Mobile Crowd Sensing

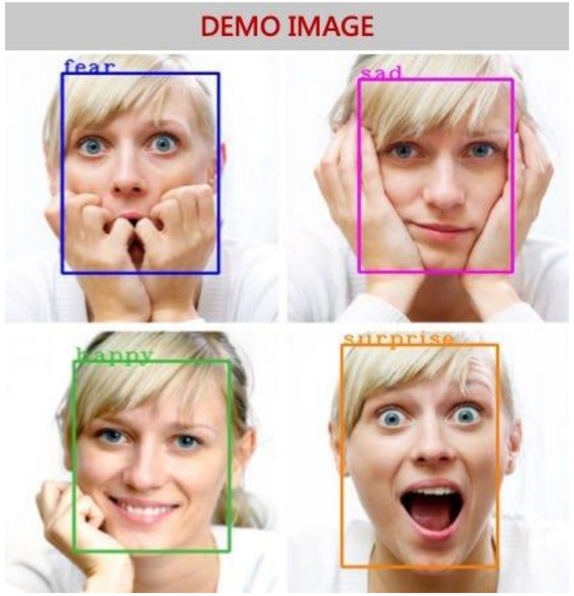
CSE 162 – Mobile Computing

Hua Huang
Department of Computer Science and Engineering
University of California, Merced

Recap: mobile personal sensing

- **Personal sensing:** phenomena for an individual.
  - Emotions from images or voices
  - Hand movement tracking
  - Activity recognition (smoking)
  - Localization
  - etc

Sensors on iPhone 4:
- Camera
- Audio
- Accelerometer
- GPS
- Gyroscope
- Compass
- Proximity
- Ambient light



Introduction

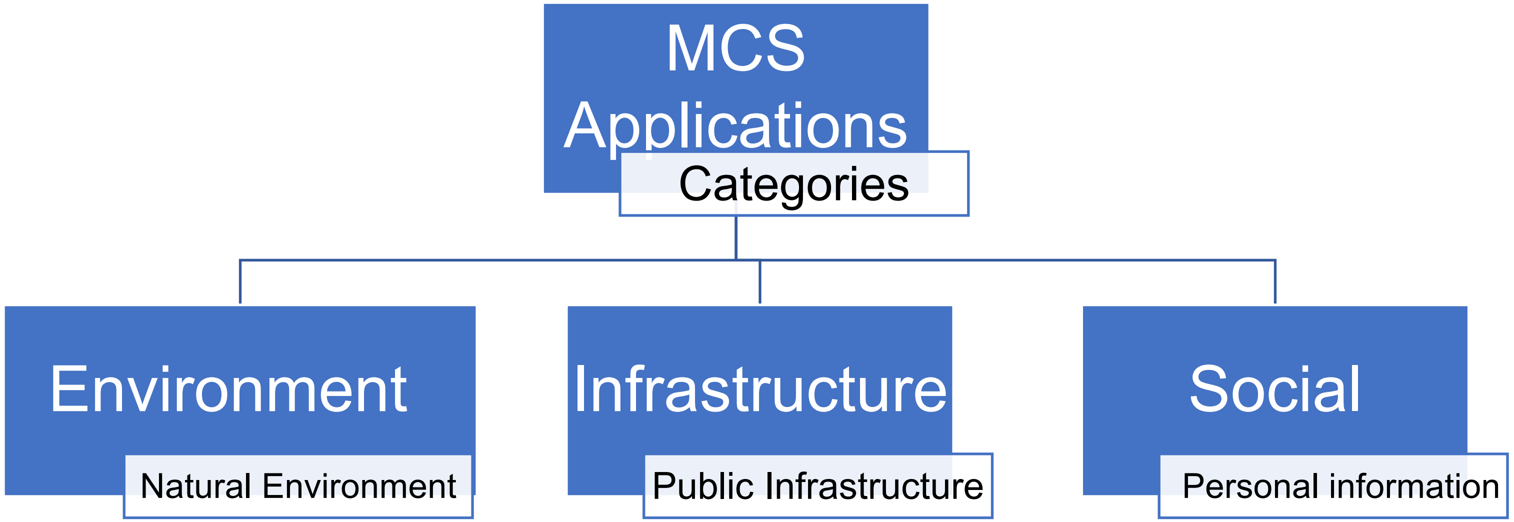
- Leveraging on crowd
- Data, services, ideas, contents, skills, money, … coming from crowds
- Crowdsourcing = Crowd + outsourcing

Introduction

- Pertains to the monitoring of large-scale phenomena that cannot be easily measured by a single individual.

