Modeling Route Choice Behavior From Smart-phone GPS data

Jingmin Chen * † Jeffrey Newman * Michel Bierlaire *

November 5, 2009

Abstract

GPS capable smart phones are emerging survey tools in transportation research field, especially in modeling individuals' mobility patterns. In route choice modeling, path observations need to be generated explicitly for the estimation. It is a challenge because the recorded data is not as dense or accurate as those from dedicated GPS devices. In this paper, we develop a methodology for generating probabilistic path observations from sparse and inaccurate location data, for state-of-the-art discrete route choice models. The difference of the proposed algorithm and the map matching algorithms is that instead of giving a unique matching result, the new algorithm generates a set of potential true paths, along with probabilities for each one to have been the true path. More importantly, the algorithm uses not only the topological measurement, but also temporal information (speed and time) in the GPS data to calculate the probability for observing the data while traveling on the proposed path. We emphasis traveling as a dynamic movement on a path, and model it as such in the algorithm. A short trip and two longer trips are used to analyze the performance of the algorithm on real data. Then, 19 trips recorded from a single user's cell phone are used in a preliminary study that estimates route choice behaviors using state-of-the-art discrete route choice modeling methodologies with the proposed probabilistic path observation generation algorithm.

Keyword: route choice modeling, path observation generation, smartphone data, GPS data, map matching

1 Introduction

Developing technology has long been harnessed to supplement or replace parts of travel behavior surveys. Tools such as GPS tracking devices have been given to survey participants, to track their movements in a systematic and unbiased way, instead of relying merely on travel diaries and prompted recall questioning. Tracking survey participants using a specialized GPS device provides numerous challenges: people may forget to charge the device, or leave it at home, and it may not receive a good signal at all times, leaving gaps in the travel record.

*Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, { jingmin.chen, jeffrey.newman, michel.bierlaire}@epfl.ch

†corresponding author

However, in the developed world, and increasingly in the developing world, many people normally carry a wireless phone with them. They already manage the tasks of charging and remembering to carry it, at least as well as they are going to manage those tasks for any special survey device. Therefore we propose, as in Stopher [2008], to bundle the survey data collection into a phone.

In collaboration with Nokia Research Center in Lausanne, we launched a data collection campaign in September 2009 to collect various kinds of data from Nokia N95 smart-phones, including GPS data. In this paper, we will focus on estimating route choice behaviors from collected GPS data by developing a new methodology for generating probabilistic path observations from recorded GPS data, and applying network-free methodology (Bierlaire and Frejinger, 2008) in discrete route choice models. Instead of deterministically matching the GPS trace onto transportation network to generate a unique path as map matching algorithms do (Ochieng et al., 2003, Quddus et al., 2007), several corresponding path observations are generated probabilistically, along with probabilities for each path to have been the true path (Bierlaire et al. [2009]). This method can avoid the biases which could be potentially introduced by deterministic map matching algorithms, especially for sparse and inaccurate location data.

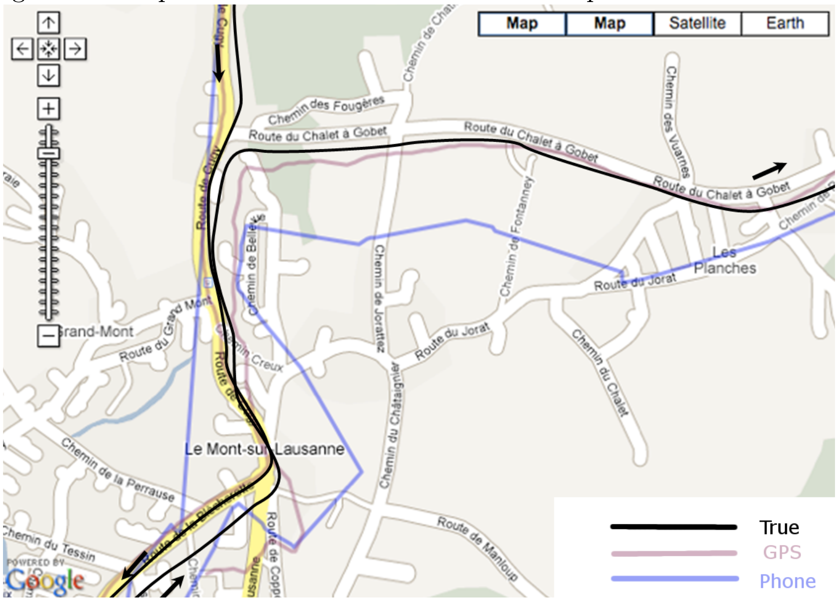
In the next section, analysis on data recording using cell phones will be presented to illustrate our motivation of developing new path observations generation algorithm. In section 3, after briefly presenting the path probability measurement proposed by Bierlaire et al. [2009], we will discuss the issues in the methodology, followed by solutions. Then the algorithm for generating path observations will be presented in section 4. The algorithm will be applied to a short real trip to illustrate the generated result. We aim at applying the algorithm in discrete route choice models to estimate individuals' route choice behaviors. In section 5, results from a preliminary study on the estimation using a small set of real GPS data will be presented.

