Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 14 (2016) 1716 – 1722

Transportation Research Procedia

www.elsevier.com/locate/procedia

6th Transport Research Arena April 18-21, 2016

Uncertainty of the od matrix's estimation in urban public transport

Balázs Horváth [a,*]

[a]Széchenyi István University, Egyetem tér 1. Győr 9026, Hungary

Abstract

The key of the planning of public transport systems is the accurate prediction of the traffic load which requires a well-functioning assignment method and reliable passenger data (i.e. time-dependent OD matrix). The paper two years ago at TRA2014 was started with this idea. Since that time the procedure was improved and real network test was performed. Based on the experiences of the real network tests several new improvements were implemented. This paper shows the newest results with the previously described method and search for connection between network structure and certainty together with sample size.

Peer-review under responsibility of Road and Bridge Research Institute (IBDiM)

Keywords: Public transport; OD matrix; sample size; modelling; travel demand estimation

1. Introduction

Planning of transport system requires the knowledge of travel demand. We have to distinguish between car traffic and public transport. Majority of car traffic measures operates with local measures as written by Polgár et al. (2013) and in case of network wide modification we can use automatic car recognition system. By public transport there is no automatic travel demand recording (only E-ticketing, but it is not widespread in Hungary). On the other hand the characteristics of travel demand are not the result of a simple mathematical equation, but a series of multi-criterion decisions as described by Csiszár (2013).

* Corresponding author. Tel.: +36-96-503-494; fax: +36-96-613-561.
E-mail address: balazs.horvath@sze.hu

Peer-review under responsibility of Road and Bridge Research Institute (IBDiM)
doi:10.1016/j.trpro.2016.05.137

The Department of Transport at Széchenyi István University is dealing for more than 30 years with public transport planning studies. During these studies there is always a big problem to have reliable passenger demand data. To solve this question the researchers group developed several methods to estimate OD matrix like Horváth et al. (1998) or Horváth (2012) or Horváth et al. (2014). The basis of the lately used method is back to the '80's when Prileszky (1995) described the original method. Since that time computation power and software environment was simpler the method was not able to test on real networks. Lately the method was improved and due to the new computation power and supporting software, the developed method is already used as real network test as described in Horváth et al. (2014).

At the other hand matrix estimation and assignment have strong connection due to the fact of complex decision process behind the route choice of the passengers as Winkler (2013) stated in his PhD thesis, this strong connection have a third player as well: the network model. As in Horváth (2013) stated the mentioned three players: network model, OD matrix and assignment method (included route choice) form together a system it is not possible (under realistic circumstances) to separated them. We are not able to check network model on its own, we are not able to proof the OD matrix as a stand-alone matrix and finally we cannot use assignment without the previous twos. Therefore the developed OD matrix estimator is only part of the game, although a very important part, due to the fact that to build up network model and to use a given assignment model is simpler than to get an OD matrix. It is true that to choose the right assignment method is not an easy task, but as it written in several papers modelling is a GIGO system, it means if the input data is wrong (like Garbage), than the results will be also wrong independently from the used assignment method.

The last five years we did huge steps to improve our matrix estimation method which now seems to be reliable enough at small or medium cities.

2. Theoretical background

The background of the developed matrix estimation method was already described in Horváth et al. (2014), although I show it here again to give a clear picture, how the later shown result was reached. The origin of the matrix estimation procedure shown in this paper was developed by Prileszky (1995). The original method, which previously sunk into oblivion was revised and improved. This section describes the operation of this method.

2.1. Theory of the matrix estimation method

Considering the fact, that the budget and elaboration time of short-term and medium-term planning is restricted, there is no possibility of detailed data collection in every case. At the majority of the projects, however, there is a full-scale cross-section passenger counting and OD survey sample during the preparation of the planning.

The essence of this method is, that the data gathered in the course of the passenger counting let us be allowed to correct the elements of the sample matrix gained from the survey.

It is possible to make use of the data of the passenger counting in two ways. On the one hand these give information about the number of boarding and alighting passengers at each stop; on the other hand these give information about the travelers' number between stops.

