Collecting mobility data with smartphones: challenges and opportunities

Michel Bierlaire
Transp-OR

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

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

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

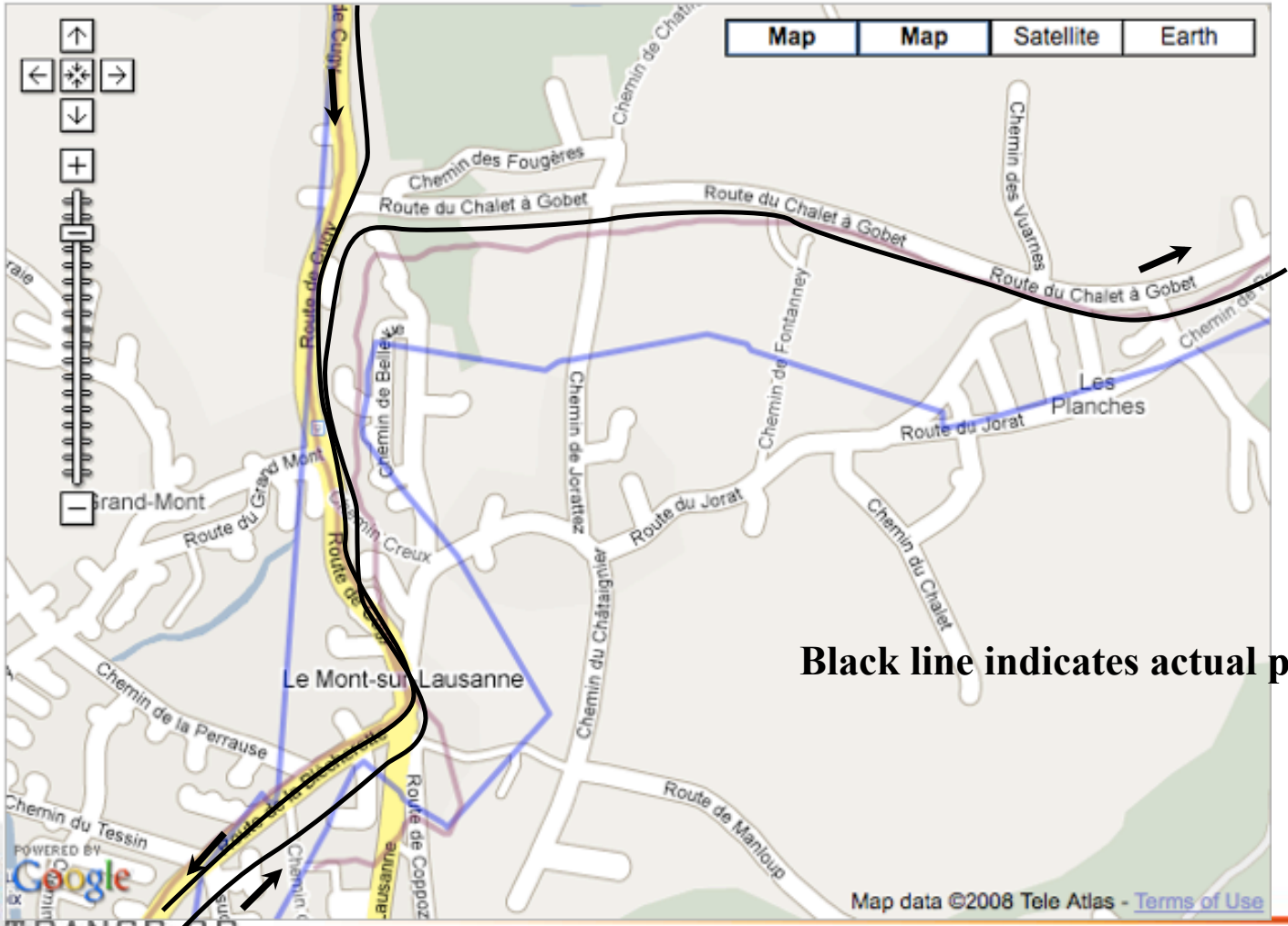
Opportunities

