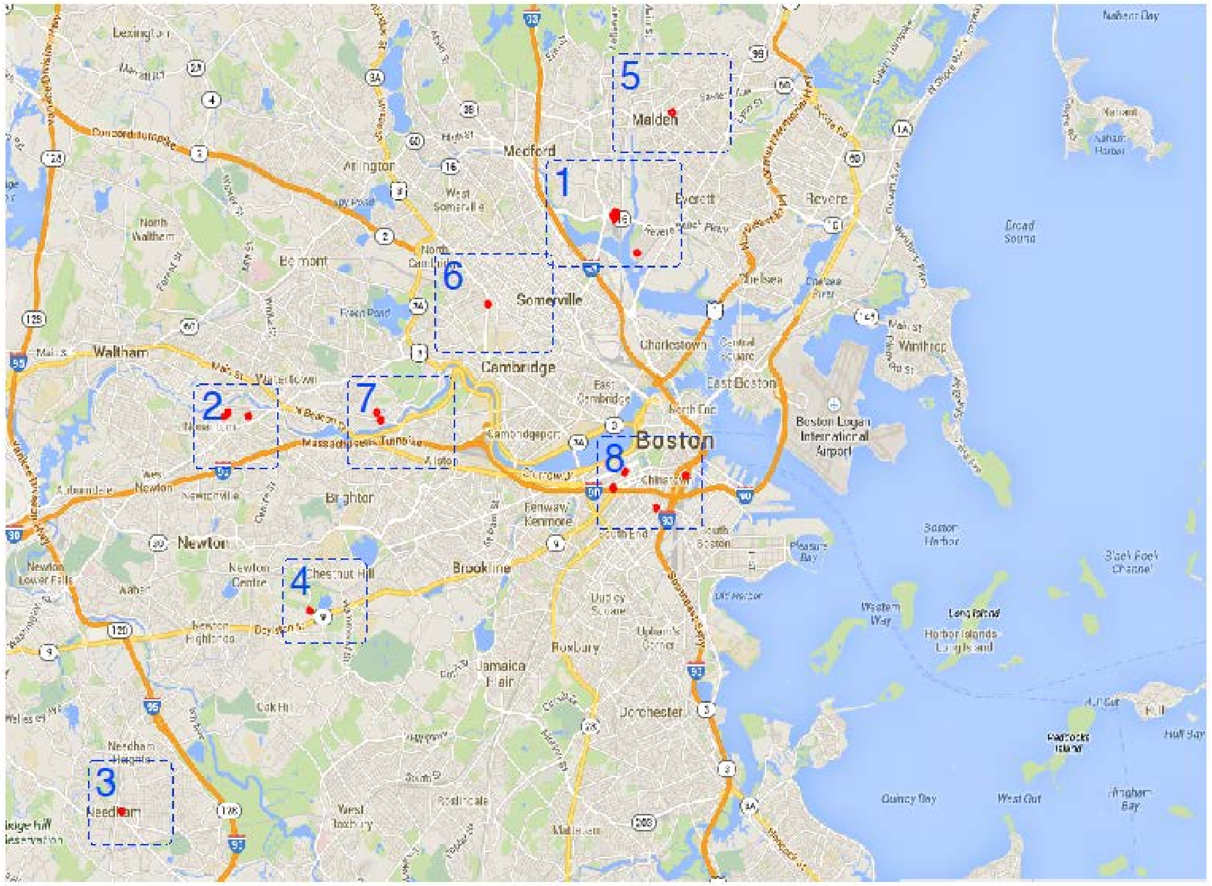
- Context = location?
- What is lifestyle? "Lifestyle is expressed in both work and leisure behavior patterns and (on an individual basis) in activities, attitudes, interests, opinions, values, and allocation of income." (from businessdictionary.com)
- Why learning lifestyle?
- Can lifestyle be learned via smartphone sensing?

