Transportation Research Part B 71 (2015) 120–137

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Probe vehicle data sampled by time or space: Consistent travel time allocation and estimation

Erik Jenelius *, Haris N. Koutsopoulos

KTH Royal Institute of Technology, Dept. of Transport Science, Div. of Traffic and Logistics, Teknikringen 72, SE-100 44 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 2 June 2014
Received in revised form 24 October 2014
Accepted 24 October 2014
Available online 18 November 2014

Keywords:
Travel time estimation
Travel time allocation
Probe vehicle
Floating car data
Sampling
Polling

ABSTRACT

A characteristic of low frequency probe vehicle data is that vehicles traverse multiple network components (e.g., links) between consecutive position samplings, creating challenges for (i) the allocation of the measured travel time to the traversed components, and (ii) the consistent estimation of component travel time distribution parameters. This paper shows that the solution to these problems depends on whether sampling is based on time (e.g., one report every minute) or space (e.g., one every 500 m). For the special case of segments with uniform space-mean speeds, explicit formulae are derived under both sampling principles for the likelihood of the measurements and the allocation of travel time. It is shown that time-based sampling is biased towards measurements where a disproportionally long time is spent on the last segment. Numerical experiments show that an incorrect likelihood formulation can lead to significantly biased parameter estimates depending on the shapes of the travel time distributions. The analysis reveals that the sampling protocol needs to be considered in travel time estimation using probe vehicle data.

1. Introduction

Information about historical, real-time and expected future traffic conditions is critical at all levels of travel planning, traffic management and transport policy. In recent years, GPS devices installed in vehicles and smartphones have emerged as a new type of traffic sensor. These probe vehicle (also called floating vehicle) sensors are often opportunistic in the sense that their original purpose is not to collect traffic data, but have a great potential for cost efficient traffic monitoring. Unlike stationary point sensors (e.g., loop detectors) and point-to-point sensors (e.g., automatic number plate recognition cameras), they can collect area-wide travel time data for any part of the network where equipped vehicles move (Antoniou et al., 2011).

However, a number of limitations mean that new sophisticated methods are needed to process the data and generate useful information, compared to traditional sensors (Leduc, 2008). One challenge is that sampling (also called reporting, or polling) frequencies are often low (less than one per minute), so that vehicles may have traversed significant distances between reports. A consequence of low sampling frequency is that the true path taken by the vehicle may be uncertain. Another consequence is that the spatial resolution of the individual travel time observations is typically lower than the resolution of the network travel time model to be estimated. For the path between two consecutive reports, only the average travel speed of the vehicle is known (derived from the traversed distance and the difference in the time stamps), while the finer details of the vehicle trajectory across different network components (e.g., network links) are unobserved. The travel times of individual network components are thus only measured in aggregated form through space-mean travel speeds.

* Corresponding author. Tel.: +46 (0)8 790 8032; fax: +46 (0)8 212 899.
E-mail addresses: erik.jenelius@abe.kth.se (E. Jenelius), haris.koutsopoulos@abe.kth.se (H.N. Koutsopoulos).

http://dx.doi.org/10.1016/j.trb.2014.10.008
0191-2615/

The travel time estimation problem on probe vehicle data therefore consists of two parts: (i) allocating fractions of the observed travel times to different network components, and (ii) estimating the travel time distributions of the network components. Since the likelihood of a particular allocation depends on the travel time distributions of the components, allocation and estimation is fundamentally a joint problem. A natural requirement of an estimation method is that the true model parameters are recovered as the number of observations increase given that the model specification is correct, i.e., that the estimator is consistent. For consistent estimation, allocation and estimation should be performed simultaneously. Such an approach based on multivariate normal travel time distributions is proposed by Jenelius and Koutsopoulos (2013).

