
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



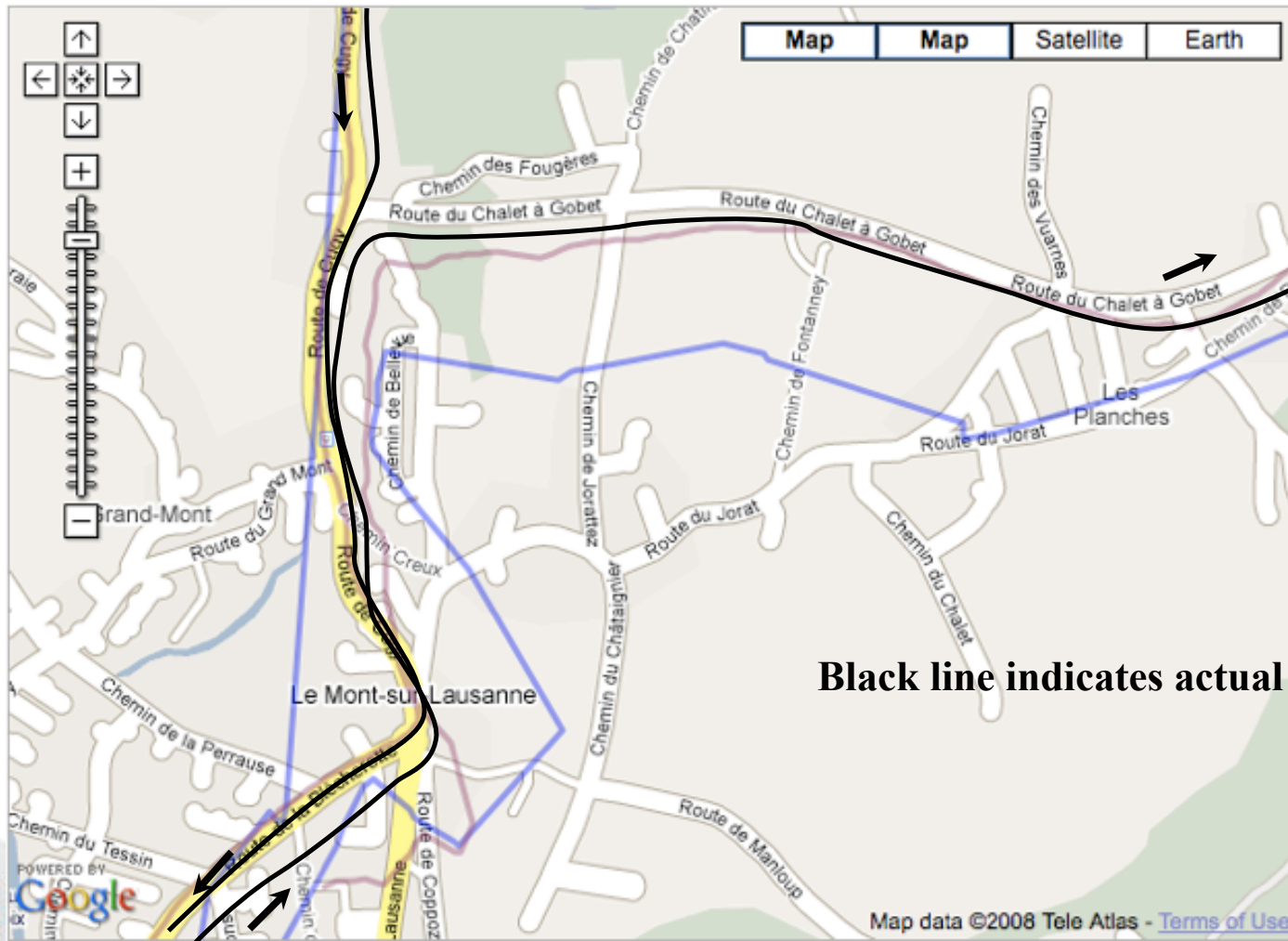
Opportunities

- 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.

Data Server

- Data is sent from cell phones to a remote server. A website is used to retrieve data from the database.