These data can be produced with the use of a suitable flow assignment procedure. If we compare the obtained data from the two sources (counting, assignment) we have an opportunity to correct the model matrix. In order to do this correction, we have to use the boarding and alighting numbers and the numbers of passengers between stops.

2.2. Details of the method

The sums of rows and columns of the OD matrix that is the number of departing and arriving passengers of each zone equals to the sum of the boarding and alighting passengers of the stops in the given zone. Because all of this, the sums of rows and columns of the OD matrix (to be determined) are known on full scale cross-section passenger counting. These are utilizable as target values through the following calculations. Boarding and alighting numbers compared to the sums of rows and columns of the model matrix expose the difference between the model (seed) and target matrices, which difference can be corrected by row and column factors. It is especially important to notice that the p_i and a_j values refer to departing and arriving passengers in the model matrix, however, the P_i and A_j values are the boarding and alighting numbers of the passenger counting, that is the latter contains transfers. Taking account this problem the planned sums of rows and columns have to be corrected before the above mentioned factor calculation. This means that the sums of rows and columns of those relations where the transfer occurs have to be corrected.

The basis of the transfer correction is the following connection:

$$\frac{p_i}{p_i + nt_i} = \frac{P_{korr,i}}{P_i} \tag{1}$$

where nt_i number of transfers in stop i.

By this correspondence it is possible to calculate the corrected target row and column sums.

$$P_{korr,i} = P_i \cdot \frac{p_i}{p_i + nt_i} \tag{2}$$

$$P_{korr,i} = P_i \cdot \frac{p_i}{p_i + nt_i} \tag{3}$$

$$P_i = \sum_{n=1}^{k} F_{i,n}; \quad A_j = \sum_{m=1}^{k} L_{j,m} \tag{4}$$

where
k number of stops in the whole area
$F_{i,n}$ boarding in n. stop of i. zone
$L_{j,m}$ alighting in m. stop of j. zone

Executing this correction step the row and column factors can be calculated as the following:

$$r_i = \frac{P_{korr,i}}{p_i} \tag{5}$$

$$c_j = \frac{A_{korr,j}}{a_j} \tag{6}$$

where
r_i i. row's factor
c_j j. column's factor

However, as it was mentioned, this matrix estimation method doesn't use only stop point data. It takes into consideration the data of passenger numbers between stops to achieve adequate precision. These are called link data, factors deriving from them are link factors. Preparation of these link factors is more complex task than of the row and column factors taking into consideration that the model link data are unknown. To determine model link data it has to assign the model matrix onto the transport network. The generated link loads give the model link data, which can be compared to the counted number of passengers between stops.

$$link_{i,j} = \frac{real\ link\ load_{i,j}}{computed\ link\ load\ \ from\ assignment_{i,j}} \tag{7}$$

After calculating the factors of every row, column and link, they have to be linked to each travel relation. Or vice versa, we have to collect the row, column and link factors of given relations for every single element of the OD matrix. While the determination of the first two is quite simple as departing and arriving zones are known in every relation, association of link factors needs another single study.

To associate the link factors, it has to know the shortest path or k shortest paths in the given relation and it has to associate all the links covered by these paths to the given relation.

After execution of the abovementioned steps, factors needed for following calculations are known regarding all relations. But these factors can be used in different ways:

- only the shortest path
- k. shortest paths
    - link factors associated to paths and weighted by assignment
    - link factors associated to paths (called later method b)
    - every factor weighted equally (called later method c)

Based on the experiences of the previous papers like Horváth et al. (2014) some of the mentioned methods can be eliminated. The practice showed that only the k. shortest path with "link factors associated to paths" and "every factor weighted equally" is well usable.

3. Latest real network test with the matrix estimator

The implemented matrix estimator works in close cooperation PTV's VISUM software, where VISUM is the databank and gives the assignment process, while the external matrix estimator calculates in several steps the OD matrix.