For computational efficiency, however, most proposed methods divide the process into two steps, performed either once or iteratively. First, each probe vehicle travel time observation is decomposed into a travel time for each traversed network component. Second, the travel time distributions for the network components are estimated. Heuristic methods for the travel time allocation are proposed by Hellinga et al. (2008), Zheng and van Zuylen (2013), and Sanaullah et al. (2013). Iterative procedures to estimate the parameters of the network travel time model (using Bayesian and maximum likelihood estimation, respectively), and to decompose the probe vehicle travel times according to the most likely allocation given the current parameter values, are proposed by Hunter et al. (2009), Westgate et al. (2013) and Hofleitner et al. (2012).

1.1. Probe vehicle sampling protocols

The times and locations where probe vehicles are sampled are determined by some form of sampling rule, or protocol. Two main sampling principles can be distinguished: *time-based sampling*, where the vehicle trajectory is sampled at certain time intervals, and *space-based sampling*, where the vehicle trajectory is sampled at certain distance intervals. Being opportunistic sensors, the protocol of the sampling typically cannot be controlled for the purpose of travel time estimation. In this situation, the method of analysis must instead be adapted to the sampling protocol.

Some studies comparing time-based and space-based probe vehicle data exist. As noted by Liu et al. (2007), Westgate et al. (2013) and others, a consequence of time-based sampling is that locations are sampled more densely in parts where the speed is low; with space-based sampling, on the other hand, the sampled locations are independent of the speed. Since congestion detection is an important application of probe vehicle measurements, this has been seen as an advantage of time-based sampling.

Liu et al. (2007) evaluate the benefits and drawbacks of time-based and space-based sampling of probe vehicle data in several dimensions: missing data ratio and GPS errors, number of useful records for travel time estimation, map-matching accuracy, and accuracy in link travel time or speed allocation. The analysis is performed using GPS data from 1570 taxis in Nagoya, Japan; the data set contained probe vehicle data sampled both by time (every 5 s) and by space (every 50 m). Regarding the allocation of the observed space-mean speed to network links, the authors find that time-based sampling leads to somewhat smaller errors when the average sampling frequency is similar. However, definite conclusions are hard to draw from the study since the "true" link speeds were estimated from probe vehicle data with 5 s and 50 m sampling intervals, respectively.

Westgate et al. (2013) present a Bayesian model for estimating link travel time distributions based on GPS data from ambulances, utilizing position and instantaneous speed information and considering that the vehicle path is unknown. The model is applied to vehicle data from Toronto, Canada, and is evaluated against a reference method based only on the speed measurements from the GPS data. In an appendix, the authors show that the inverse of the harmonic mean of instantaneous speeds from probe vehicles is an unbiased and consistent estimator of the mean segment travel time when sampling by space, whereas it is biased upwards when sampling by time. This is a direct parallel to the case with stationary loop detectors, where the harmonic mean of the observed speeds is the appropriate quantity to use for travel time estimation. The paper does not derive an unbiased and consistent estimator of the mean segment travel time (nor other statistics of the travel time distribution) under time-based sampling.

Although the studies above have addressed certain aspects of the impact of the sampling protocol for travel time estimation, the two major problems of (i) travel time allocation and (ii) travel time estimation have not been analyzed systematically under different sampling protocols previously.

1.2. Objectives

The purpose of this paper is therefore to highlight the role of the sampling protocol for consistent travel time allocation and estimation. Using a general travel time model formulation for a network path, it is shown that whether sampling is space-based or time-based determines the proper specification of the likelihood function of the generated data. In particular, time-based sampling induces a selection bias in the observed travel times between two given positions which needs to be corrected for. This bias is not an effect of GPS measurement errors or uncertainty about the true traversed path but an integral feature of the time-based sampling process that exists also when true vehicle positions and paths are perfectly known.

The paper then considers the special case where the travel time model is partitioned into segments with uniform space-mean speed within each segment. Explicit expressions are derived under both sampling principles for the probability distribution of the allocation of the measured travel times among the traversed segments, and the likelihood function of the measurements. In particular, time-based sampling is biased towards measurements where a disproportionally large