LIPS: Features

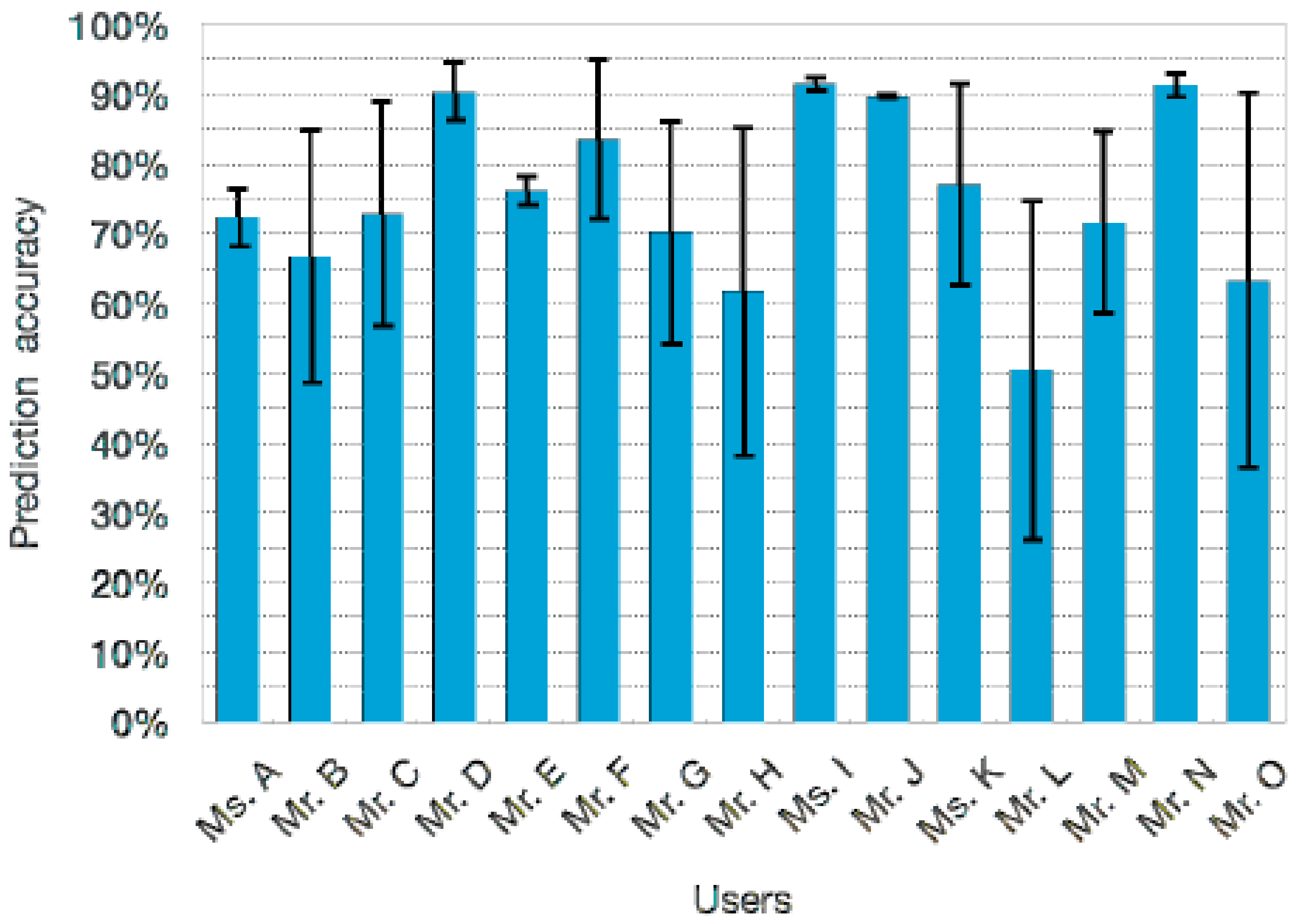
- Day and time
- Location
- Moving state: driving/still/running/walking/biking
- Weather
- Temperature
- Active/Inactive

LIPS: Lifestyle Characterization using PoIs





LIPS: Prediction Accuracy



Outline

- Vision
- System: Design and Implementation of A Unified Platform
- An App: Objective Ranking based on Mobile Phone Sensing (SOR)
- ***New Research Directions***

Future Research Directions

- New killer applications!!!
- GPS-less sensing task scheduling on a sensing server
- Energy-efficient sensing task scheduling on a mobile phone
- Mobile user reputation and truthfulness issues
- Mobile user privacy
- Processing and analyzing sensor data in the cloud

Thank You

Questions?