- 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

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

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

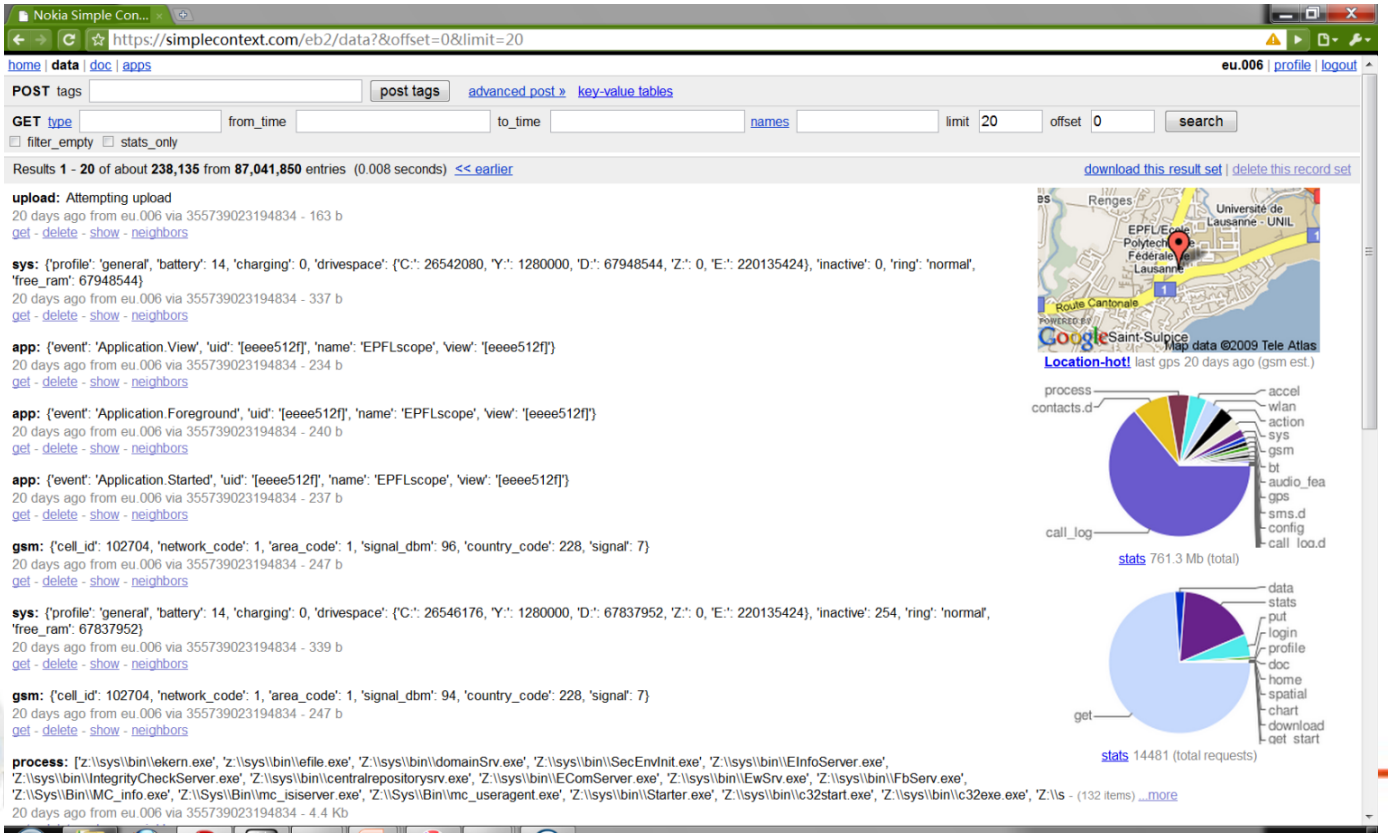
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

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.

