Mobile Sensing Fundamentals

Yingying (Jennifer) Chen
Rutgers University

Learning Objectives

- What is mobile sensing? Which steps are involved in mobile sensing applications?
- Which dimensions can be used to categorize mobile sensing applications?
- How ground truth can be collected and how its validity can be assessed?
- How sensing performance can be assessed?

What is Mobile Sensing?

- **Sensing:** perceiving/becoming aware of external or internal stimulus
- **Mobile sensing:** the use of sensors embedded on mobile devices for sensing stimulus
- What can be sensed?
  - Environmental characteristics: where is the user? Who are around the user?
  - User characteristics: what mood is the user in? What personality type is the user? What does the user like?
  - Device characteristics: remaining battery, current network connectivity, etc.

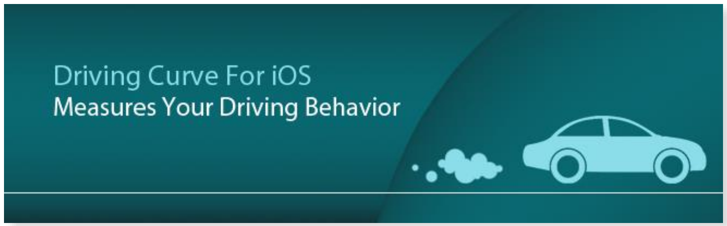
Types of Sensors

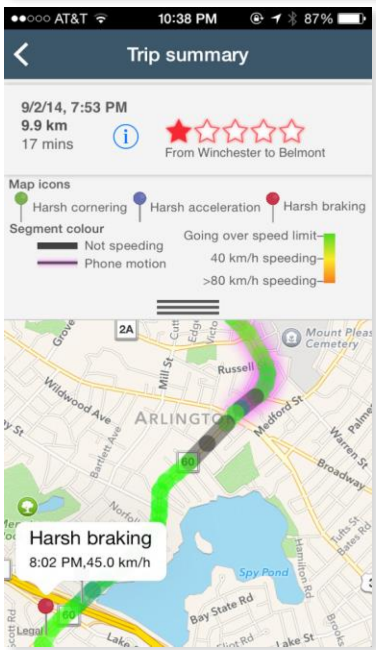
- Inertial sensors: accelerometer and gyroscope
  - Landscape/portrait mode detection
  - Health/fitness tracking
- Environment sensors
  - Temperature, humidity, barometer (air pressure)
  - Proximity and ambient light sensors (used, e.g., for adapting screen and activate energy saving policies)
- Physiological sensors
  - Heart beat, blood pressure

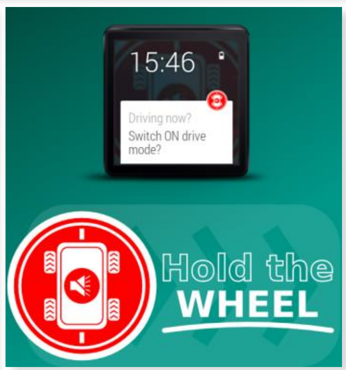
Types of Sensors (Cont.)

- "Non-sensor sensors"
  - Microphone: prosodic analysis, ambiance sensing
  - Cameras (front and/or back)
- Location-related sensors:
  - GPS (and Glonass)
  - GSM/Cellular sensor (base station, signal strength)
  - WiFi (access points and signal strengths)
- Behavior sensors
- Application usage behavior, extent of calls made
- ...and many others

Examples: Transportation Behavior Modeling







- Transportation behavior monitoring
  - Persuasive mobility applications
  - Mobility tracking applications
- Driving-related application
  - Monitoring gas consumption, driving skill, driver drowsiness etc.
- Traffic flow and congestion estimation
- Sensors:
  - Accelerometer and gyroscope most common, followed by GPS and radio environment (GSM, WiFi)
  - Other sensors considered include magnetometer, cameras, and microphones