Nokia Simple Con...
https://simplecontext.com/eb2/data?&offset=0&limit=20

home | data | doc | apps eu.006 | profile | logout

POST tags [advanced post »](#) [key-value tables](#)

GET type from_time to_time names limit 20 offset 0

filter_empty stats_only

Results 1 - 20 of about 238,135 from 87,041,850 entries (0.008 seconds) << earlier [download this result set](#) | [delete this record set](#)

upload: Attempting upload
20 days ago from eu.006 via 355739023194834 - 163 b
[get](#) - [delete](#) - [show](#) - [neighbors](#)

sys: {'profile': 'general', 'battery': 14, 'charging': 0, 'drivespace': {'C:': 26542080, 'Y:': 1280000, 'D:': 67948544, 'Z:': 0, 'E:': 220135424}, 'inactive': 0, 'ring': 'normal', 'free_ram': 67948544}
20 days ago from eu.006 via 355739023194834 - 337 b
[get](#) - [delete](#) - [show](#) - [neighbors](#)

app: {'event': 'Application.View', 'uid': '[eeee512f]', 'name': 'EPFLscope', 'view': '[eeee512f]'}
20 days ago from eu.006 via 355739023194834 - 234 b
[get](#) - [delete](#) - [show](#) - [neighbors](#)

app: {'event': 'Application.Foreground', 'uid': '[eeee512f]', 'name': 'EPFLscope', 'view': '[eeee512f]'}
20 days ago from eu.006 via 355739023194834 - 240 b
[get](#) - [delete](#) - [show](#) - [neighbors](#)

app: {'event': 'Application.Started', 'uid': '[eeee512f]', 'name': 'EPFLscope', 'view': '[eeee512f]'}
20 days ago from eu.006 via 355739023194834 - 237 b
[get](#) - [delete](#) - [show](#) - [neighbors](#)

gsm: {'cell_id': 102704, 'network_code': 1, 'area_code': 1, 'signal_dbm': 96, 'country_code': 228, 'signal': 7}
20 days ago from eu.006 via 355739023194834 - 247 b
[get](#) - [delete](#) - [show](#) - [neighbors](#)

sys: {'profile': 'general', 'battery': 14, 'charging': 0, 'drivespace': {'C:': 26546176, 'Y:': 1280000, 'D:': 67837952, 'Z:': 0, 'E:': 220135424}, 'inactive': 254, 'ring': 'normal', 'free_ram': 67837952}
20 days ago from eu.006 via 355739023194834 - 339 b
[get](#) - [delete](#) - [show](#) - [neighbors](#)

gsm: {'cell_id': 102704, 'network_code': 1, 'area_code': 1, 'signal_dbm': 94, 'country_code': 228, 'signal': 7}
20 days ago from eu.006 via 355739023194834 - 247 b
[get](#) - [delete](#) - [show](#) - [neighbors](#)

process: [z:\sys\bin\vekern.exe', z:\sys\bin\efile.exe', z:\sys\bin\domainSrv.exe', z:\sys\bin\SecEnvInit.exe', z:\sys\bin\EInfoServer.exe', z:\sys\bin\IntegrityCheckServer.exe', z:\sys\bin\centralrepositorysrv.exe', z:\sys\bin\EComServer.exe', z:\sys\bin\EwSrv.exe', z:\sys\bin\FbServ.exe', z:\sys\bin\IMC_info.exe', z:\sys\bin\mc_isiserver.exe', z:\sys\bin\mc_useragent.exe', z:\sys\bin\Starter.exe', z:\sys\bin\c32start.exe', z:\sys\bin\c32exe.exe', z:\sys - (132 items) ...more

20 days ago from eu.006 via 355739023194834 - 4.4 Kb

Map of EPFL and Saint-Sulpice area. Location-hot! last gps 20 days ago (gsm est.)

stats 761.3 Mb (total)

get 14481 (total requests)



Visualization Tools (Activity Survey)

Location Search

eu.006 09/18/2009 Admin Friends Configuration Contact Logout guided report 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) [Edit]
home BY bike [Delete]

GPS Track

Floating Calendar, indicates days having GPS data

Time Tooltip when the mouse moves over a travelled location

Time Slider

9:59

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

EPFL Ecole Polytechnique

POWERED BY Google

1000 英尺
200 米

EPFL
POLYTECHNIQUE
ÉCOLE 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 user information (eu.006, 09/17/2009) and navigation links (Admin, Friends, Configuration, Contact, Logout). A "free report" section contains a search box with "what,where" and buttons for "show discard", "show activities", and "hide".

The central map shows a route in yellow and green. A blue box with a white background and black text is overlaid on the map, containing the text: "Yes: performed an activity; No: not an activity, just passed by." A green box with a white background and black text is also overlaid, containing the text: "A stop is automatically detected with location and time information." Arrows point from these boxes to specific locations on 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 labeled "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/recreat, 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 scale bar (1000 英尺 / 200 米) and a time axis from 00:00 to 15:00. The bottom right corner features 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 address bar shows the URL `http://transporpc1.epfl.ch/noko/survey.php#`. The page title is "Tool Box".

