Route choice behavior based on license plate observations in the Dutch city of Enschede

Tom Thomas

Centre of Transport Studies

University of Twente

Bas Tutert

Centre of Transport Studies

University of Twente

Tom Thomas

Centre of Transport Studies

University of Twente, p.o.box 217, Enschede, The Netherlands

Email: t.thomas@utwente.nl

1 Introduction

Traffic assignment models consists of two main components: routeset generation and route choice behavior [1]. Traditionally, a routeset is generated in advance. Alternative routes are often chosen by Monte Carlo simulations, in which link resistances are changed randomly [2]. However, routeset generation actually depends on route choice behavior. When travelers choose alternative routes, these routes should be included in the route set.

The problem is that route choice is not sufficiently modeled in traditional assignments. The models are often theoretical, and sometimes calibrated by stated preference surveys, but they are seldom validated by oberved route choices. Studies that link observed route choice behavior to underlying attributes, e.g. [3], [4] and [5], are rare. These studies, however, are useful, because they show to what extend real choices depend on 'objective' attributes, like travel time, and to what extend these choices are based on individual preferences of travelers.

When travelers make choices based on individual preferences rather than economics, this will have consequences for traffic loads on the network. In this study, we use license plate data from the Dutch city of Enschede to analyze empirical route choice behavior We generate a route set of many alternative routes and show how the observed distribution of routes depends on the travel times along these routes. From this we can also estimate the effects of individual preferences on route choice and how this influences the traffic loads on the network.

2 Method

This study is based on the registration of license plates at observation posts along all main roads of the city of Enschede during the off peak (14.00 – 16.00h), and evening rush hour (16.00 – 18.00h) on a Tuesday, and during a Saturday afternoon (13.00 – 15.00h). In total, about 26000 observed cases, evenly distributed over these three periods, were used in this analysis.

We defined a link as an imaginary line between two successive (chronological) observations, and a route as a chain of links. When a license plate, for example, is first registered at observation post A, and last registered at B, then the origin-destination (OD) pair for that case is AB. If the car is also registered at C, the route would consist of the links AC and CB.

Journey times were also registered. By aggregating all measurements from a link, average travel times for that link were estimated. These average travel times are quite accurate, because of the large number of cases used in their estimates. The average travel time of a route was obtained by simply adding all the average link travel times of that route.

Our route set contains all observed routes, but also routes that were not observed. In this case, we avoid a bias, because we also include routes that could have been chosen, but which are not in the observed sample by accident. The route set generation was done as follows. The links themselves form the first set of routes and OD pairs. A new route is generated when a new link connects to the previous (sequence of) link(s). By chaining links, the number of routes and OD pairs is extended. This process is stopped until no new routes are formed, or when a route becomes circular (i.e. when one observation post occurs twice in the sequence), or when a (part of a) route is more than 20 minutes longer than the fastest route between the same posts. In total, we generated about 80 routes on average per OD pair.

In theory, we can obtain the frequency distribution of routes per OD pair. However, because frequencies are small for individual OD pairs, we aggregated OD pairs in groups. We distinguish different travel time classes (4 – 7 min, 7 – 12 min, 12 – 17 min, and 17 – 25 minutes) and grouped OD pairs based on the travel time along the fastest route. We discard the class with very short travel times, because this class only contains few OD pairs for which the route choice may very well depend on other attributes. These OD pairs are therefore not representative.

We then estimated the travel time difference between the fastest route and every other route per OD pair. Based on the travel time difference, we grouped the routes in travel time difference classes (0 – 1 min, 1 – 2 min, 2 – 4 min, 4 – 7 min, 7 – 12 min and 12 – 20 minutes difference). Per class, n_1 is the sum of freqeuncies along the fastest routes (the frequencies for all comparisons are added, i.e. if a fastest route is compared with two other routes, its freqeuncy is added twice), and n_2 is the sum of frequencies along the other routes. Similarly, T_1 is the average travel time along the fastest routes, and T_2 is the average travel time along the other routes in the same class. The aggregated frequencies (n_1 and n_2), and average travel times (T_1 and T_2) are the principal parameters, considered in this study.