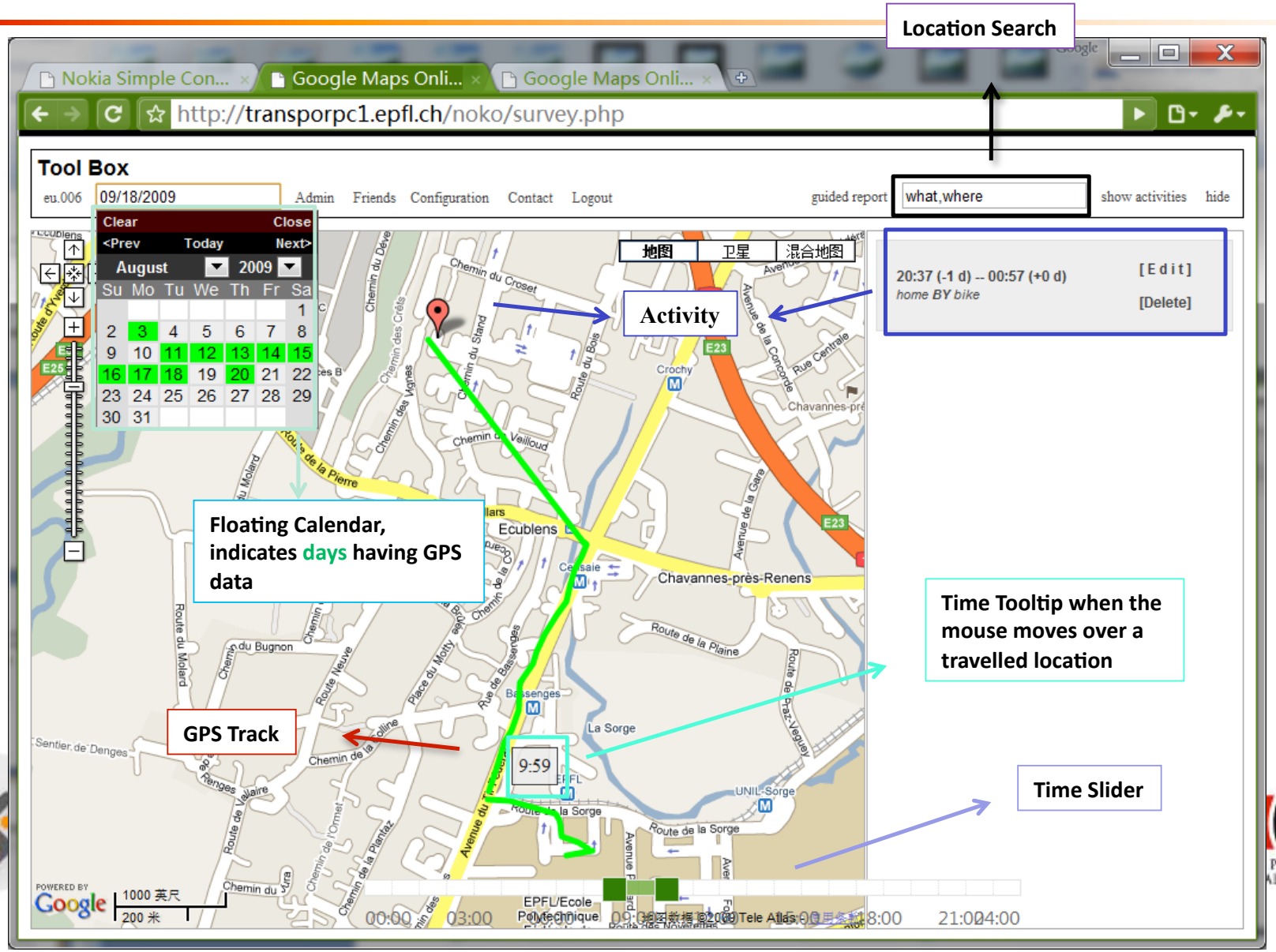
Data Server

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

