

---

# Collecting mobility data with smartphones: challenges and opportunities

Michel Bierlaire  
Transp-OR

# Outline

---

- Data recording
- Modeling mobility patterns
- Data visualization

---

Opportunities, Challenges and Solutions

# DATA RECORDING ON CELL PHONES

# Opportunities

- Rich data available on smart phones.

## Context

Ambient  
Sound



GSM  
GLOBAL SYSTEM FOR  
MOBILE COMMUNICATIONS



# Opportunities

---

- Rich data available on smart phones.

## Usage of the phone



1. Calendar Entries
2. Phone Log
3. Media Play Log
4. Contacts

# Opportunities

---

- Rich data available on smart phones.

## Others



1. Accelerometer
2. Snapshot of the screen

# Opportunities

---

- Rich data available on smart phones.
- Data is collected from individuals.
- Data is constantly recorded, because users take along their cell phones all the time.

# Challenges

---

- A cell phone software recording all available data.
- Huge battery consumption by GPS data retrieving.
  - Only works less than 6 hours continuously.
- Privacy issues concerned by cell phone users.

# Solutions

---

- A symbian S60 software records data constantly and sends data automatically via wireless network to a remote data server.
- The software combines data from accelerometer, GSM, BT and WIFI to determine when to start and to stop recording GPS data. With this improvement, the software can run a day with normal usage.

# Data Collection Campaign

---

- In collaboration with Nokia Research Center at Lausanne, a data collection campaign has been launched since September, 2009. It will last until summer 2010.
- Currently  $>75$  participants. We expect 120 in the near future.
- An agreement is signed by participants concerning privacy issues. And the data is anonymized before usage.

---

# MODELING MOBILITY PATTERNS FROM DATA



# Opportunities

---

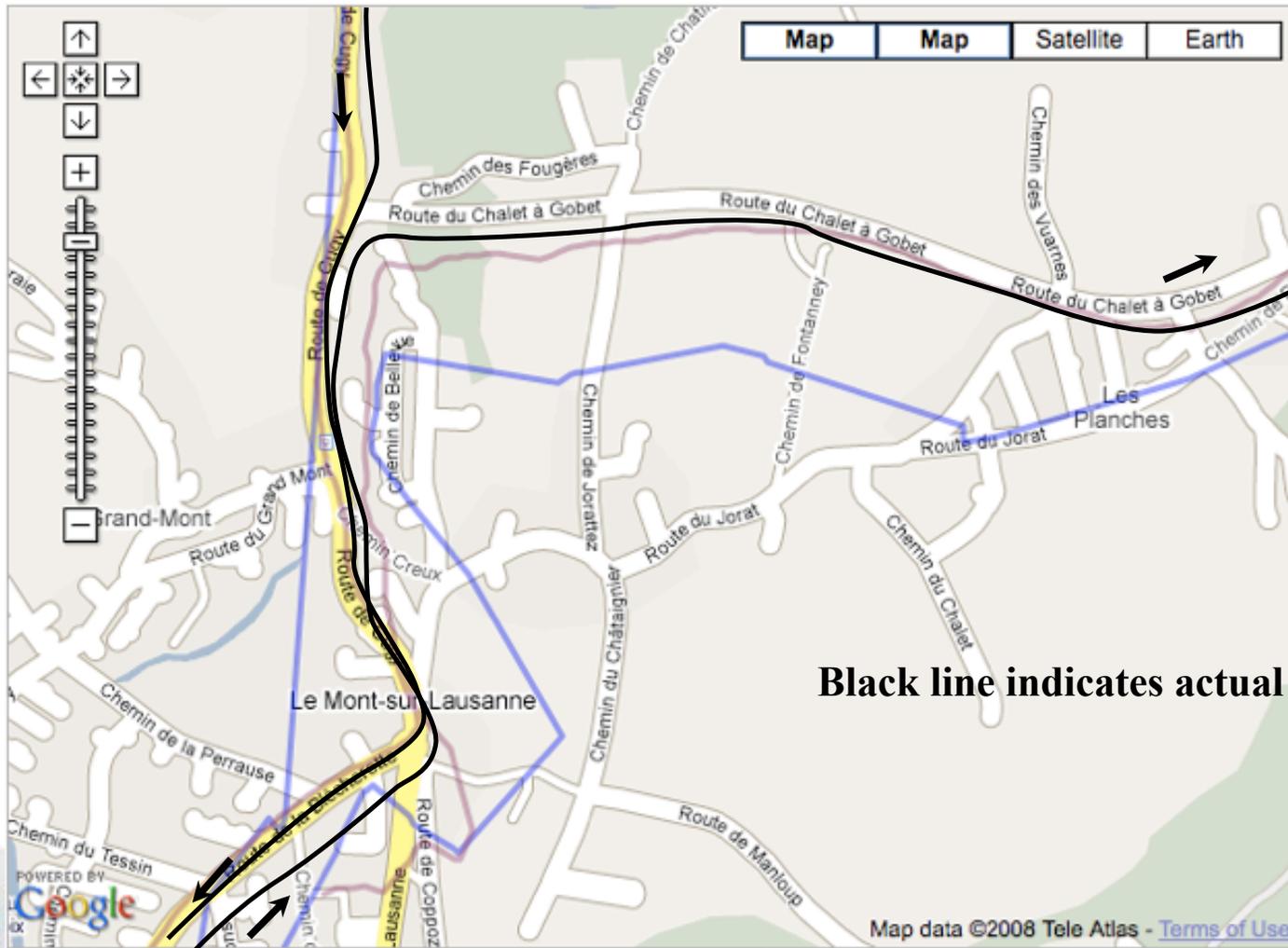
- Some pieces of data are available to transportation researchers for the first time .
- Rich data reflects individuals' mobility patterns, surroundings and characteristics.
- With location data, other data can be tagged with locations.

