RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

# MINING DATA FROM MOBILE DEVICES

Introduction

*Spiros Papadimitriou*, Tina Eliassi-Rad





## Tutorial plan

Introduction (Spiros)
- Overview
- Mobile technology overview

Algorithms & Applications I (Spiros)
- Mobile sensing
- Urban sensing
- Healthcare

Break (30 minutes)

Algorithms & Applications II (Tina)
- Location
- Context
- Ads
- Privacy

Conclusions (Spiros)



## Mobile devices



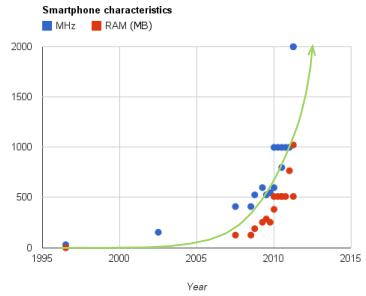

## Mobile devices: smartphones

Embedded sensors:
- GPS & compass
- Accelerometer & gyro
- Proximity
- Camera
- Speech recognition
- (Humidity, Temperature, Barometer/altimeter)
- …

(more later)



## So what?




…you have a pretty powerful computer in your pocket!

…and it's connected!



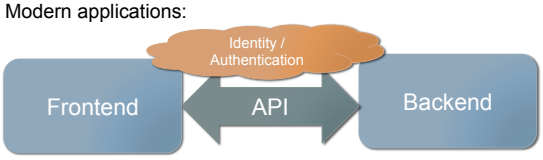
## So what?

> 'It's what I and many others have worked towards our entire careers. It's just happening *now*.'
> – Eric Schmidt (on cloud computing)

- The same could be said about mobile sensing and mining
  - Sensing & sensor networks
  - Ubiquitous computing
  - Mobility tracking
  - …
- But all are becoming mainstream now!

# Mobile "vs" web

Modern applications:




- Browser (e.g., HTML5 + Javascript)
- Smartphone
- …

- Typically JSON (increasincly, authenticated)

- Amazon EC2
- Microsoft Azure
- Google GCE
- …



# Mobile "vs" web

Identity / authentication (e.g., OAuth):

- Users
- Applications

- Mobile APIs for managing identity/accounts & content



# Mobile "vs" web

E.g.: what is the difference between Facebook in your web-browser, vs Facebook on your smartphone

Not much:

- It's the same backend & API, just running a different frontend

A lot:

- Access to content and data only on the device (e.g., photos, location, accelerometer, etc…)



# Example applications

- Geo-location
- Urban computing
- Quantified self
- Healthcare
- Security
- …many more!

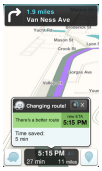


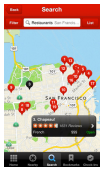
# Geo-location

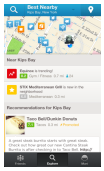
## Example applications

What most people think (mainstream applications):


Google Maps


Waze


Yelp


Foursquare

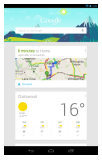
- Maps
- Navigation
- Local search (+ social)



# Geo-location

## Example applications


Locale


Google Now


Glympse

- Context-based:
  - Locale: e.g., "if I'm within 0.5mi of work address and I have a meeting on my calendar, then set my phone to silent"
  - Google Now: "if I have a dentist appointment on my calendar, notify me when I need to leave, based on current traffic conditions, to be on time" or "if my email contains records of a booked flight, show flight status"
- Location reporting and sharing: Glympse, Google Latitude, etc.

# Urban computing

Use broadly collected data for:

- Zoning and planning
- Traffic monitoring and management
- Public transportation planning
- Crisis detection and management
- Energy consumption sensing
- Air quality monitoring

Much of this data comes from traces of mobile activity!



# Quantified self

## Example applications

Withings devices

Sleep Cycle Instant HR

- Measure "self", visualize, and correlate
- Idea dates back to 70s; term coined ~2007 by Kevin Kelly
- Both peripheral sensors as well as just apps; e.g.
  - Heart rate, Sleep quality
  - Weight, Activity
  - …

http://quantifiedself.com/



# Healthcare


Quantified self: log everything

Medical applications: glucose, asthma, ECG, …

- Related to quantified self
  - Many of these services can send data to your doctor
- Our distinction: specific goal vs. "log everything" approach



# Healthcare

Micro level / personal; e.g.

- Fall detection
- Activity detection
- Mood detection
- …

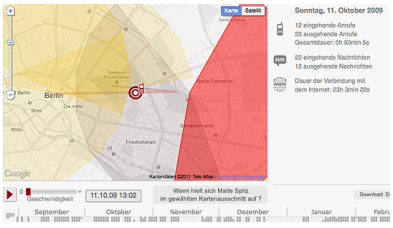
Macro level / population; e.g.

- Disease propagation (mobile, Twitter & Foursquare, …)



# Privacy

## Examples



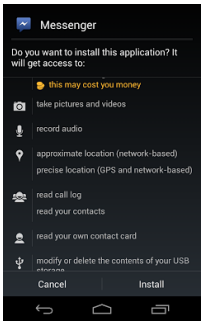

- Vast data that allows quite accurate activity tracking or inferences
- Clearly raises privacy concerns
- Policy ( & technology ?)

"Tell-all telephone" – Die Zeit & Malte Spitz



# Security & Malware

## Examples




- Mobile malware: 6x [Juniper]
- E.g., BadNews: malware on Google Play (30+ apps, 2M downloads, fake app update prompts, mobile "pickpocketing")

Some challenges:

- Role mining: characterize groups of permissions more meaningfully
- Unusual activity detection

## Mobile mining

- The mobile "revolution" (like the "PC revolution") brings together many disciplines and touches many areas
- So, we had to draw some (occasionally arbitrary) divisions, and leave several things out

This tutorial focuses on:

- Work with a substantial analytics component
- Data collected via smartphones (although we'll touch on others sensors briefly, but we won't go into sensing or ubiquitous computing territories—much)



## The rest of this tutorial

Algorithms & Applications I:

- Mobile sensing
- Urban computing
- Healthcare

Algorithms & applications II:

- Location and context
- Advertising
- Social



RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

# MINING DATA FROM MOBILE DEVICES

Introduction

*Spiros Papadimitriou*, Tina Eliassi-Rad



## License