- For example, intelligent transportation systems may require traffic congestion monitoring and air pollution level monitoring.
  - These phenomena can be measured accurately only when many individuals provide speed and air quality information from their daily commutes

Mobile Crowd Sensing Applications



Environmental MCS applications

- Environmental MCS applications,
  - pollution levels in a city
  - water levels in creeks
  - monitoring wildlife habitats

- Example: The **CommonSense** system uses specialized handheld air quality sensing devices that communicate with mobile phones (using Bluetooth) to measure various air pollutants (e.g. CO_2, NO_x).
  - These devices when deployed across a large population, collectively measure the air quality of a community or a large area.

Infrastructure Applications

- Infrastructure Applications involve the measurement of large scale phenomena related to public infrastructure.
  - Traffic congestion
  - road conditions
  - parking availability
  - outages of public works (e.g. malfunctioning fire hydrants, broken traffic lights)
  - real-time transit tracking.

Infrastructure MCS Example

- **CarTel** utilizes specialized devices installed in cars to measure the location and speed of cars and transmit the measured values using public WiFi hotspots to a central server
  - This central server can then be queried to provide information such as least delay routes or traffic hotspots
- **Nericell** utilizes individuals' mobile phones to not only determine average speed or traffic delays, but also detect honking levels (especially in countries like India where honking is common) and potholes on roads.

Social applications

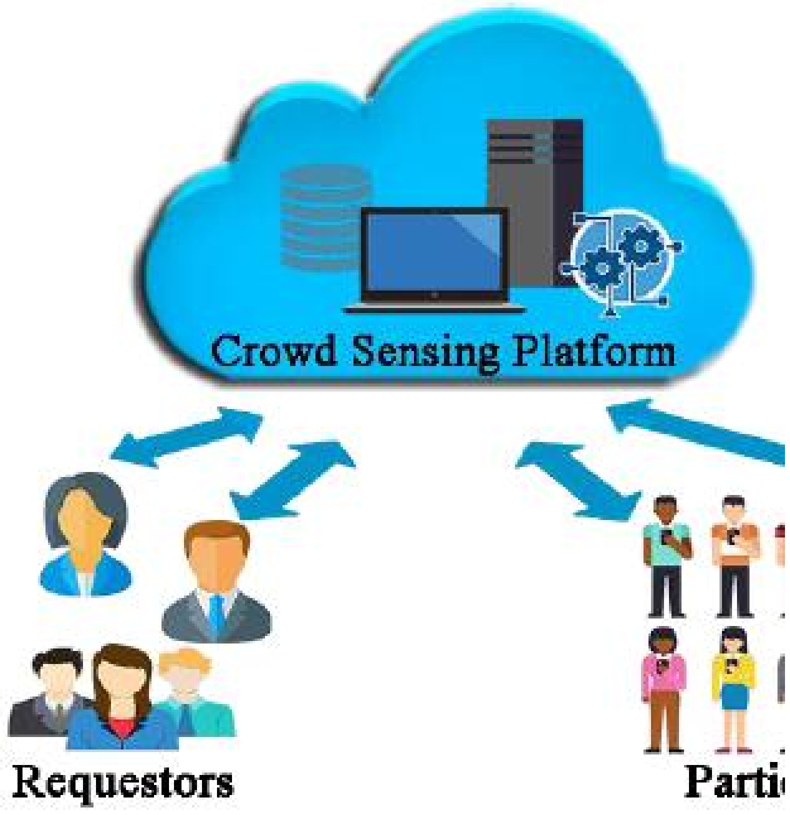
- Social applications involves individuals sharing sensed information amongst themselves.
  - Individuals can share their exercise data (e.g. how much time one exercises in a day) and compare their exercise levels with the rest of the community.
  - **BikeNet**, individuals measure location and bike route quality (e.g. CO_2 content on route, bumpiness of ride) and aggregate the data to obtain “most” bikeable routes
  - **DietSense**, individuals take pictures of what they eat and share it within a community to compare their eating habits.

MCS: Unique Characteristics

- Compared to traditional mote-class sensor networks, mobile crowdsensing has a number of unique characteristics that bring both new opportunities and problems.
  1. today’s mobile devices have significantly more computing, communication and storage resources than mote-class sensors, and they are usually equipped with multimodality sensing capabilities.
  2. millions of mobile devices are already “deployed in the field”: people carry these devices wherever they go and whatever they do. By leveraging these devices, we could potentially build large scale sensing applications efficiently
     - For example, instead of installing road-side cameras and loop detectors, we can collect traffic data and detect congestion levels using smartphones carried by drivers.