2 Data collection using smart phones

By mid of October 2009, more than 75 Nokia N95 smart phones had been given out to volunteers as their personal cell phones. As each individual uses his cell phone in daily life, a pre-installed software records and sends various kind of data to a remote data server automatically. Those data include GPS readings, nearby WIFI stations, nearby Bluetooth devices, available GSM tower, accelerometer, media play log, calendar entries and phone log. Once collected, the rich data represents a unique opportunity to estimate mobility patterns of individuals. In this paper, we will focus on using GPS data to estimate route choice behaviors.

Another important feature of most of GPS capable cell phones is Assisted-GPS, which reduce warm-up time for getting the first GPS reading to seconds. This advantage provides more opportunities to observe full tracks of the user's trips without losing the beginning parts of trips. Undoubtedly, cell phones have significant advantages over dedicated GPS devices, however, they also have own shortcomings. Firstly, due to battery constraint, the data is recorded with time interval of 10 seconds. This sampling rate is generally lower than that of dedicated GPS devices, so the collected data is more sparse. Secondly, the data is not as accurate as those collected from dedicated GPS device. In the N95 phones used in this study, the GPS antenna is embedded under the keyboard,

Figure 1: Comparison between GPS data from cell phone and GPS device



which is covered by the screen when the phone is not in use. Furthermore, most of people have the habit that putting cellphones in pockets or handbags when not in use. Both of these factors generally weaken the GPS signal. The difference in data can be easily understood by looking at a map (see Figure 1) with two traces of GPS data, which are recorded at the same time by the same person who carried both a dedicated GPS device (the same device used in Flamm et al. [2007]) and a Nokia N95 smart-phone. The cell phone's recorded trace deviates significantly from the true path, while the data from GPS device shows much better adherence to the true path.

A main application of map matching algorithms is on GPS navigation tools, which are required to locate the user on a certain arc or even position on an arc (see, for example Quddus et al. [2006]). Even though Dead-Reckoning sensors are integrated with some navigation tools to improve the accuracy of the detection, it is still possible that for a entire trip an incorrectly matched arc leads to an incorrectly matched path, and it is even possible that various matched arcs can not be connected to form a path. However, in travel behavior study, especially in route choice modeling, researchers are not interested in associating every single GPS point to an arc. Instead, modelers are concerned about the correctness of the generated path representing the whole process of a trip, because incorrectly matched paths can result in biases in models.

Route choice modeling frameworks have been adapted to accept a probabilistic representation of the actual path (Bierlaire and Frejinger, 2008). An observation need not be a unique path, but can be represented by a set of potential paths, along with a probability for each path to have been the actual one. In map matching algorithms where the Multiple Hypothesis Technique (MHT) is used (Marchal et al., 2005), a set of route candidates is maintained at each GPS point. Then for each route candidate, a score is calculated based on distance, speed or/and heading difference between GPS points and arcs, though heading has found to be unreliable for this usage (Schuessler and Axhausen [2009]). Further, the score calculations, while often heuristically effective, lack the theoretical grounding necessary to serve as the probability that the corresponding path is the true path. Moreover, the simplicity of score calculation can not ensure its correctness if there are outliers in the GPS data. Besides, in such a post-processing algorithm (w.r.t. real time algorithm for navigation tools), "inaccurate" data is eliminated from map matching in the process of data filtering (Schuessler and Axhausen [2009, 2008]). Consequently, some useful information in those "inaccurate" data is also excluded.