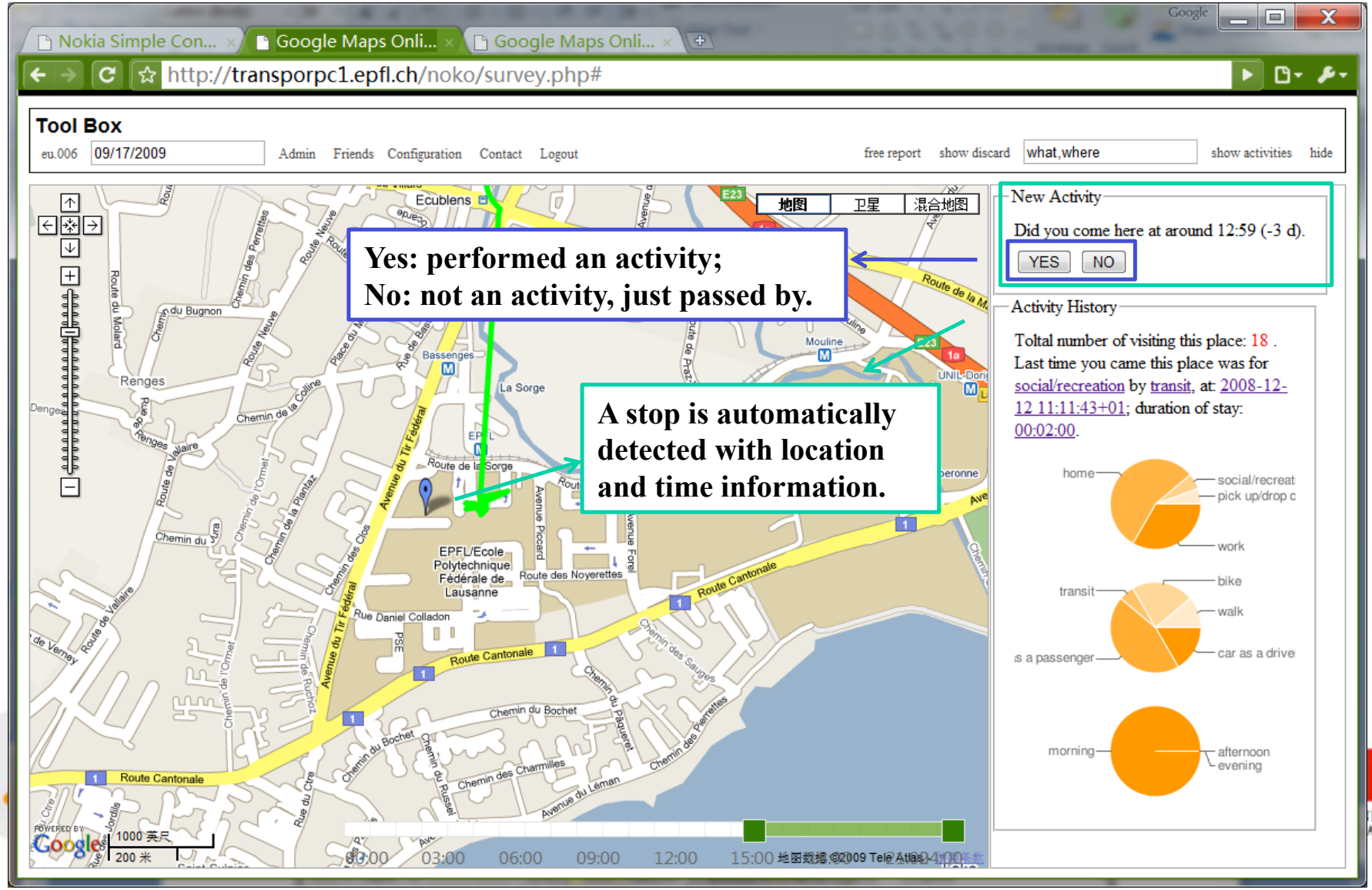


ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