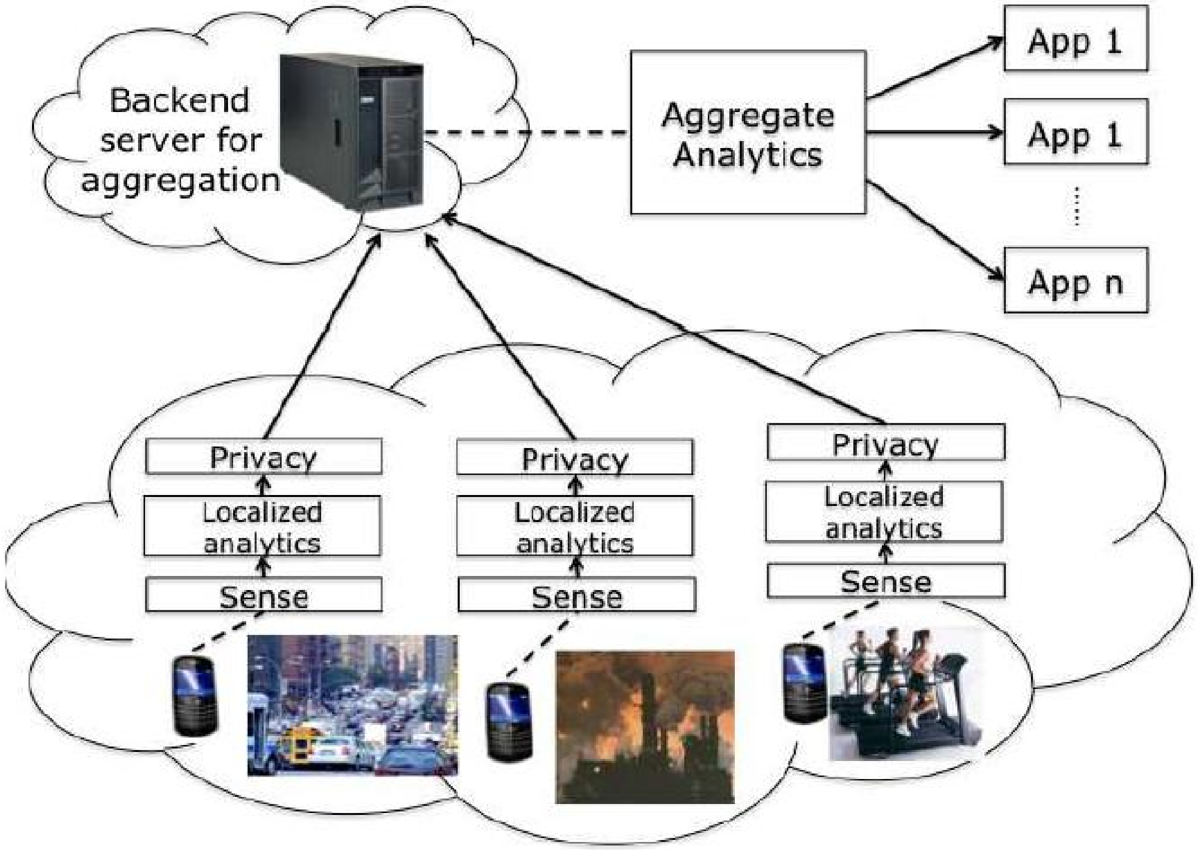
MCS Architechture

- The general architectural design of an MCS system comprises the following main entities:
  - **Requestors**, initiate targeted data collection process by publishing a task pertinent to their interests to the crowd sensing platform and analyze the collected data from the Workers.
  - **Workers** are assigned several tasks pertinent to data collection, thus are the main source of information and play a major role in data collection
  - **Crowd sensing Platform** is the main communication link between Requestors and Workers. The platform stores, processes and analyzes data provided by Workers and the Requestors.