Survey Form with prefilled information and suggestions

The survey form is titled "Create New Activity". It contains the following fields and options:

- What activity did you perform here?
--select an activity-- [social/recreation?](#)
- How did you come here?
--select a mode-- [transit?](#)
- ARRIVE LEAVE
- Time: 12:59 -- 18:21 +0 d
- Buttons: previous stop, next stop
- Buttons: Save, Discard

History of visiting the same location

The activity history section shows:

- 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](#).

A pie chart below the history section shows the distribution of activities:

- home
- social/recreat
- pick up/drop c
- work

The map shows a location in Lausanne, Switzerland, with a green pin and a green line indicating a route. The map includes labels for streets like "Rue du Villars", "Avenue du TIR Fédéral", and "Route Cantonale".

Visualization Tool (Friends' Tracks)

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

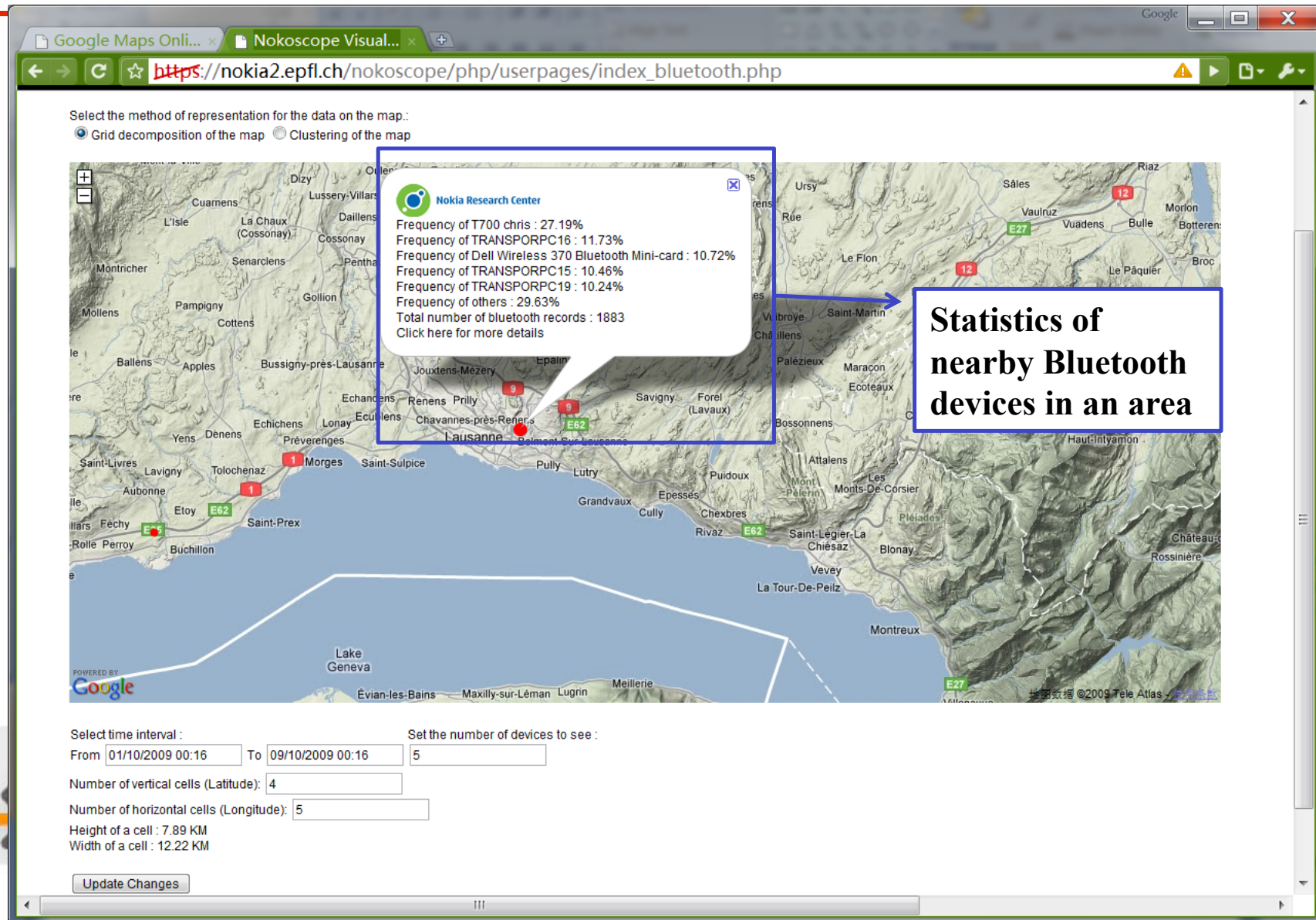
eu.005 at 2009-09-23	**
eu.003 at 2009-10-01	**
eu.000 at 2009-10-06	**
eu.010 at 2009-10-04	**
eu.006 at 2009-09-17	**

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.