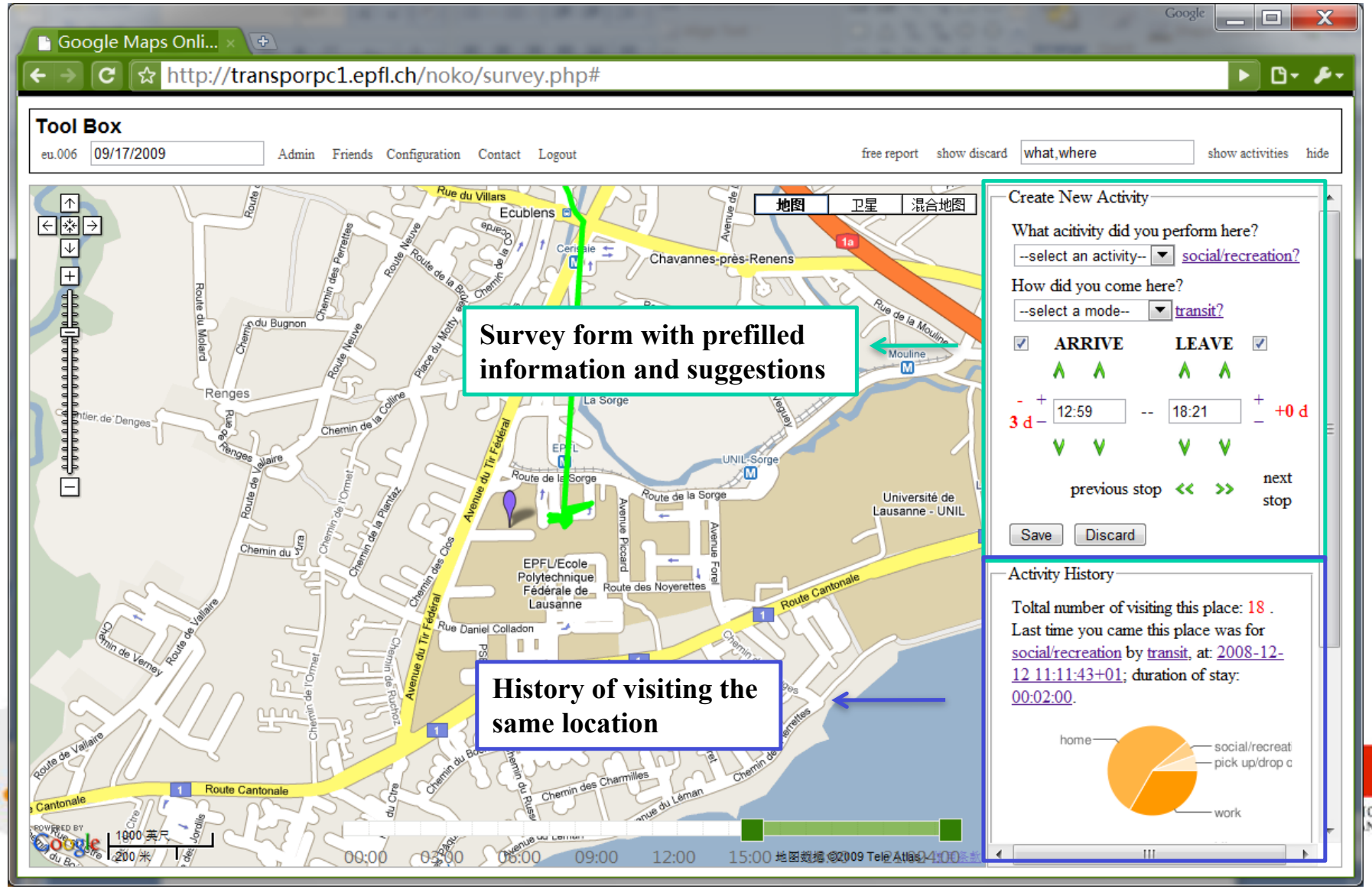
Visualization Tools (Activity Survey)