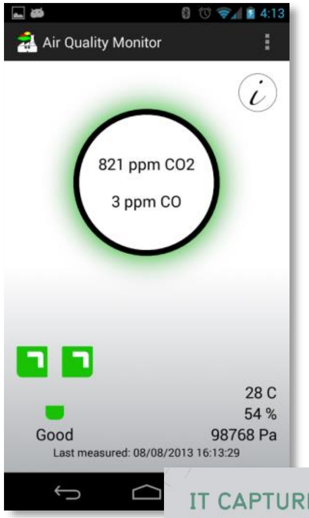
Examples: Driving Behavior Correction App

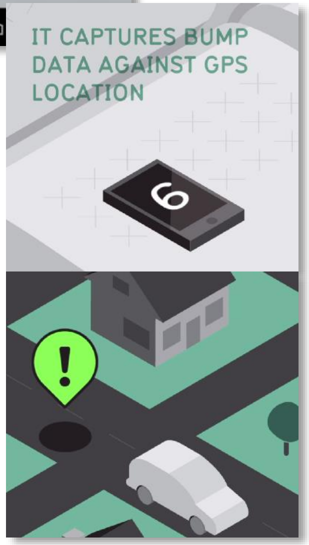
- Dash:
  https://www.youtube.com/watch?v=PdEY3LcTNEM

- Drive Score:
  https://www.youtube.com/watch?v=NZspgGTnK9c

Examples: Participatory Sensing

- Mapping phenomena in the urban environment
  - User reporting based applications, e.g., measuring CO2 at different locations or taking infrared images of buildings for structure analysis
- Road condition estimation
  - E.g., Pothole Patrol and Nericell, automatically detect and report problems in roads
- Sensors:
  - Location information (GPS)
  - Inertial and magnetic sensors (accelerometer, gyroscope, magnetometer)
  - Audio (e.g., Nericell detects honking intensity)





Other examples: SoundSense

- https://www.youtube.com/v/VK9AE_7dhc4

Anatomy of Mobile Sensing Application: Stages

Three common components can be identified from the examples:

- **Sensing:** each application relied on one or more sensor
- **Intelligence:** some information extracted from sensor data
- **Interface level**
  - Adaptation: Interface modified based on extracted information (e.g., rotating screen)
  - Sharing: providing information to friends/other users
  - Persuasion: trying to influence user behavior based on extracted information

Often one or more parts rely on information in external sources

Anatomy of Mobile Sensing Applications: Modes

Sensing applications can operate in two modes

- **On-demand:** sensing performed when user interacts with an application/device
  - Rotate the screen when the user turns the phone,
  - Answer the phone when performing a specific gesture
  - Location-based services (**LBS**): facilitate applications based on the user's location
- **Continuous:** sensor data collected and processed continually
  - Requires support for multitasking and background applications (not possible on all operating systems)
  - Resource-efficiency critical requirement, discussed in a separate lecture

Anatomy of Mobile Sensing Applications: Scale and Paradigms

- Two sensing scales:
  - Personal sensing: focus on individual user
  - Community sensing: focus on a group of users that share a common goal (and trusted relationship)
- Two sensing paradigms (based on user involvement):
  - Participatory sensing: users actively contribute data
    - Common in community sensing applications
  - Opportunistic sensing: sensor data collected (and processed) without user involvement
    - Most common mode for personal sensing application

Distributing Sensing Apps: Application Stores

- Different app stores are main distribution channel for mobile sensing applications
  - Apple App Store, Google Play, Microsoft Mobile Marketplace, etc.
- Several challenges:
  - Biases in data (population biases, context/usage biases, device biases)
  - Reliability of measurements collected from applications difficult to assess
  - Applications need to be “attractive” to users (good usability, scalability, energy-efficiency, ...)