MCS Workflow

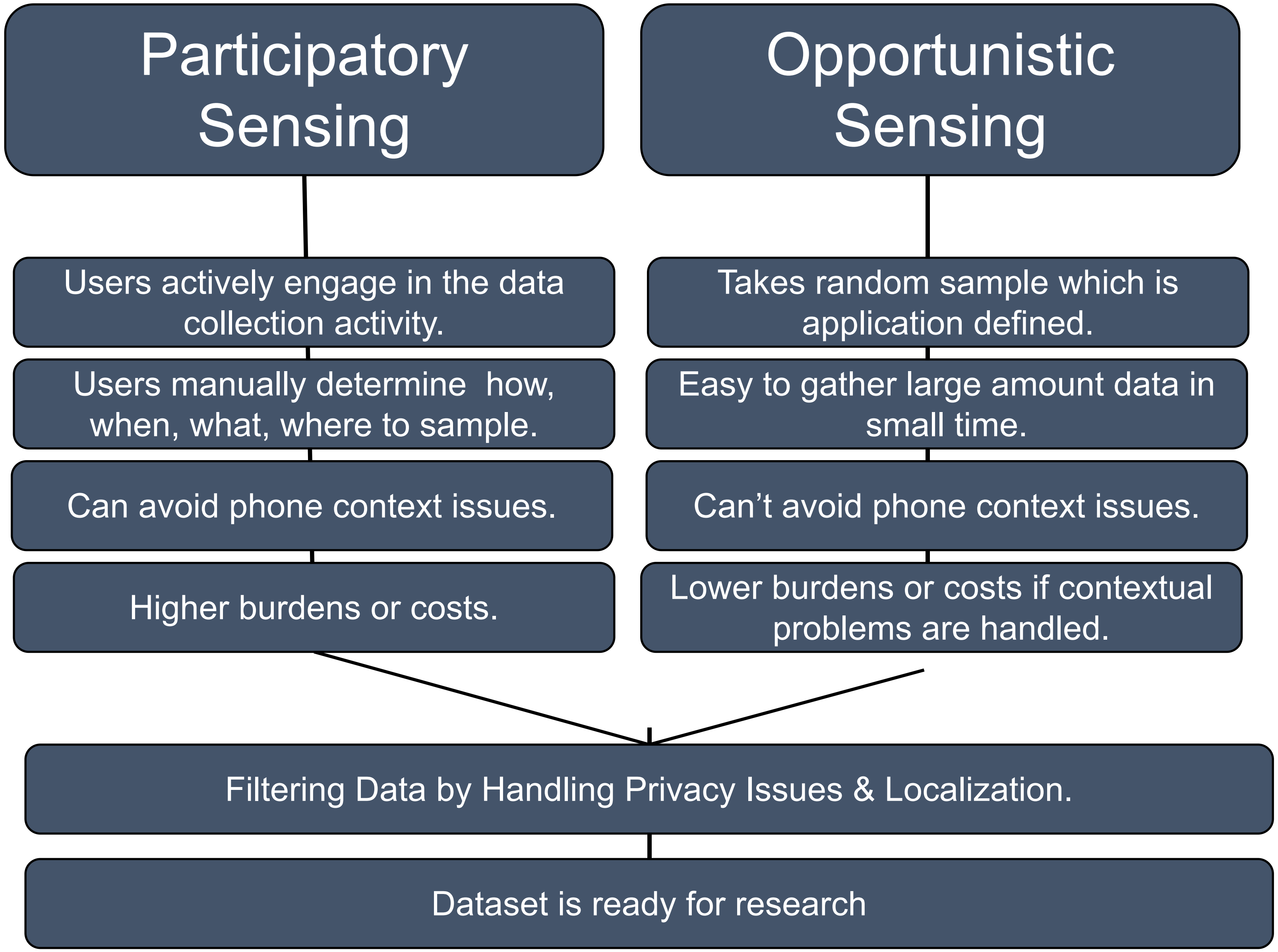
- Raw sensing data is collected on devices.
- Local analytics process it to produce consumable data for applications.
- After privacy preservation, the data is sent to the backend.
- Aggregate analytics will further process it for different applications.



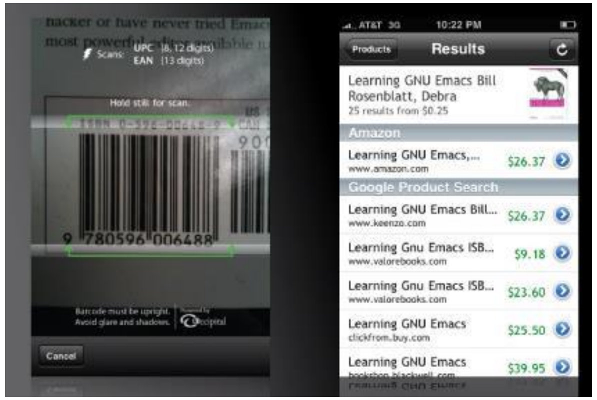
Participatory and opportunistic sensing

Participatory and opportunistic sensing

- Participatory sensing requires the **active involvement** of individuals to contribute sensor data (e.g. taking a picture, reporting a road closure) related to a large-scale phenomena.
- Opportunistic sensing is more autonomous and user involvement is minimal (e.g. continuous location sampling without the explicit action from the user).