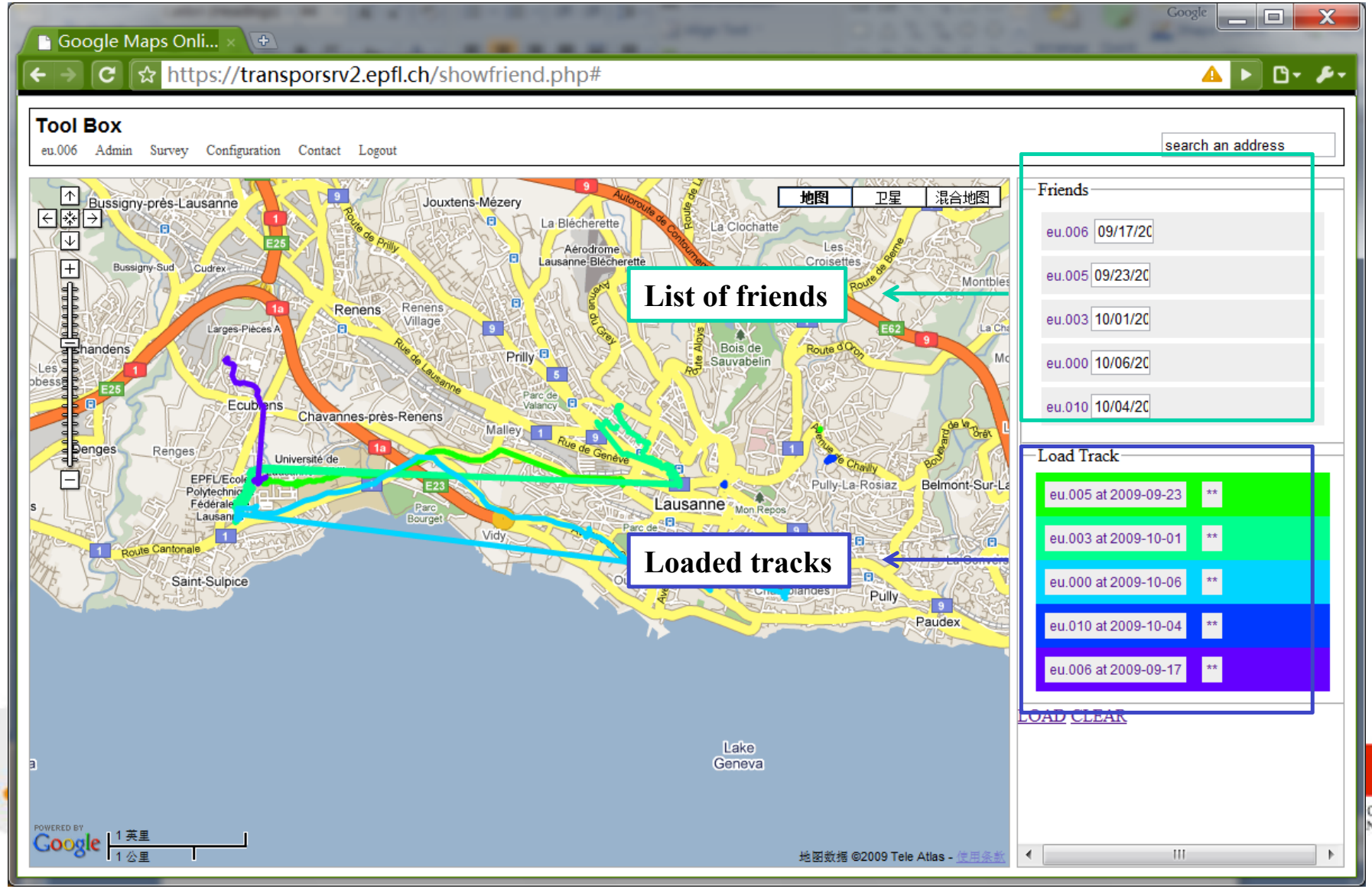
Visualization Tools (Activity Survey)