In this paper, we present an innovative algorithm which generates probabilistic representation of traveled path from cell phone GPS data. The calculation of the path probability involves the simulation of travel on potential true path. The algorithm is designed to be able to take advantage as much as possible the information in the raw data, since the data is already sparse from cell phones. The result from the algorithm will be probabilistic path observations for state-of-the-art discrete route choice models.

3 Path Probability Measurement

In Bierlaire et al. [2009], a path probability measurement is developed to calculate the probability that a proposed path is the true path given a trace of GPS data. It utilizes both spatial relevance between GPS data and network elements, but also temporal relationship among them. The horizontal accuracy information is used to measure the spatial relevance. Bayesian inference is used to calculate the probability that the traveler was traveling on the proposed path, while the device recorded the location, speed, and time information. In this paper, we briefly introduce this method and then present some improvements which we learned from applying it to real data. For more details about the algorithm, we refer to Bierlaire et al. [2009].

3.1 Algorithm framework

By simulating traveler's movement on a path, the algorithm calculates the probability for generating the observed GPS trace on the path. The probability value is updated when a new GPS point enters,

$$\Pr(\check{g}_j, \check{g}_{j-1}, \cdots, \check{g}_1 | p) = \Pr(\check{g}_j | \check{g}_{j-1}, \cdots, \check{g}_1, p) \cdot \Pr(\check{g}_{j-1}, \cdots, \check{g}_1 | p), \quad (1)$$

where $\check{g}_j$ denotes the jth recorded GPS point in a trip and p is the proposed path. For each GPS point $\check{g}_j$, the point probability $\Pr(\check{g}_j | \check{g}_{j-1}, \cdots, \check{g}_1, p)$ is calculated by integrating, in $\check{g}_j$'s domain of relevance D_j, the probability for generating the GPS point at each possible location, which we term the location probability. The domain of relevance of GPS point (or domain shortly) is a set of arcs, which are close to the GPS point and satisfy some conditions which will be described in the next section. By Bayes, the location probability is the product of the conditional probability for arriving at the location at the observation time, and the marginal probability for generating the GPS point at the location. According to these, we derive

$$\Pr(\breve{g}_j|\breve{g}_{j-1},\cdots,\breve{g}_1,p) = \sum_{a\in(D_j\cap p)} \Pr(\breve{g}_j,a|\breve{g}_{j-1},\cdots,\breve{g}_1,p), \tag{2}$$

$$= \sum_{a\in(D_j\cap p)} l_a \cdot \int_0^1 f_{\boldsymbol{g},\boldsymbol{\epsilon}^j}(\breve{g}_j,\epsilon_a)\cdot f_{\boldsymbol{\epsilon}^j}(\epsilon_a|\breve{g}_{j-1},\cdots,\breve{g}_1,p)d\epsilon_a, \tag{3}$$

where l_a is the length of arc a. The marginal probability, $f_{\boldsymbol{g},\boldsymbol{\epsilon}^j}(\breve{g}_j,\epsilon_a)$, is the spatial measurement for the relevance between GPS point and location in the network, which means the probability that the $\breve{g}_j$ is recorded at a position ϵ_a on the arc a, given by

$$f_{\boldsymbol{g},\boldsymbol{\epsilon}^j}(\breve{g}_j,\epsilon_a) = \frac{1}{\pi(\breve{\sigma_j^x})^2}\exp(-\frac{||\breve{x}_j - x||^2}{2(\breve{\sigma_j^x})^2}), \tag{4}$$

in which $\breve{x}_j$ and x are horizontal coordinates of $\breve{g}_j$ and ϵ_a respectively, $||\breve{x}_j - x||$ is their Euclidean distance, and $\breve{\sigma_j^x}$ is the horizontal accuracy of $\breve{g}_j$. The conditional probability, $f_{\boldsymbol{\epsilon}^j}(\epsilon_a|\breve{g}_{j-1},\cdots,\breve{g}_1,p)$, stands for the probability that the traveler arrives at the current location, given the condition that he departed from the previous GPS point. It is calculated by integrating the probability that the traveler arrives at the current location given each location in the previous domain as the departure location,