Participatory Sensing
Opportunistic Sensing
Users actively engage in the data collection activity.
Users manually determine how, when, what, where to sample.
Can avoid phone context issues.
Higher burdens or costs.
Takes random sample which is application defined.
Easy to gather large amount data in small time.
Can’t avoid phone context issues.
Lower burdens or costs if contextual problems are handled.
Filtering Data by Handling Privacy Issues & Localization.
Dataset is ready for research

Participatory crowdsensing examples



- **LiveCompare**
  - User-created database of UPCs and prices
  - GPS and cell tower info used to find nearby stores
- **PetrolWatch**
  - Uses phone to photograph gasoline price
  - Uses GPS to know when gas station is near

Opportunistic crowdsensing examples

- **Pothole Monitor**
  - Combines GPS and accelerometer

The Challenges of MCS…

Incentivize Users

Resource Limitations

Privacy

Aggregate Analytics

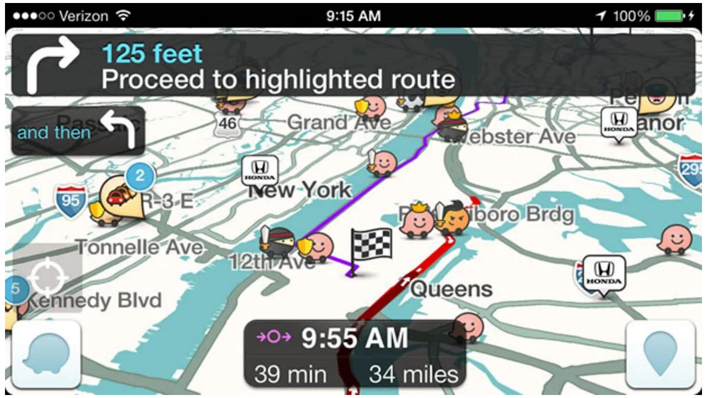
Challenge: Incentivize users

- Three basic approaches:
  - To make **entertainment an incentive**, crowd sensing tasks are turned into sensing games, such that users can contribute computation or sensing abilities of their mobile devices when they play these games. This paradigm makes users feel enjoyable when they perform tasks
  - The rationality of taking **service as an incentive** roots in the mutual-benefit principle. Service consumers are also service providers. In other words, if a user wants to benefit from the service provided by the system, she also has to contribute to the system.
  - The last category is based on **monetary incentives**. In this case, the system has to pay a certain amount of money to motivate potential participants, such that the participants can use their resources, usually smartphone sensors, to complete the distributed tasks.

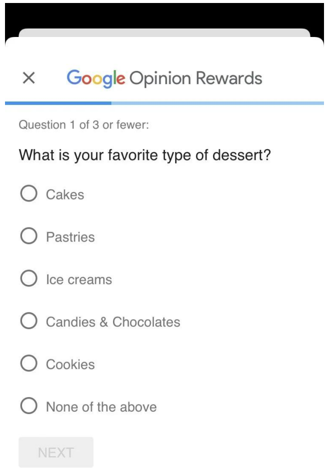
examples: Incentivize users

- Entertainment as an incentive
  - Pokeman Go

- Service as incentive
  - Traffic monitoring



- Monetary incentive
  - measure web content usage by paying users money based on page visits to a site.



Entertainment as an incentive

- Example: mobile game **Treasure** to build WiFi coverage maps of a given game area.
- Players carry mobile devices with GPS and WiFi. They need to pick up virtual coins scattered over the game area and then upload the coins to a server to gain game points. Better network connections give larger probabilities of uploading the collected coins successfully.
- Therefore players are motivated to find areas with stronger WiFi coverage.

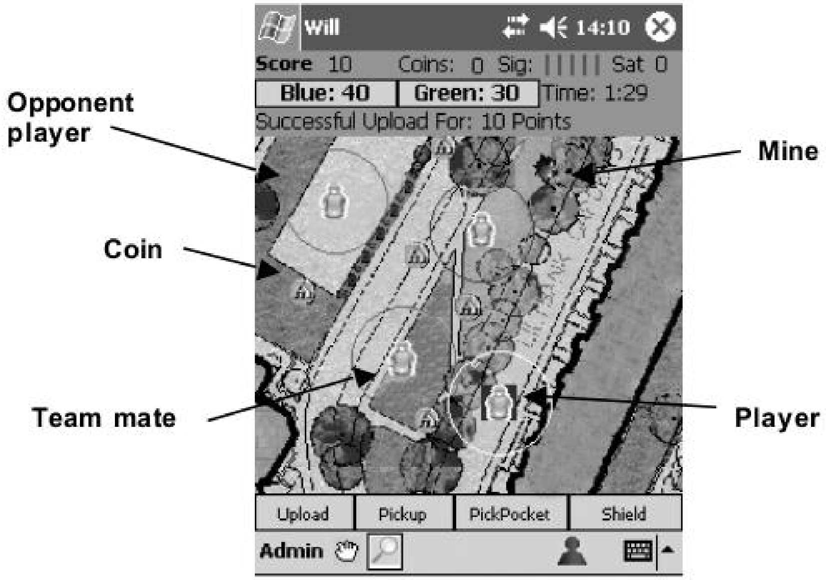
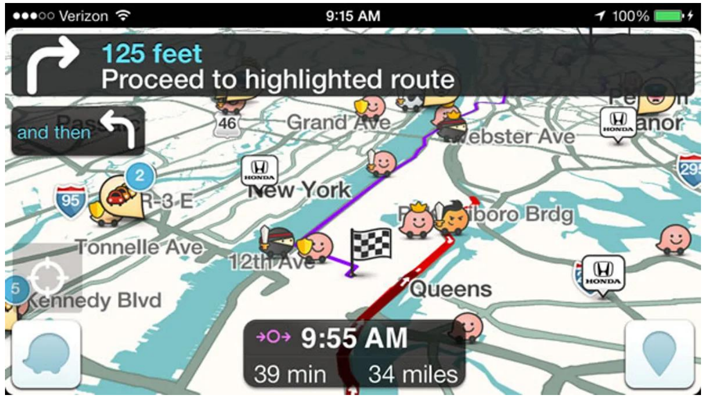