Sensing Pipeline

- Sensor data processing steps can be divided into **a pipeline**
  - **Sensor data extraction**
    - Typically extracted in frames, e.g., all measurements within a second window (common window sizes within range 0.5 – 4 seconds)
  - **Preprocessing**
    - Removing errors from data, filtering
    - Typical operations: low/high-pass filtering, jitter removal
  - **Feature extraction**
  - **Classification / Regression**
    - Extracting the relevant information from the sensor data

Developing Sensing Pipeline

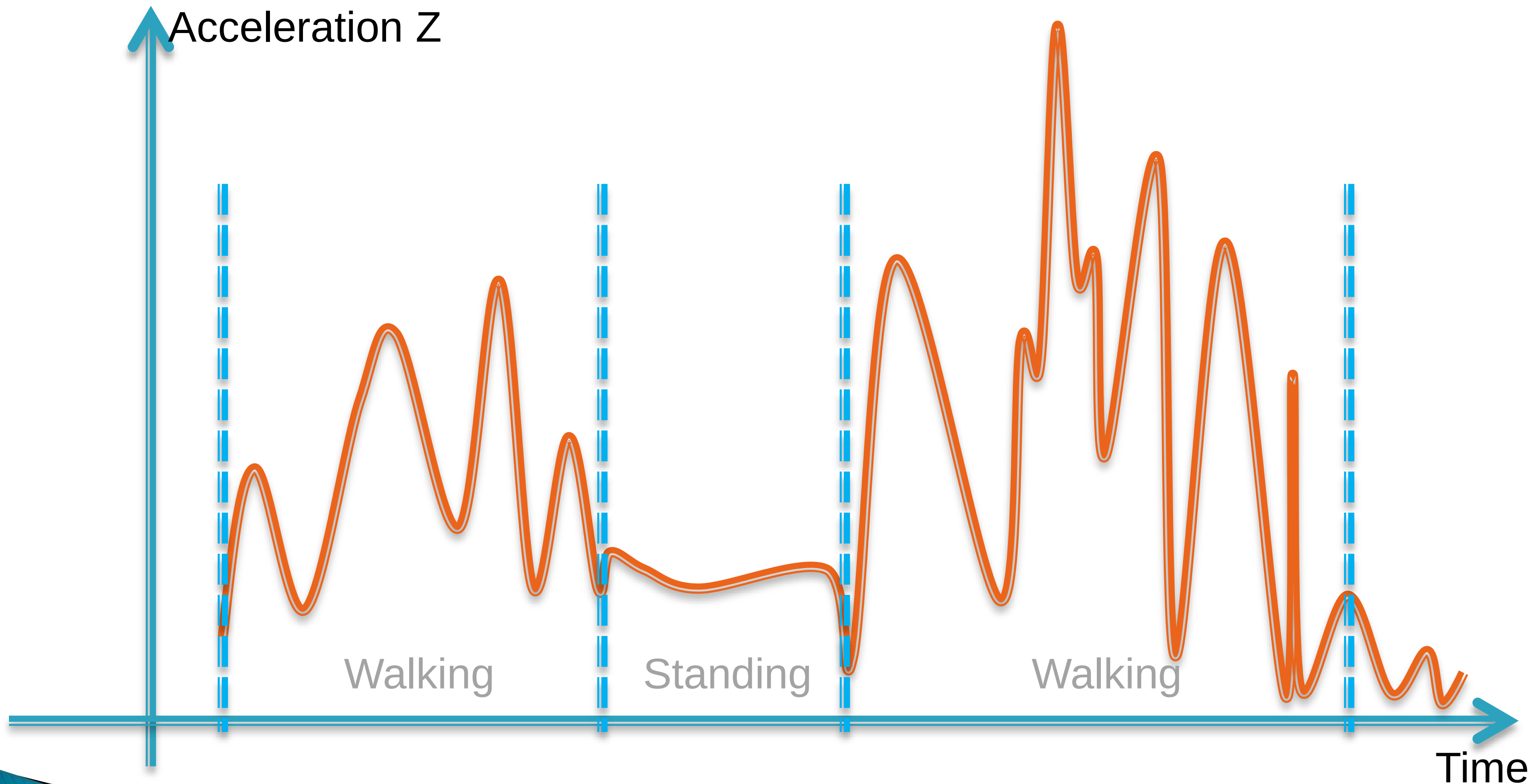
Developing sensing pipeline is a **multi-stage process**