$$f_{\boldsymbol{\epsilon}^j}(\epsilon_a|\breve{g}_{j-1},\cdots,\breve{g}_1,p) = \sum_{b\in(D_{j-1}\cap p)} f_{\boldsymbol{a}}(b|\breve{g}_{j-1},\cdots,\breve{g}_1,p)\cdot f_{\boldsymbol{\epsilon}^j}(\epsilon_a|b,\breve{g}_{j-1},\cdots,\breve{g}_1,p) \tag{5}$$

$$= \sum_{b\in(D_{j-1}\cap p)} \frac{\Pr(\breve{g}_{j-1},b|\breve{g}_{j-2},\cdots,\breve{g}_1,p)}{\Pr(\breve{g}_{j-1}|\breve{g}_{j-2},\cdots,\breve{g}_1,p)}\cdot \int_{\epsilon_b=0}^1 f_{\boldsymbol{\epsilon}^j}(\epsilon_a|\epsilon_b,\breve{g}_{j-1},b)\cdot f_{\boldsymbol{\epsilon}^{j-1}}(\epsilon_b|\breve{g}_{j-1},b)d\epsilon_b \tag{6}$$

in which ϵ_b is a position on arc b belonging to the domain of $\breve{g}_{j-1}$. All terms except $f_{\boldsymbol{\epsilon}^j}(\epsilon_a|\epsilon_b,\breve{g}_{j-1},b)$ are calculated when $\breve{g}_{j-1}$ was dealt (see Bierlaire et al. [2009] for more details). And we calculate the position transition probability $f_{\boldsymbol{\epsilon}^j}(\epsilon_a|\epsilon_b,\breve{g}_{j-1},b)$ by simulating traveler's movement from $\breve{g}_{j-1}$ to $\breve{g}_j$. Specifically, it equals to the probability that traveler uses observed time difference between $\breve{g}_{j-1}$ and $\breve{g}_j$ to travel from the previous position to the current position,

$$f_{\boldsymbol{\epsilon}^j}(\epsilon_a|\epsilon_b,\breve{g}_{j-1},b) = f_{\boldsymbol{t}_{b\to a}}(\breve{t}_j - \breve{t}_{j-1}). \tag{7}$$

3.2 Issues and Improvements

3.2.1 Low speed GPS point

The calculation of position transition probability (7) relies on the speed information from GPS data. The speed data are used to calculate the travel time on the arcs where the GPS points are assumed to be observed. However, if a GPS point has a very low speed value, which means that the traveler may have been stopped somewhere, we are not able to derive either the travel time or the stopped time from GPS data. In order to solve such problem, we simplify the position transition probability to,

$$f_{\boldsymbol{\epsilon}^j}(\epsilon_a|\epsilon_b, \breve{g}_{j-1}, b) = \begin{cases} 1, & \epsilon_b \to \epsilon_a \\ 0, & \epsilon_a \to \epsilon_b \end{cases}, \tag{8}$$

in which $\epsilon_b \to \epsilon_a$ means ϵ_a is in the downstream of ϵ_b.

3.2.2 Preference to longer path

There is a drawback using (2) as the point probability. If there are two paths with difference only in the domain of $\breve{g}_j$ that p_2 contains one more arc q, i.e.

$$(D_j \cap p_2) = (D_j \cap p_1) \cup \{q\}, \tag{9}$$

from (2) we will have $\Pr(\breve{g}_j|\breve{g}_{j-1}, \cdots, \breve{g}_1, p_2) \geq \Pr(\breve{g}_j|\breve{g}_{j-1}, \cdots, \breve{g}_1, p_1)$, because $\Pr(\breve{g}_j, q|\breve{g}_{j-1}, \cdots, \breve{g}_1, p) = \vartheta \geq 0$. It leads to an inappropriate result that longer path is always more preferable to shorter one, even if ϑ is very low. Hence, in order to solve this issue, we weight $\Pr(\breve{g}_j|\breve{g}_{j-1}, \cdots, \breve{g}_1, p)$ by the length of path segment which lies in the current domain, then (2) becomes,

$$\Pr(\breve{g}_j|\breve{g}_{j-1}, \cdots, \breve{g}_1, p) = \frac{\sum_{a \in (D_j \cap p)} \Pr(\breve{g}_j, a|\breve{g}_{j-1}, \cdots, \breve{g}_1, p)}{\sum_{a \in (D_j \cap p)} l_a}. \tag{10}$$