Fig. 1. The game interface. The map shows players' locations, along with coins that are often positioned outside the network. A semi–transparent map layer of green, yellow and red squares builds up as players move around, revealing network coverage, however for printing reasons these have not been included in this figure.

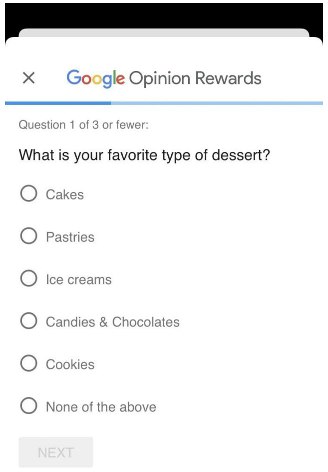
Service as incentive



- For some mobile crowd sensing systems, a participant (e.g., a smartphone user) may have two roles concurrently: a contributor and a consumer.
- Traffic monitoring is a typical example.
  - A participant acts as a contributor when she travels on a bus or car if she collects traffic data (e.g., GPS traces) to a service provider via networks such as WiFi, GPRS, and 3G.
  - The service provider then processes the data crowdsourced from a large amount of users and provides a real-time traffic information service, such as querying on traffic jams and bus crowdedness
  - In such sensing applications, to attract more users to contribute sensed data such that the system can provide services of good quality, the service provider will usually grant a participant some service quota, which determines how much service that user can receive.
  - In essence, this strategy is an exchange of contribution and consumption for each participant.

Monetary incentive

- Paying for sensed data in crowd sensing tasks is the most intuitive incentive, as it has made sensed data become goods in a free market. Any user who would like to make some money can sell her sensed data for crowd sensing tasks.