1. Determine phenomena to sense and sensors to use
2. Collect sensor data of the phenomena together with ground truth labels
3. Determine appropriate preprocessing steps and features that are good for the phenomena
4. Train classifiers / regression models based on the features and ground truth
5. Deploy trained classifiers as part of the application
6. Repeat steps 2 – 5 as needed

Obtaining Ground Truth

- Developing sensing applications requires **ground truth** of the phenomena that is being sensed
  - Examples: current transportation mode, number of steps taken, personality type of a user
- Typically the ground truth is given as **a discrete category/class** or as **a numeric value**
  - Discrete class: transportation mode (walking, running, train, …) → classification problem
  - Numeric value: questionnaire-based measures, e.g., extroversion (value between 1 – 48) → regression
  - Labels assigned to frames (not individual measurements)

Illustration of the Importance of the Ground Truth



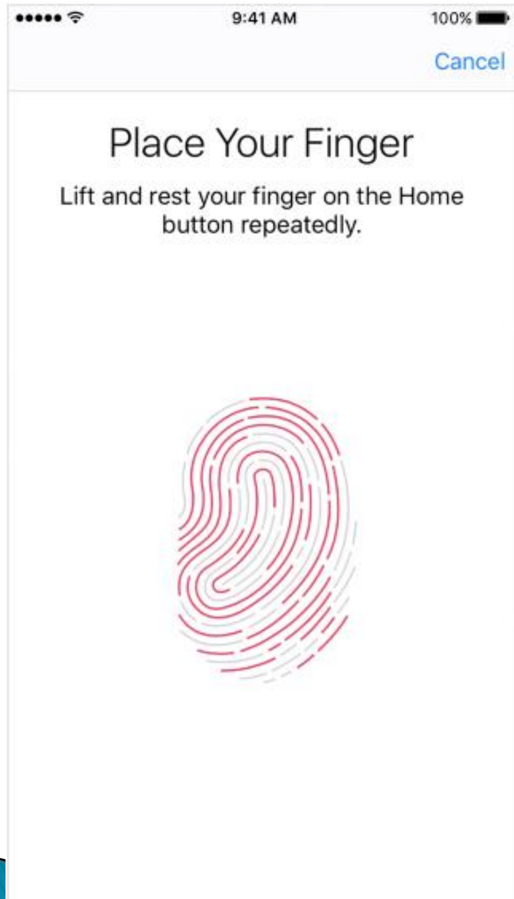
Obtaining Ground Truth: Labeling / Annotation

- When labels are classes/categories, most common way is to **manually assign** the corresponding labels
  - Self-annotations: user collecting the data marks the appropriate class / category
  - Diary: periodic labeling
  - Experience sampling: prompting the user to label
  - Shadowing: external person observes the user and assigns the appropriate label
  - Video-annotation: activity recorded and an external person assigns labels based on the recording
- Typically “event-based” annotation
  - Only starting/ending points of activities are labeled

Error Sources in Reporting

- Self-reporting biases:
  - Recall: Forgetting to annotate some event or remembering something wrong
  - Self-bias: Intentionally manipulating labels, e.g., leaving out information or modifying information
- Participant observation (shadowing) biases:
  - Behavior bias: user modifies his/her behavior due to awareness of being observed
- Coding errors
  - Ambiguity and non-uniqueness of labels/actions (e.g., walking in a train)
- Latency
  - Start/end of activity starts/ends too early/late

Example: Authentication (Apple Touch ID)



- Example of experience sampling-based labeling
  - Prompts the user and asks for appropriate label at pre-defined intervals
  - Binary labels: Yes / No
- In this case, no logging of motion sensors → user interactions have no effect on collected data
- Hence **prompting** the user valid approach

Example: Transportation Mode Detection

- Discrete set of different classes (9+1)
  - Separate NULL class for data that does not conform to any of the activities of interest
- Assigning labels during the activity (not desirable)
  - Requires interactions with the device, which influence collected data
- Alternatives
  - Assigning labels post activity
  - Separate person (and device) responsible for labeling