4 Path Observation Generation Algorithm

Path candidates should be generated before the probability can be calculated. The full set of generated paths is enormous, including all paths connecting the domains of all GPS points. In this paper, we define a policy to limit the number of generated paths dynamically at each GPS point. At each GPS point $\breve{g}_j$, a set of best path candidates are chosen according to the cumulative value of their path probability$\Pr(p|\breve{g}_j, \breve{g}_{j-1}, \cdots, \breve{g}_1)$. And for each arc in D_j, at least one path going through it is selected. At the next GPS point$\breve{g}_{j+1}$, by searching from the end nodes of the old paths via shortest path trees, the domain of relevance is determined, and the path candidates are updated by appending the shortest paths to the old path candidates. If an arc satisfy the following criteria, it will be included in the domain of relevance: firstly, the heading difference between recorded GPS data and the arc is less than 60 degrees if the observed speed is greater than $5km/h$, and secondly the degree of relevance

$$f(\breve{g}_j, a) = \int_0^1 f_{\boldsymbol{g},\boldsymbol{\epsilon}^j}(\breve{g}_j, \epsilon_a) d\epsilon_a$$

is greater than a pre-defined threshold. Using shortest path trees is appropriate because the domains' sizes are generally large, so that the generated paths are capable of representing topology of the network. The process of the algorithm is:

1. Initialization:

   (a) Search over the network for the domain of first GPS point $\check{g}_0$;

   (b) Connect arcs in the domain to generate a full set of path candidates P_0;

2. At each GPS point $\check{g}_j$:

   If the GPS point with traveling speed ($\check{v}_j > 5km/h$), tag it as normal speed:

   (a) Search from the end nodes of path candidates (P_{nr}) of previous normal speed GPS point for D_j, with distance limitation $2 \cdot \check{v}_{nr}$;

   (b) Loop over each old path $p_k \in P_{nr}$:

      i. Loop over each segment s connecting from p_k to D_j ;

      ii. If the first arc of s, $a_0 \in D_j$, or a_0 is not the reverse arc of the last arc of p_k, append s to p_k to generate a new path in P_j;

   (c) If L is not empty:

      i. over all arcs in all path candidates, determine the domain of each low speed GPS point;

      ii. empty L;

   (d) If number of paths in P_j is greater than a fixed number:

      i. calculate path probability value for each path candidate;

      ii. from higher to lower, choose path candidates with cumulative probability value of 90%;

      iii. for each arc a in the domain, preserve in P_j at least one path with a as the last arc, according to the probability value;

      iv. update P_j;

   (e) Set $j = j + 1$, and $nr = j$ as index of previous normal speed GPS point go to 2;

   Else:

   (a) Push $\check{g}_j$ to the list of temporal low speed GPS points L

   (b) Set $j = j + 1$, go to 2;

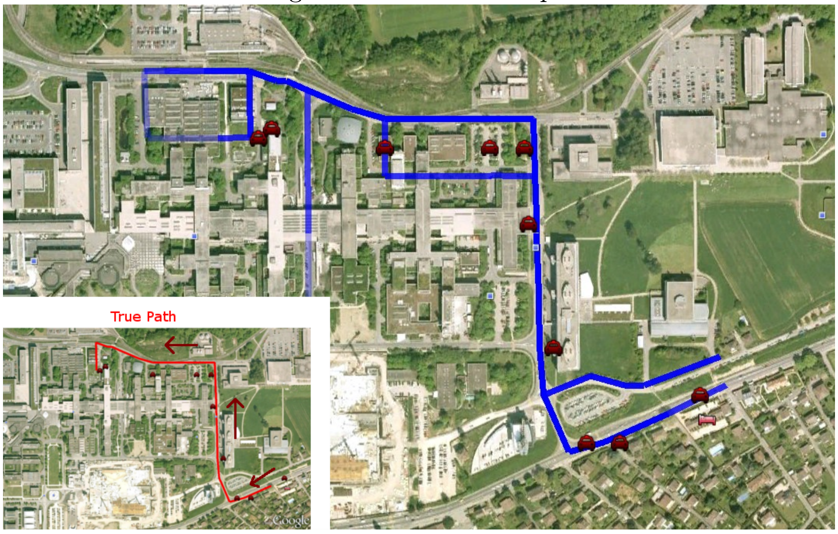
3. Finalization:

   (a) If L is not empty, over all arcs in all path candidates, determine the domain of each low speed GPS point;

   (b) Calculate probability for each path candidate;

   (c) From higher to lower, choose path candidates with cumulative probability value of 90% as final result.

Figure 2: A small example



5 Case studies

5.1 Examples for path observation generation algorithm

In order to illustrate the performance of the algorithm, we apply it to a short real GPS track, which was recorded when a traveler was in a car running in a dense network, see Figure 2. The network data used is downloaded from a free map data source, open street map (`www.openstreetmap.org`). The true path on which the GPS points (red cars) were recorded are indicated in the mini map at the left-bottom corner of the graph. In the main view, blue lines indicate path observations generated. A large number of paths, 19 specifically, are generated, however, most of them overlap over each other, with different origins and destinations but share almost the same intermediate segments.

In the result, the beginning and ending of the trip is ambiguous because the path probability value relies highly on the inter-dependency among GPS records, but there is no information before the first GPS point, and no information after the last GPS point. The problem of origin can be easily solved if the trip is longer, since the latter GPS point strengthen the selection of the entire path, which is one of the characteristics of the path probability algorithm. The ambiguous trip end is an issue of the algorithm if only GPS data is used, because when the user stops traveling, the GPS records will appear as a cloud distributed around the true position. However, the rich available data from the cell phone gives us an opportunity to infer the trip end from the user's daily habits and cell phone's connection to other devices, such as WIFI stations and others' bluetooth devices.

The generated results shows some paths deviate from the true path in the middle of the trip. This is a result of the inaccuracy of the GPS data. The path probability algorithm shows its strength in this case, as it assigns fairly low probability value to those paths. Only 4 out of 21 paths goes through the wrong segment, with probability value 16% in total.

Table 1: Statistics of the trips

	Min	Average	Max
Numbers of GPS points per trip	16	35.6	58
Travel time per trip [second]	179	395	795
Length of generated path (Length) [km]	1.71	3.98	6.44
Number of traffic signals (NbTS)	0	2.86	5.44
Path size (PS)	0.02	0.08	0.77

We further apply the algorithm to two longer trips with 72 and 50 recorded GPS points respectively. Results shown in Figure 3 and Figure 4 are better than the small example. But we observe in 3 a small deviation from main road in some generated paths. Although the calculated probability value for those paths are low (5 out of 33 paths,13% in total), they are still possible to have been the true path. From Figure 4, we can easily recognize the the true path, because those GPS points are relatively accurate and close to the generated path.

5.2 Modeling route choice behavior from real data

The purpose of developing the probabilistic path generation algorithm is that generated path observations can be used in discrete choice models to estimate users' route choice behaviors. We extract from the cell phone database a single user's 19 trips in Lausanne area. A Path Size Logit (PSL) model with network-free data approach (Bierlaire and Frejinger [2008]) is used for the estimation. The deterministic term of the utility function is specified by

$$V_i = \beta_{PS} \cdot \ln PS_i + \beta_l Length_i + \beta_{TS} NbTS_i, \tag{11}$$

in which $Length_i$ and $NbTS_i$ denote the length and number of traffic signals of the path p_i respectively. The path size attribute (PS_i) is calculated by,

$$PS_i^U = \sum_{a \in p_i} \frac{l_a}{l_{p_i}} \frac{1}{\sum_{p_j \in U} \delta_{aj}}, \tag{12}$$

in which U is the path choice set of the origin and destination of path observation p_i, and δ_{aj} equals one if path p_j contains arc a, zero otherwise. 1 shows the statistics of the trips and generated path observations. For each trip, the attributes are weighted by path probabilities.