# Challenges 1

---

- Inferring the travelled path from GPS data.
  - GPS data collected from cell phones is not as accurate as dedicated GPS devices.
  - Map matching doesn't work well for inaccurate GPS data.

# GPS device VS Cell phone



**Black line indicates actual path**

# Ongoing work

---

- A new methodology takes advantage of
  - the spatial relationship between GPS points and network elements, and
  - the temporal relationship underlying the observations and network structures.
- It accounts for poor quality of GPS data.
- It generates probabilistic path observations from GPS data. (Bierlaire et al., 2009)

# Challenge 2

---

- How to infer mobility patterns from various kinds of data?
  - Nearby Bluetooth devices, WIFI stations, GSM towers reflect the environment?
  - Media play history reflects a user's characteristics?
  - Phone log and calendar entries?
  - Others?

# Inferring users' activity

---

- Each individual has the habit that he performs a certain kind of activity (e.g. work),
  - with a certain group of people, (nearby cell phones by BT)
  - in a certain environment (nearby computers and wifi spots by BT and WIFI),
  - at a certain location (location by GPS),
  - in certain time range (time stamp).

# Ongoing work

---

- Estimating activities by using Bluetooth data (Hurtubia et al., 2009).
  - With activity survey data and land use data, Bayesian inference and random utility models are used to infer the activity type of a user at a location and time.

# Future works

---

- Fuse various kinds of data to infer user's activity type.
- Use generated path observations to model users' route choice and transportation mode choice behavior.
- Fuse various kinds of data to model users' mobility patterns under different situations (modeled from context data).

---

Servers and Tools

# VISUALIZING DATA TO USERS



# Developed Tools

---

- A data storing server.
- A GPS track visualization page with activity survey function.
- A page visualizing friends' GPS tracks.
- Pages visualize statistics of data in different areas.



# Visualization Tools (Activity Survey)

**Location Search**

eu.006 09/18/2009 Admin Friends Configuration Contact Logout guided report what,where show activities hide

**Tool Box**

Clear Close  
 <Prev Today Next>  
 August 2009  
 Su Mo Tu We Th Fr Sa  
 1  
 2 3 4 5 6 7 8  
 9 10 11 12 13 14 15  
 16 17 18 19 20 21 22  
 23 24 25 26 27 28 29  
 30 31

**Activity**

20:37 (-1 d) – 00:57 (+0 d)  
 home BY bike  
 [ Edit ]  
 [ Delete ]

**Floating Calendar, indicates days having GPS data**

**GPS Track**

**Time Tooltip when the mouse moves over a travelled location**

**Time Slider**

00:00 03:00 06:00 09:00 12:00 15:00 18:00 21:00 24:00

EPFL Ecole Polytechnique

EPFL  
 POLYTECHNIQUE  
 DE LAUSANNE

# Visualization Tools (Activity Survey)

The screenshot displays a web browser window with the URL `http://transporpc1.epfl.ch/noko/survey.php#`. The interface includes a "Tool Box" at the top with navigation links like "Admin", "Friends", "Configuration", "Contact", and "Logout". A date field shows "09/17/2009".

