Using Floating Car Data in Route Choice Modelling - Field Study

Hekmat Dabbas¹ *, Walid Fourati¹, Bernhard Friedrich³

¹ Institute of Transportation and Urban Engineering, Technische Universität Braunschweig, Hermann-Blenk-Straße 42, 38108 Braunschweig, Germany

Abstract

Route choice assessment is an essential step in classic traffic modeling techniques, which have long been based on theoretical, empirically calibrated models. As floating car data (FCD) are becoming widely available, remarkable ongoing efforts attempt to improve the existing route choice models to produce more accurate estimations. This paper documents a study, conducted with real trajectory data collected on a motorway network in Germany. The aim was to analyze the usage of FCD to estimate route choices on different spatial scales. The logit model was used as a starting point. Then different levels of use of trajectory data in the process of traffic assignment were tested and compared. The validation is made against reference data collected through Automatic Number Plate Recognition (ANPR) cameras on selected routes during consecutive weekdays. The drawback of using the logit model is discussed as well as the advantages of the FCD involvement in the estimation process. We conducted that an aggregated set of FCD can estimate highly accurate and realistic route choice proportions.

Peer-review under responsibility of the scientific committee of the 23rd EURO Working Group on Transportation Meeting.

Keywords: Traffic assignment; Probe vehicles, Study case, real data

1. Introduction

Assigning traffic between an origin and a destination to different alternative routes is one of the essential steps in the practice of traffic planning. Its purpose is to approximate the distribution of traffic volume between routes, which can be used in several traffic management and traffic planning applications. A particular attention has long been given to this task as it affects directly the validity of other models and estimations. Wardrop (1952) was the first to discuss the problem of the route choice by introducing the deterministic user equilibrium principle (UE), stating that a network reaches the UE state when the infrastructure users cannot reduce their travel time by using alternative routes. The deterministic UE choice model was not found to be realistic, as users of the network do not always choose the shortest route. Daganzo and Sheffi (1977) generalized the UE model with their stochastic user equilibrium model (SUE), which takes the perceived travel time as the main factor for their route choice model. More explicitly, travellers choose the route that is perceived to be the shortest in terms of generalized costs. These models rely widely on hypotheses with certain assumptions that can be unidealistic for certain cases which results in unrealistic assignment. The conventional data sources (e.g. detector loops) were not able to cover all necessary information such as trip start/end and route choices. These necessary information can be delivered by e.g. surveys or ANPR devices. However these sources are costly and demanding in terms of man-power and in case of surveys not always up to date. Recently a noticeable effort can be recognised in the field of big data. New available data such as FCD offer the chance to cover missing information which results in mitigating the reliance on some assumptions. FCD act as dynamic sensors which collect many information about the trip such as vehicle positioning records. With data analytic techniques many information such as trip start/end, route choices, travel time and average speed can be extracted.

Matsumoto et al. (2005) used artificial probe vehicles to guess the actual traveled path at the first stage of their two-stage model. In the second stage, origin-destination (OD) flows are modified to estimate the travel demand. Ben-Akiva et al. (2014) used GPS trajectories extracted from taxis to estimate travel time and calibrate the path size logit

* Corresponding author. Tel.: +49-531-391-66817; fax: +49-531-391-66828.
E-mail address: h.dabbas@tu-braunschweig.de

Peer-review under responsibility of the scientific committee of the 23rd EURO Working Group on Transportation Meeting.
model to estimate traffic assignment. FCD usually has low penetration rate, it cannot be used to estimate the route choice directly. This paper presents a comparative case study of using floating car data to estimate traffic assignment. To overcome the problem of the low penetration rate, FCD is aggregated over a longer time period and validated in this paper. The real route choice is extracted from ANPR data and used for the validation. FCD involvement is divided to five levels tested in five analyses. The involvement level at which the model reveals accurate assignment is chosen according to conventional statistical analyses; namely: Root mean square error and Correlation coefficient.

2. Analysis Setup

The network used in the study is located in the highly urbanized metropolitan area of Duisburg, Germany. It is originally extracted from Open-Street-Map and transformed into a graph made of nodes and links.

The reference data used for the analysis originate from a data collection campaign that used (ANPR) cameras. Hashed (anonymized) license plate numbers were collected during three successive weekdays in November 2017 during daytime hours.

The detector data was provided directly by the road authority from its centralized data collection system. The data is aggregated per minute, providing counts and average speed per vehicle class (car/truck). The detector data is only used in this study as input to the U.S. Bureau of Public Roads (BPR) formula.

A commercial provider provided probe vehicle trajectories from different origins e.g. fleet management platforms, taxi-tracking data, navigation devices, and smartphone apps. Trajectories are map-matched to the studied network. FCD is sorted in this research in two categories. FCD data derived from probe vehicles travelled on the link on the same date and time of the ANPR measurement i.e. not aggregated. The other category is FCD derived from the aggregation of the probe vehicles on all regular days of November 2017, which corresponds to the time of the ANPR measurements. The aggregated FCD set was proved to be unbiased.

3. Empirical Analysis

The conducted analyses are scaled from the least to the most involving FCD, which can be used in different ways and at different levels in the process of obtaining route portions between origin-destination (OD) pairs. The purpose of the analysis is to observe the effect of FCD on the route choice problem. The results of the different analysis are compared by means of the statistical methods: Root Mean Square Error and The Correlation Coefficient.

Analysis 1: this is conducted to highlight the effect of FCD in comparison to the conventional statistical methods. Therefore, the Logit model is used in this analysis to estimate the traffic assignment.

Analysis 2: in this analysis the logit model is exclusively responsible for solving the route choice. However, in comparison to the first analysis, the travel time, which is a critical input in the Logit model, is directly obtained from FCD. Another contribution of FCD in this analysis is the calibration of the parameter used in the logit model.

Analysis 3: the OD pairs are split in this analysis into two sets. The first set is made of each OD pair with a minimum count of detected probe vehicles (17 in this analysis). It represents in this case the minimum amount of probe vehicles that each OD pair should have in order to describe the floating data as unbiased data. For this first set, the probability of the route usage is derived directly from the aggregated set of FCD. The rest of the OD pairs are classified in the second set, for which the route choice problem is solved following the same procedure as analysis 2.

Analysis 4: in this analysis all OD pairs are considered to have an unbiased and representative set of FCD. Therefore, the portions of the probe vehicles on a route are recognized as the usage proportion of that route among the available alternative routes of an OD pair.

Analysis 5: this additional analysis is conducted to shed a light on the impact of FCD aggregation on the validity of extracted route portions. Logit model here, as in the previous analyses, is completely neglected. All choice probabilities for analysis 5 are derived from the segregated FCD set.

References