The choice set is sampled by using biased random walk algorithm (Frejinger et al., 2009) with settings: 50 draws, Kumaraswamy parameters $b_1 = 30$ and $b_2 = 0.4$, length is used as generalized cost for the shortest path computations.

Table 2 reports the coefficient estimates for all attributes. All coefficients have their expected signs and they are all significantly different from zero. The positive value of Path Size coefficient is consistent with the theory (Frejinger, 2008). The negative signs of coefficients for path length and number of traffic signals indicate that the traveler is more willing to choose shorter path, and less traffic signals for his trip, which is reasonable and understandable in the real world.

Figure 3: Longer trips-1

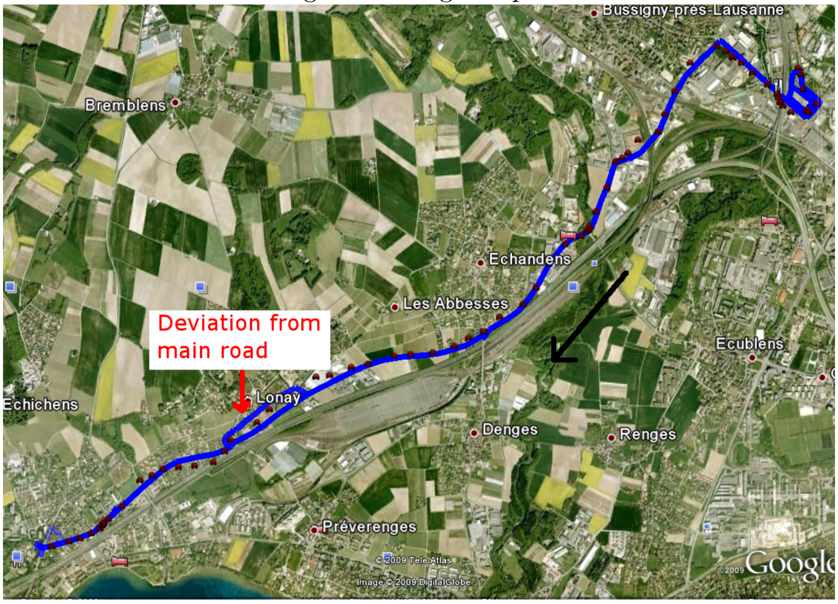


Figure 4: Longer trips-2

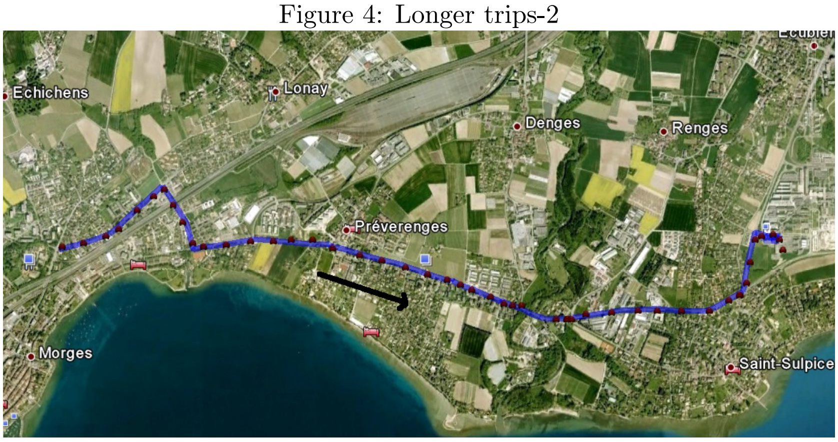


Table 2: Estimation result

Coefficient	Value	Rob. Std. Error	Rob. t-test
ln(Path Size)	3.56	1.656	2.16
Path length	-41.9	15.4	-2.72
Number of traffic signals	-1.23	0.402	-3.05

Number of observations: 21
Null log-likelihood: -81.619
Final log-likelihood: -28.592
Adjusted rho-square: 0.613
Model estimated by BIOGEME (Bierlaire 2007)