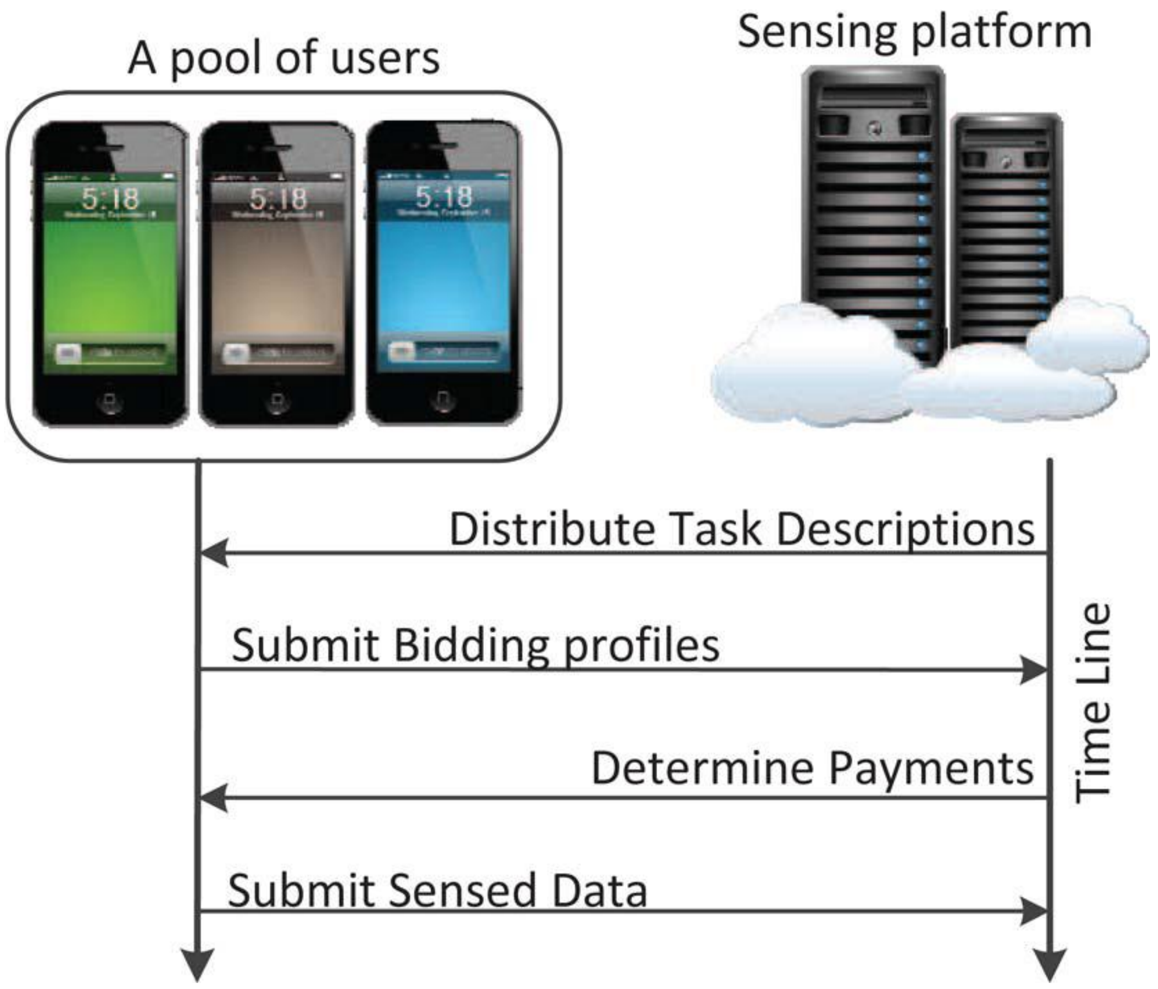


Reverse Auction Mechanism

- The system involves two participating roles: a platform that distributes sensing tasks and the mobile phone users who constitute potential labor force.
- The objective is to design a task assignment and payment negotiation scheme, which ensures that both the platform and users are satisfied

Reverse Auction Mechanism

- The platform initiates one round of task distribution by sending task descriptions.
- A set of n users are assumed to be interested in the sensing tasks after receiving the requests.
- If users participate in sensing tasks, they will consume multiple resources, including computation, communication, and energy. Thus it is rational for a user to expect certain profit based on her cost and sensing plan (e.g., sensing time).
- A participating user then submits a bidding profile (including a bidding price and a sensing plan) to the platform.
- After collecting all bidding profiles from the n users, the platform selects a subset of them and determines the payments for them.
- Finally, the selected users perform the assigned tasks and upload the sensed data to the platform.