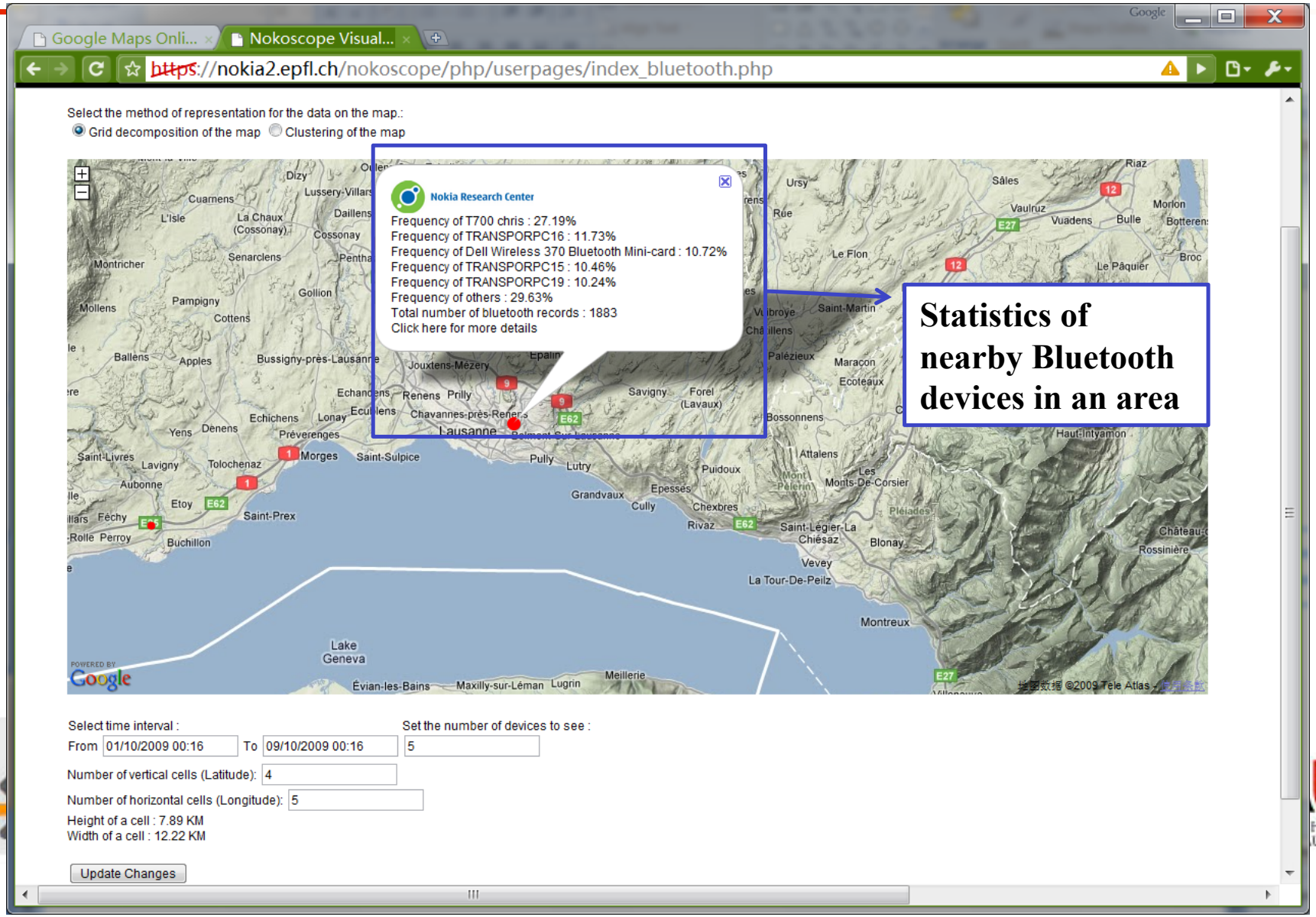
Visualization Tools (Activity Survey)