Numeric Labels: Psychometric Instruments

- Psychometrics measures psychological aspects of the user (personality, mood, emotion)
  - One of the most important categories of applications where labels are numeric (instead of categorical)
- Ground truth collected using standardized questionnaires
  - Consist of several items or questions
  - Responses typically elicited using Likert-scales (e.g., 1-5, 1-7 or 1-9)
  - Typically multiple items grouped together to form a scale
    - Example:Big-5 (personality) questionnaire has 5 scales measuring different personality traits

Example: EPQ-BV extroversion scale

Eysenck's Personality Questionnaire - Brief Version

	1	2	3	4	5
Are you a talkative person?					
Are you rather lively?					
Do you enjoy meeting new people?					
Can you usually let yourself go and enjoy yourself at a lively party?					
Do you usually take the initiative in making new friends?					
Can you easily get some life into a dull party?					
Do you tend to keep in the background on social occasions? (Reversed)					
Do you like mixing with people?					
Do you like to plenty of action and excitement around you?					
Are you mostly quiet when you are with other people? (Reversed)					
Do other people think of you as being very lively?					
Can you get a party going?					

Evaluation Criteria

- Performance of sensing algorithms:
  - Standard classification/regression metrics (precision, recall, F1-score, mean absolute error)
  - Event-based metrics (underfill and overfill rate, fragmentation rate)
- End-user performance criteria(covered in a later lecture):
  - Energy-efficiency / battery drain
  - Usability

Classification Metrics

	Class Positive	Class Negative
Test positive	True positive (TP)	False positive (FP)
Test negative	False negative (FN)	True negative (TN)

- Most common approach for evaluation sensing algorithms
- Consider the matrix on the left:
  - **Precision** = TP / (TP + FP)
  - **Recall** = TP/(TP+FN)
  - **F1-score** = 2TP/(2TP+FP+FN)
    - Harmonic mean of precision and recall
    - Equal to 2*(precision*recall)/(precision + recall)

Classification Metrics and Frame-Based Scoring

- Precision and recall examples of frame based metrics
- Classifier applied separately for each frame and metrics calculated by comparing output to ground truth
- Scoring: for each class and sample, determine whether output is true positive (FP), false positive (FP), true negative (TN), or false negative (FN)
- Precision/recall/F1-score calculated from scoring
- Example:

Ground truth	W	W	W	W	W	W	S
Classifier I	W	S	W	S	W	S	S
Walking (W)	TP	FN	TP	FN	TP	FN	TN
Standing (S)	TN	FP	TN	FP	TN	FP	TP

Confusion Matrix

- Often classification metrics are combined with a confusion matrix
- Table that shows what kinds of misclassifications occur
- Can be used for calculating precision, recall and F1-score
- Example:
  - Precision(Still) = 35654 / (35654 + 14 + 564 + 976 + 1500 + 3187) = 0.85
  - Recall(Still) = 35654 / (35654 + 58 + 218 + 235 + 433 + 1784) = 0.929

	Predictions								
	Still	Walking	Bus	Train	Metro	Tram	Precision	Recall	F_1-score
Still	35,654	58	218	235	433	1,784	85.1	92.9	88.8
Walking	14	12,823	371	10	11	61	96.2	96.5	96.3
Bus	564	318	3,811	207	356	1,540	58.9	56.1	57.4
Train	976	50	238	319	423	540	23.4	12.5	16.3
Metro	1,500	11	444	347	823	621	34.6	22.0	26.9
Tram	3,187	70	1,393	244	332	2,762	37.8	34.6	36.1

Other (sensing) metrics

- Regression metrics (continuous/numeric targets)
  - Mean absolute error (MAE): Σ_i | $predicted_i - value_i$ |
- Latency
  - Difference in the start time of event and the first true positive of classifier
  - Effectively the extent of underfills
- Energy and computational cost
  - Resource requirements of the algorithms
  - Computational overhead

Summary

- **Mobile Sensing:**
  - The use of mobile devices for acquiring information about phenomena (user/device/environment)
  - Sensing pipeline, scale, and paradigms
- **Ground truth collection:** several options with different pros and cons
- **Several evaluation criteria:**
  - Sensing performance: classification metrics and event-based metrics
  - Usability, end-user factors, energy-efficiency