

Mobile Century data documentation

December 16, 2009

Contents

1	USER AGREEMENT, TERMS OF USE	1
2	DESCRIPTION OF THE DATA	2
2.1	Inductive loop detector data	2
2.1.1	Inductive loop detector properties	3
2.2	Virtual Trip Line data	3
2.3	Ground truth travel time data	4
2.4	Vehicle trajectory data	4
2.4.1	Individual trip data	5
2.4.2	Individual phone log data	5

1 USER AGREEMENT, TERMS OF USE

The *Mobile Century* data was collected on February 8, 2008, as part of a joint UC Berkeley - Nokia project, funded by the California Department of Transportation, to support the exploration of uses of GPS enabled phones to monitor traffic. In addition to the cell phone GPS data, two additional data sources are available for the experiment site. Inductive loop detector data obtained through the *Freeway Performance Measurement System* (PeMS), and travel time data obtained through vehicle re-identification using high resolution video data are included with this release. All identifiers assigned to the cell phones used during the *Mobile Century* experiment have been randomized to protect the participants in the experiment. The video data

is also processed and random number has been assigned to represent each vehicle.

An extensive description of the experiment and data is available in the following article: Herrera, J.C., et al. "Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment". *Transport. Res. Part C* (2009), doi:10.1016/j.trc.2009.10.006. Additional information about the successor project, known as *Mobile Millennium*, is available at <http://traffic.berkeley.edu>.

By downloading the data, the user acknowledges that:

- **The data is available for use for research and analysis purposes only.**
- **He/she will not redistribute the data.**
- **Any publication using the data should refer the following article: Herrera, J.C., et al. "Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment". *Transport. Res. Part C* (2009), doi:10.1016/j.trc.2009.10.006.**

2 DESCRIPTION OF THE DATA

The data was collected during the Mobile Century experiment on Feb 8th 2008 between 10:00am and 18:00pm (PST) on Interstate 880, CA.

2.1 Inductive loop detector data

Inductive loop detector data, obtained from the *Freeway Performance Measurement System* (PeMS) database, consist of two sets: Northbound (NB) and Southbound (SB) 30 second raw loop detector data. Data are stored in two files:

1. `pems_rawdata_SB.csv` and
2. `pems_rawdata_NB.csv`.

These files contain the following columns: "pems_id", "unixtime", "flows" and "occs", see Table 1 for the detailed explanation.

Table 1: PeMS raw data columns

"pems_id"	refers to the <i>Vehicle Detector Station</i> (VDS) ID of the loop detector station in the PeMS database.
"unixtime"	refers to Unix time (seconds, UTC).
"flows"	refers to vehicle counts per lane over 30 second period.
"occs"	refers to occupancy of the sensor per lane over a 30 second period.

Table 2: PeMS properties file columns

"pems_id"	refers to VDS ID of the loop detector station given by PeMS.
"abs_pm"	refers to an absolute postmile (on Interstate 880) of the loop detector station given by PeMS.
"lat", "lon"	refer to the latitude and longitude (degrees) of the loop detector station, respectively. Note, that lat/lon coordinates are approximate locations of the loop detector stations.

2.1.1 Inductive loop detector properties

Geographical locations of the Northbound and Southbound inductive loop detectors extracted from the PeMS database are given in two files:

1. pems_prop_NB.csv and
2. pems_prop_SB.csv.

These files have the following columns: "pems_id", "abs_pm", "lat" and "lon", see Table 2 for detailed explanation.

2.2 Virtual Trip Line data

VTL (Virtual Loop Detector) data consist of two sets: Northbound (NB) and Southbound (SB) VTL speed data. Data are stored in two files:

Table 3: VTL data columns

"vtl_id"	refers to the integer ID of the VTL that the mobile device crossed when the data was sent.
"unixtime"	refers to Unix time (seconds, UTC).
"coordinate"	refers to a latitude-longitude (degrees) coordinate pair that defines the geographical location of the measurement.
"vel_mph"	refers to the reported speed (mph) of the mobile device when crossing the VTL.

1. vtl_data_SB.csv and
2. vtl_data_NB.csv.

These files contain the following columns: "vtl_id", "unixtime", "coordinate" and "vel_mph", see Table 3 for detailed explanation.

2.3 Ground truth travel time data

Ground truth travel time data is provided between Stevenson Blvd and Decoto Road and between Decoto Road and Winton Avenue for the Northbound direction of Interstate 880. Data are stored in files

1. ground_truth_traveltimes_1_NB.csv and
2. ground_truth_traveltimes_2_NB.csv,

respectively. The files contain two columns: "departure_time" and "travel_time", see Table 4 for detailed explanation.

2.4 Vehicle trajectory data

Two different types of vehicle trajectory data are provided. The first type described in Section 2.4.1 consists of individual trips on one direction (NB and SB) of the highway.

The second type described in Section 2.4.2 consists of 77 individual phone logs. These logs contain the unprocessed latitude and longitude coordinates of the vehicles throughout the experiment.

Table 4: Ground truth travel time data columns

"departure_time"	refers to Unix time (seconds, UTC) when vehicle started the trip (from Stevenson Blvd or Decoto Rd).
"travel_time"	refers to the travel time (seconds) experienced by the vehicle recorded (at Decoto Rd or Winton Ave).

2.4.1 Individual trip data

These data consist of individual “trips” on one direction of the highway. Northbound trips are in the "NB_veh_files" folder and the southbound trips are in the "SB_veh_files".

Each file contains the following five columns: "unixtime", "latitude", "longitude", "postmile"¹ and "speed"², see Table 5 for detailed explanation.

2.4.2 Individual phone log data

Individual phone logs consist of 77 GPS log files extracted from the Nokia N95 mobile devices. Data are stored under the GPS_logs folder in 77 files named vehXXX.csv where XXX is ranging from 101 to 177.

Each file contains the following four columns: "unixtime", "latitude", "longitude" and "speed"², see Table 6 for detailed explanation.

¹**Important notice:** The postmile values are approximated using a projection from the GPS coordinates stored by the mobile devices, onto Interstate 880 postmiles. The postmile references for Interstate 880 are obtained from PeMS loop detector station locations. The highway between two loop detector stations is assumed to be a straight line. Hence, the postmile values may contain errors. Parts of the data when vehicle is not on the highway are excluded from the individual trip data.

²**Important notice:** The GPS chipset in the Nokia N95 cell phones used during the experiment did not estimate the velocity of the phone directly. Instead, the speeds are computed in software by applying a finite difference approximation of the position estimates provided by the GPS chipset. Additional filtering and outlier removal is applied to the speeds released in this dataset. The latitude and longitude data has not been altered from its original form.

Table 5: Vehicle trajectory data columns

"unixtime"	refers to Unix time (seconds, UTC).
"latitude", "longitude"	refer to latitude and longitude coordinates (degrees), respectively.
"postmile"	refers to the number of miles from the start of Interstate 880 in California; a positive value represents the vehicle is on the Northbound side of the highway, a negative value represents the vehicle is on the Southbound side of the highway.
"speed"	refers to the speed (mph) derived from the lat-lon coordinates.

Table 6: Individual phone log data columns

"unixtime"	refers to Unix time (milliseconds, UTC).
"latitude", "longitude"	refer to latitude and longitude coordinates (degrees), respectively.
"speed"	refers to the speed (mph) derived from the lat-lon coordinates.