3 Results

Route choice may depend on many attributes. Their influence is often hardly known, and if known from for example stated preference surveys, most of these attributes, like for example travel time reliability, cannot be estimated in a straightforward way. In applied assignment models, it is therefore often assumed that travelers choose the fastest route, which will lead to an user equilibrium [6]. However, without impicitly modeling individual preferences, their aggregated effect can be taken into account by estimating the relation between the distribution of observed frequencies and real travel times. We did this, and find the following tight relation for the Enschede data.

$$\frac{n_2}{n_1} = \exp[-1.33(T_2 - T_1)^{0.7} + 1.33(0.5)^{0.7}] \qquad \text{voor } T_2 - T_1 > 0.5 \text{ min} \quad (1a)$$

$$\frac{n_2}{n_1} = 1 \qquad \text{voor } T_2 - T_1 \leq 0.5 \text{ min} \quad (1b)$$

This relation is more or less valid for all travel time classes. We therefore conclude that the route choice probablity depends on absolute rather than relative travel time difference. Equations (1a) and (1b) are also valid for the different periods (off peak, rush hour and Saturday afternoon), although the average travel times are different, e.g. the average travel time is on average 10% larger during rush hour than during the off peak.

The calculation of the choice probabilities is now straightforward. If for example an OD pair has two alternatives that are 1 and 3 minutes longer than the fastest route, then according to equation (1a), $n_2/n_1 = 0.60$ and $n_3/n_1 = 0.13$. Thus, in that case, the assigned fractions over routes 1, 2 and 3 are 58%, 35% and 7% respectively.

The probability that a longer route is chosen, declines rapidly with travel time difference. However, because there are many alternatives, a significant fraction of 25% (of the observed cases) does not follow the fastest route. Regarding the network performance, the detour time over the alternative routes is an even more important parameter. This 'occupancy measure' indicates how much of extra load the network has to process due to detours. We find that the average detour time (a combination of number of alternative routes, route choice probability and travel time difference) is maximal for alternatives with a travel time difference of about 5 minutes with respect to the fastest route.

The average detour time (aggregated over all routes) is 0.57 minutes. This is about 8% of the travel time (the average travel time along the fastest route was 7.5 min), which implies that the network has to process 8% of extra load compared to a traditional equilibrium assignment.

This study is based on observed route choice behavior in Enschede. It would be useful to compare these results with observations from other cities, and with observed route choices on highways. Because the number of alternatives is lower for a highway network, we expect that detour factors should also be lower, and thus that traditional assignments will probably show more reliable results for highways. However, this can only be validated by route choice and travel time observations from highways

As mentioned before, most other attributes cannot be estimated in a straightforward way. The hierarchy of roads may be one of the few measurable attributes that has an effect on route choice. Small roads with speed bumps are less comfortable than highways (without congestion). Thus, given similar travel times, it is quite likely that travelers prefer larger roads. Another factor that can be taken into account, is the rate of overlap between routes, e.g. two almost identical routes may be seen as one route by the traveler. The hierarchy of roads and the overlap of routes are to be analyzed in a follow-up study.

Including other attributes, such as those mentioned above, may significantly improve route choice predictions for individual cases. However, it is not likely that this will have an effect on equation (1). In fact, we do not expect that we have introduced a systematic bias by not including other attributes, because they are implicitly in the observed choices, and also not correlated with travel times of different alternative routes. We therefore think that our simple route choice model can already be used to improve traffic assignments in a structural way.

References

[1] P.H.L. Bovy, "On Modelling Route Choice Sets in Transportation Networks: A Synthesis", *Transport Reviews, 29*(*1*), 43-68 (2009)

[2] Y. Sheffi and W.B. Powell, "An algorithm, for the Equilibrium Assignment Problem with Random Link Times", *Networks I2*(*2*), 191-207 (1982)

[3] D. Papinski and D.M. Scott, "GIS Toolkit design and application for route choice analysis: a comparison of observed routes and shortest paths", in TRB proceedings (2009)

[4] R.P. Batley and R.G. Clegg, "Driver Route and Departure Time Choices: The Evidence and the Models", Proceedings of the 33rd Annula UTSG conference (2001)

[5] J.R. Duffell and A. Kalombaris, "Empirical Studies of Car Drive Route Choice in Hertfordshire", *Traffic Engineering and Control, 29* (*7/8*), 398 – 408, 1988

[6] J.G. Wardrop, "Some theoretical aspects of road traffic research, *Proceedings Institute of Civil Engineers*, *part II* (*vol. 1*), 325-378 (1952)