The main map area shows a street grid with a highlighted yellow route. A blue location pin is placed on the route near the EPFL building. A green line indicates a path along the route. A blue box with white text is overlaid on the map, containing the text: "Yes: performed an activity; No: not an activity, just passed by." A green box with white text is also overlaid, containing: "A stop is automatically detected with location and time information." Arrows point from these boxes to the map.

On the right side, there are three panels:

- New Activity:** A form with the question "Did you come here at around 12:59 (-3 d)." and two buttons: "YES" and "NO".
- Activity History:** Text indicating "Total number of visiting this place: 18." and "Last time you came this place was for [social/recreation](#) by [transit](#), at: [2008-12-12 11:11:43+01](#); duration of stay: [00:02:00](#)." Below this are three pie charts: the first for "home" (social/recreation, pick up/drop c, work), the second for "transit" (bike, walk, car as a drive), and the third for "morning" (afternoon, evening).

At the bottom, there is a time axis from 00:00 to 15:00, with a green bar indicating the current time. A scale bar shows 1000 feet and 200 meters. The bottom right corner has a red square logo with the text "TIQUE ANNE".

# Visualization Tools (Activity Survey)

The screenshot shows a web browser window displaying a Google Maps interface. The URL is <http://transporpc1.epfl.ch/noko/survey.php#>. The page features a "Tool Box" at the top with navigation links and a search bar. The main area is a map of a location, with a green line indicating a path. A red pin is placed on the map, and a blue arrow points to it from the "History of visiting the same location" box. The sidebar on the right contains two main sections: "Create New Activity" and "Activity History".

**Survey form with prefilled information and suggestions**

**History of visiting the same location**

**Create New Activity**

What activity did you perform here?  
--select an activity-- [social/recreation?](#)

How did you come here?  
--select a mode-- [transit?](#)

ARRIVE      LEAVE

↑ ↑      ↑ ↑

- +      - +

3 d - 12:59 -- 18:21 + +0 d

↓ ↓      ↓ ↓

previous stop << >> next stop

**Activity History**

Total number of visiting this place: 18 .  
Last time you came this place was for [social/recreation](#) by [transit](#), at 2008-12-12 11:11:43+01; duration of stay: 00:02:00.

home      social/recreat  
pick up/drop c  
work

100%

# Visualization Tool (Friends' Tracks)

The screenshot shows a web browser window displaying a Google Maps interface. The URL is <https://transporsrv2.epfl.ch/showfriend.php#>. The page features a 'Tool Box' with navigation links: eu.006, Admin, Survey, Configuration, Contact, and Logout. A search bar is present with the placeholder text 'search an address'. The main map area shows the city of Lausanne and surrounding areas, with several colored tracks overlaid: orange, green, blue, and purple. A 'List of friends' table is visible on the right side of the map, listing user IDs and dates. Below it is a 'Load Track' table with colored rows corresponding to the tracks on the map. A 'LOAD CLEAR' button is located at the bottom of the right-hand panel. The bottom of the page includes a Google logo, a scale bar (1 mile / 1公里), and a copyright notice: '地图数据 ©2009 Tele Atlas - 使用条款'.