6 Conclusions and Future Works

A proposed path observation generation algorithm, supplemented to a path probability measurement, is presented in this paper to generate path observations from GPS data. By applying to real trips, we show the viability of the algorithm and we analyze some characteristics of the algorithm in order to explore any possibility of improving it. The results from longer trips are generally better because more data, and hence more information, can be used by the path probability algorithm to capture the dependency among consecutively recorded GPS data. However, the efficiency of the algorithm is still an issue to be solved for long trips. The accuracy of the result will also need to be examined using data recorded in more complicated situations, such as GPS cloud generated in congested intersections and road segments. Preliminary tests comparing the algorithm with state-of-the-art map matching algorithms reveal that our algorithm is superior in the presence of sparse data. With dense and accurate data, both approaches give results of similar quality. More data and testes will be used for the comparison of the two methodologies.

The path probability can be understood as the probability that the device records the same data if we simulate a travel on a proposed path. A prior knowledge of transportation mode is useful for the simulation because each transportation mode is restricted to use certain types of roads in the transportation network, and different transportation modes have different running patterns. There are already some works on detecting transportation mode purely from GPS data, for instance, Schuessler and Axhausen [2008]. However, the performance of applying such algorithm to cell phone data hasn't been examined. The availability of acceleration data and other data from cell phones provides a prospect of exploring new transportation mode detection methods. For example, it is usually hard to distinguish bus from car, since both of them are motor vehicles running in the same transportation network. However, the cell phone can see more nearby bluetooth devices, i.e. passengers' cell phones, in a bus. Additionally, we can learn from the fact that passengers on a bus changes frequently, leading to the frequent change of nearby bluetooth devices. This will remain to be our further research.

The estimation of route choice behavior from a single users' trips is a preliminary study on applying the probabilistic path observation generation algorithm in discrete choice modeling framework. The estimation result is consistent with the theory and reasonable in terms of estimated behavior. Although the data set is relatively small and utility function is simple, it shows the viability of utilizing state-of-the-art discrete choice modeling methodologies in the estimation of route choice behavior from cell phone data. In the future, we will use the proposed algorithm to generate more path observations from cell phone data, when they are available, to estimate behaviors using more models.

References

M. Bierlaire and E. Frejinger. Route choice modeling with network-free data. *Transportation Research Part C: Emerging Technologies*, 16(2):187–198, 2008.

M. Bierlaire, J. Newman, and J. Chen. A method of probabilistic map distribution of path likelihood. In *Paper Presented in Swiss Transportation Research Conference 2009*, 2009.

M. Flamm, C. Jemelin, and V. Kaufmann. Combining person based gps tracking and prompted recall interviews for a comprehensive investigation of travel behaviour adaptation processes during life course transitions. In *7 th Swiss Transport Research Conference*, 2007.

E. Frejinger. *Route choice analysis: data, models, algorithms and applications*. PhD thesis, EPFL,, 2008.

E. Frejinger, M. Bierlaire, and M. Ben-Akiva. Sampling of alternatives for route choice modeling. *Transportation Research Part B*, 2009.

F. Marchal, J. Hackney, and K. Axhausen. Efficient map matching of large global positioning system data sets: Tests on speed-monitoring experiment in zürich. *Transportation Research Record: Journal of the Transportation Research Board*, 1935:93–100, 2005.

W. Ochieng, M. Quddus, and R. Noland. Map-matching in complex urban road networks. *Brazilian Journal of Cartography (Revista Brasileira de Cartografia)*, 55(2):1–18, 2003.

M. Quddus, W. Ochieng, and R. Noland. Current map-matching algorithms for transport applications: State-of-the art and future research directions. *Transportation Research Part C*, 15(5):312–328, 2007.

M. A. Quddus, R. B. Noland, and W. Y. Ochieng. A high accuracy fuzzy logic based map matching algorithm for road transport. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 10(3):103–115, 2006.

N. Schuessler and K. Axhausen. Map-matching of gps traces on high-resolution navigation networks using the multiple hypothesis technique. Working paper, 2009.

N. Schuessler and K. W. Axhausen. Identifying trips and activities and their characteristics from gpsraw data without further information. In *8th International Conference on Survey Methods in Transport*, Annecy, France, 2008.

P. R. Stopher. Collecting and processing data from mobile technologies. In *8th International Conference on Survey Methods in Transport*, Annecy, France, 2008.