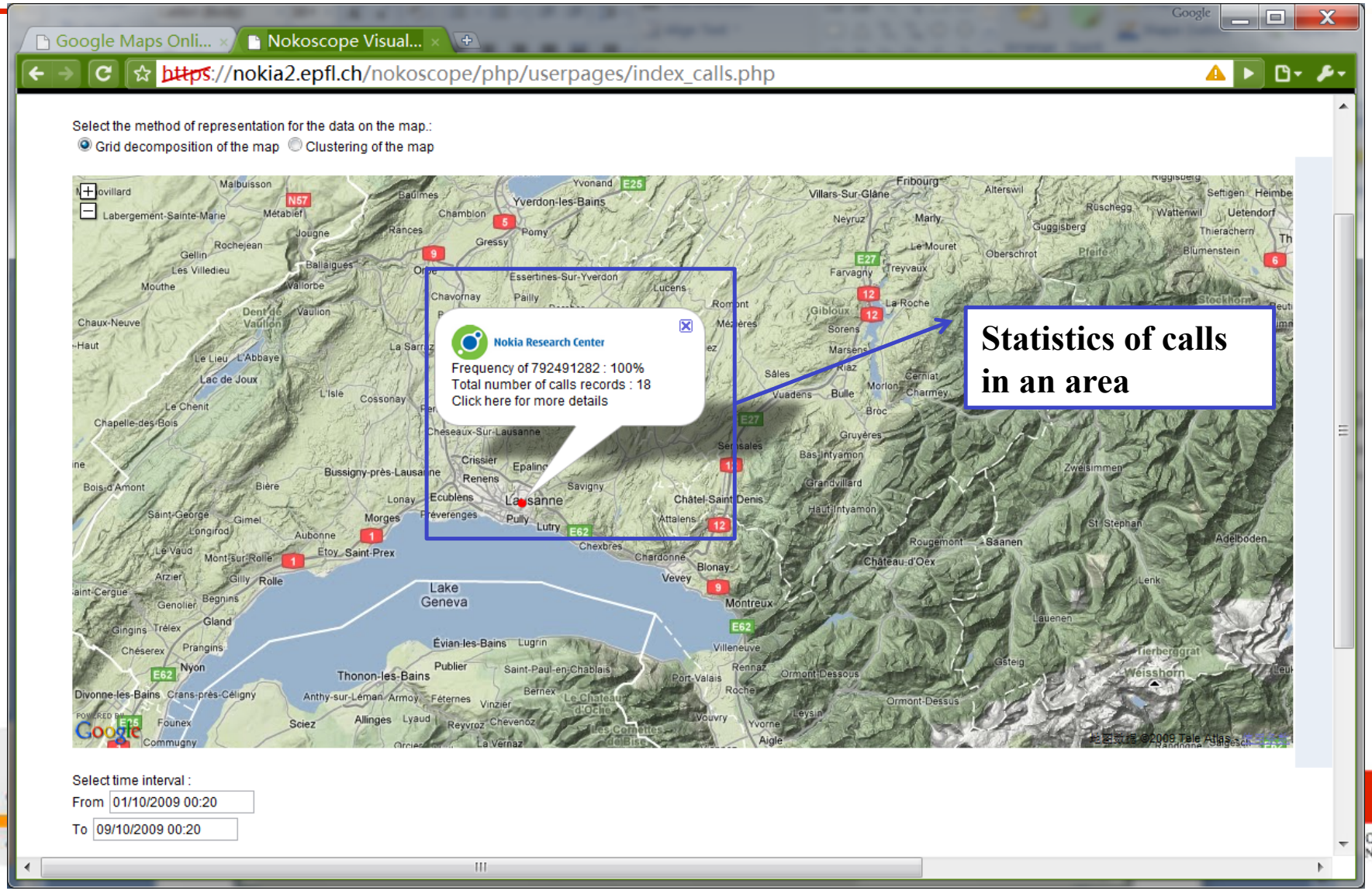
Visualization Tool (Friends' Tracks)



Visualization Tools (Nearby BT)



Visualization Tools (Calls)



Ongoing works

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