**Tool Box**  
eu.006 Admin Survey Configuration Contact Logout

search an address

**List of friends**

eu.006	09/17/20
eu.005	09/23/20
eu.003	10/01/20
eu.000	10/06/20
eu.010	10/04/20

**Loaded tracks**

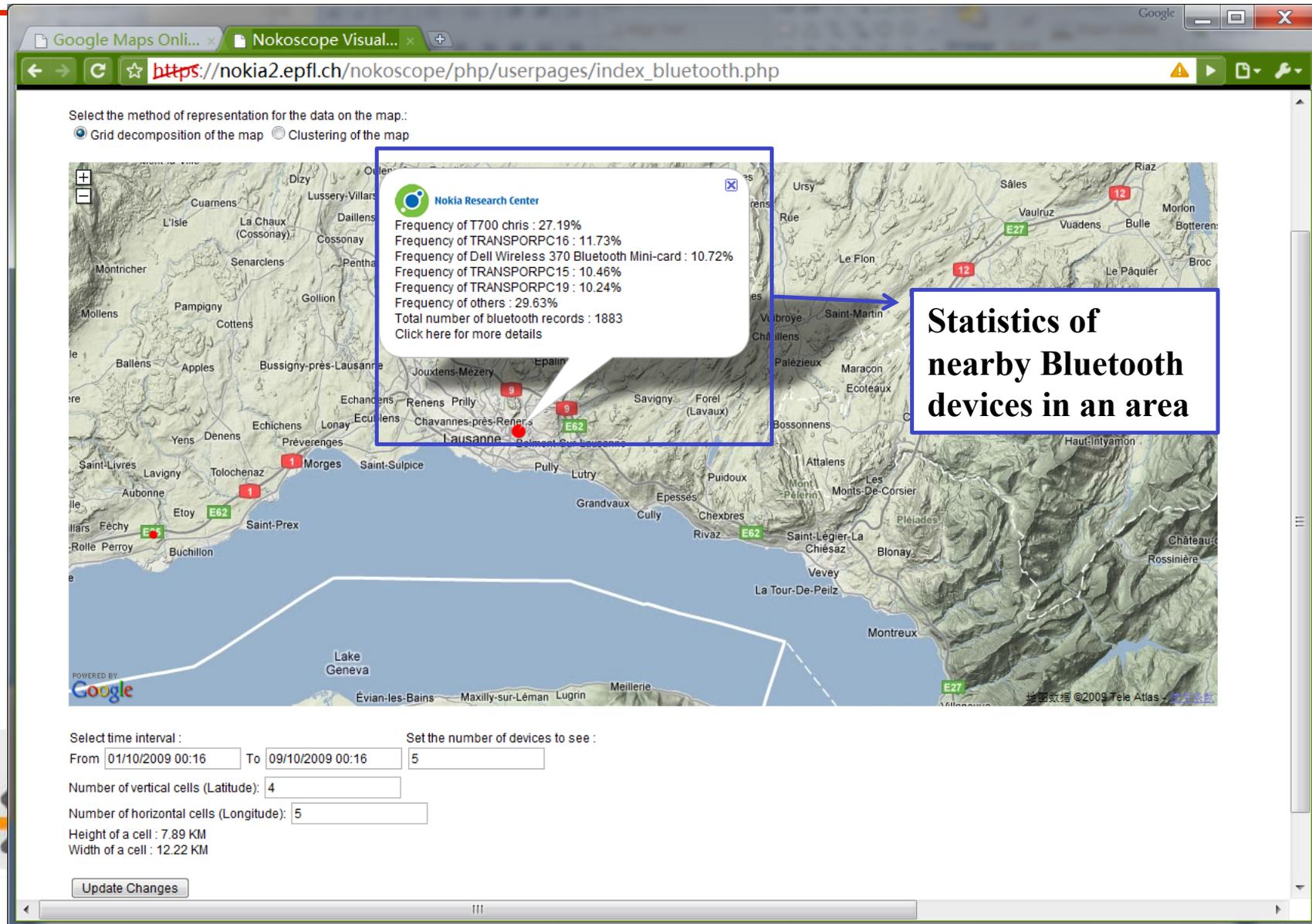
eu.005 at 2009-09-23	**
eu.003 at 2009-10-01	**
eu.000 at 2009-10-06	**
eu.010 at 2009-10-04	**
eu.006 at 2009-09-17	**

LOAD CLEAR

POWERED BY Google 1 英里 1 公里

地图数据 ©2009 Tele Atlas - 使用条款

# Visualization Tools (Nearby BT)



# Visualization Tools (Calls)

Select the method of representation for the data on the map.:  
 Grid decomposition of the map  Clustering of the map

**Nokia Research Center**  
Frequency of 792491282 : 100%  
Total number of calls records : 18  
[Click here for more details](#)

**Statistics of calls in an area**

Select time interval :  
From   
To

# Ongoing works

---

- Visualizations of other data.
- Visualization on google earth with trip animations.