The Challenges of MCS…

Incentivize Users

Resource Limitations

Privacy

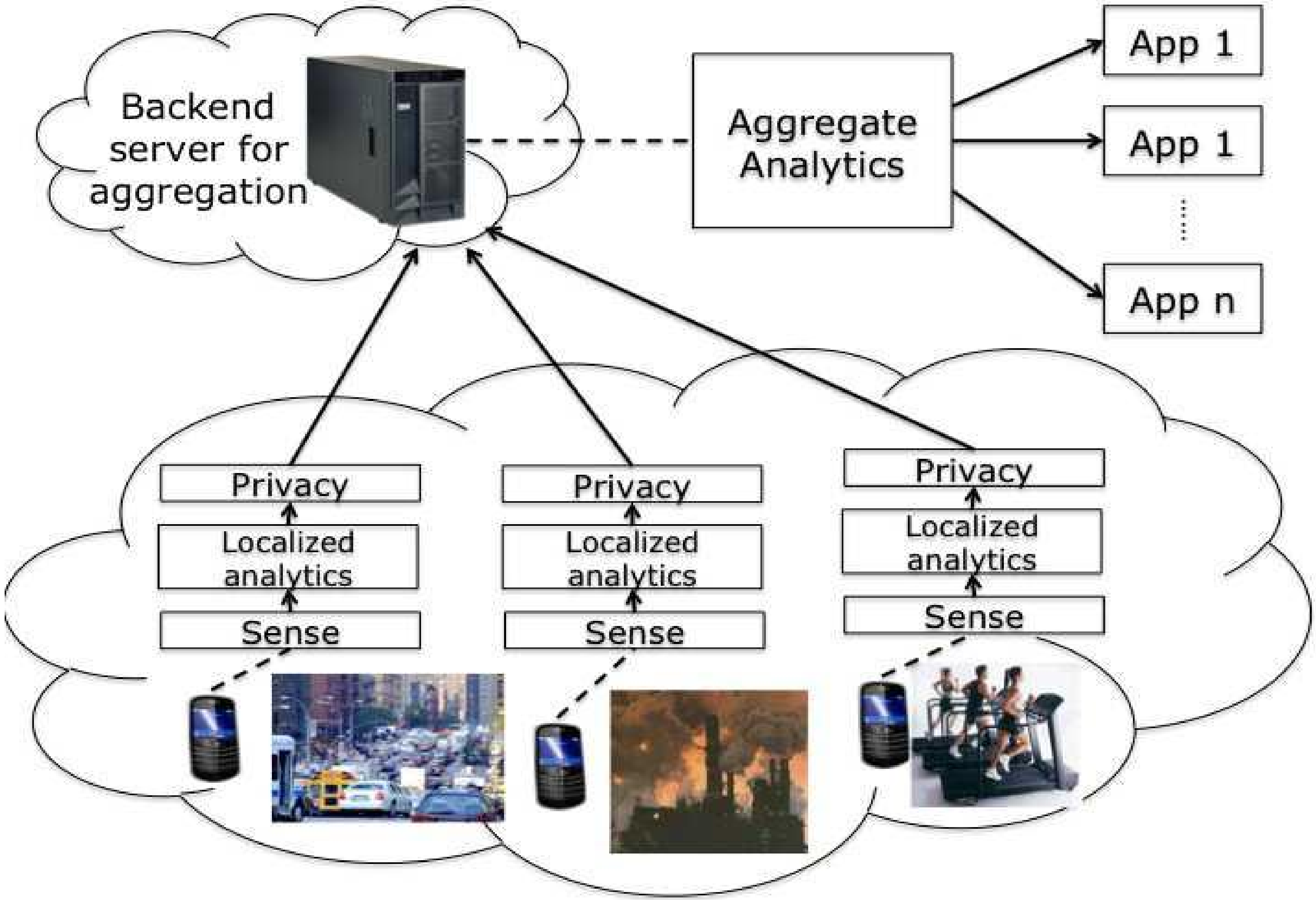
Aggregate Analytics

Challenge: Resource limitations

- How to run the MCS application without significantly affecting the normal device operation?

- How to avoid overloading the server?
  - Single point failure problem

LOCALIZED ANALYTICS



LOCALIZED ANALYTICS

- Upload less amount and appropriately processed data
  - *Data mediation*, such as filtering of outliers, elimination of noise, or makeup for data gaps.

- Reduces the amount of processing that the backend has to perform

MCS: Localized Analytics

- Benefits
  - Transmitting data summary consumes lesser energy and bandwidth than transmitting the raw sensor readings.
  - Do the processing locally saves time.
    - Transmitting raw sensor data on intermittently connected channels can be time consuming.
    - *Example: Context inference*
      - *Transportation mode*
      - Kinetic modes of humans
      - Social settings

The Challenges of MCS…

Incentivize Users

Resource Limitations

Privacy

Aggregate Analytics

Privacy

- Potentially collect sensitive sensor data pertaining to individuals
  - the routes they take in daily commutes
  - their home and work locations

Privacy Preservation techniques

- Anonymization
- Cryptographic techniques
- Data Perturbation

Privacy Preserving Approach: Anonymization

- Remove any identifying information from the sensor data before sharing it with a third party.
  - Removes any identifying information from the sensor data before sharing it
  - Drawback: anonymized GPS (or location) sensor measurements can still be used to infer the frequently visited locations of an individual and derive their personal details.

Privacy Preserving Approach: cryptographic techniques

- Cryptographic techniques are used to transform the data to preserve the privacy of an individual.
  - Example: privacy leakage from text auto-completion
  - Example: suggested location of interest
  - are compute intensive and are not scalable because they require the generation and maintenance of multiple keys, which also leads to higher energy consumption.

Conformal encryption.
Talk about it.

Privacy Preserving Approach: Data Perturbation

- Add noise to sensor data before sharing it with the community to preserve privacy of an individual
  - Requirement on the noise: Preserve the privacy of an individual, while not influencing the statistics accuracy
- For example, in a weight watchers application, it is important to compute the average weight of the population.
  - Add a random number to the weight
  - When these values are averaged, the randomized component vanishes and the average weight of the community remain accurate

The Challenges of MCS…

Incentivize Users

Resource Limitations

Privacy

Aggregate Analytics

Challenge: Aggregate Analytics

- MCS applications rely on analyzing the data from a collection of mobile devices, identifying spatio-temporal patterns.
- There are two possible approaches for data mining.
  - Offline approach
  - Online approach
- Consider the restaurant recommendation app based on user ratings

Offline Aggregate Analytics

- One is a traditional approach where data is stored in a database first and then one can apply various mining algorithms against the database to detect patterns.
- However, if the application requires fast detection of patterns, this approach will not work.
- e.g., how to quickly recommend the best route considering traffic?

Online Aggregate Analytics

- Stream data mining algorithms take as input continuous data streams and identify patterns, without the need to first store the data.
- These algorithms are less accurate due to incomplete data.
- e.g., do we recommend a lunch restaurant based on the breakfast traffice data?