3.1. City of Eger

There are some test results for the Hungarian city Eger. The city has 55.000 inhabitants. The public transport network has 77 stops and 41 lines. The basis of the calculation is a full-scope cross-section counting done in 2012.

Figures 1. and 2. show the connection between link load counted at the passenger counting and link load calculated during the assignment based on the estimated OD matrix with method b and c. As it can be seen method b has a correlation coefficient of 0,95 while the method c only 0,91. Both are very high value.

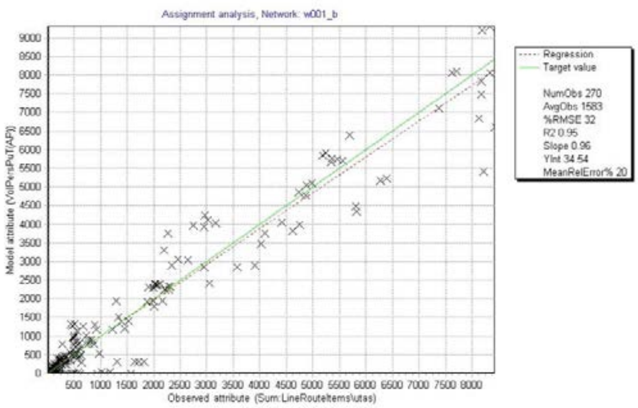

Fig. 1. Connection between observed and modelled values at the Eger network with method b.

In the graph it is clear that method c overestimated some of the link loads, especially at lower values. It can be the cause of the worse results.

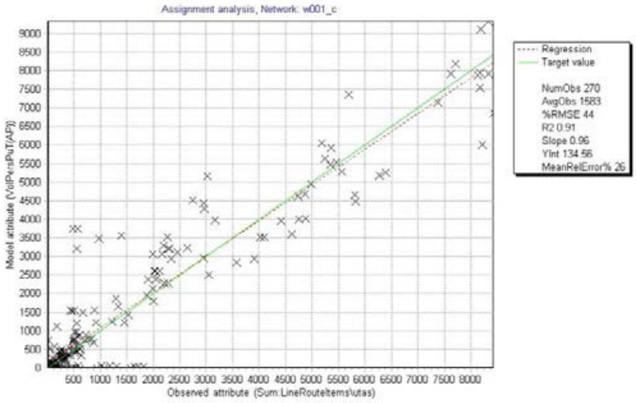

Fig. 2. Connection between observed and modelled values at the Eger network with method c.

3.2. City of Dunaújváros

There are also some test results for the Hungarian city Dunaújváros. The city has 40.000 inhabitants. The public transport network has 89 stops and 43 lines. The basis of the calculation is a full-scope cross-section counting done in 2008.

Figures 3. and 4. show the connection between link load counted at the passenger counting and link load calculated during the assignment based on the estimated OD matrix with method b and c. As it can be seen method b has a correlation coefficient of 0,98 and method c has the same valus of 0,98.

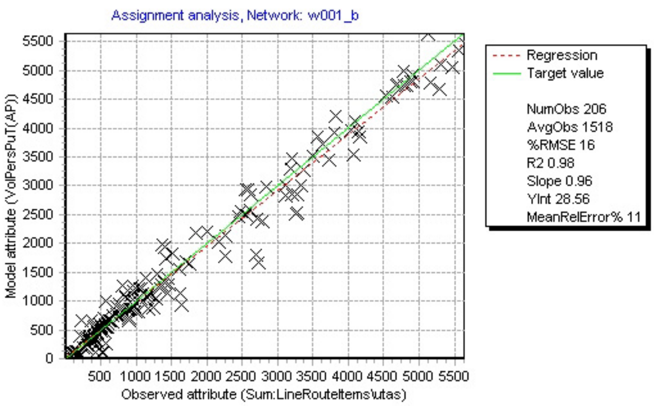

Fig. 3. Connection between observed and modelled values at the Dunaújváros network with method b.

In the graph it is clear that method both estimation are very good maybe at method b there is an underestimation at middle values.

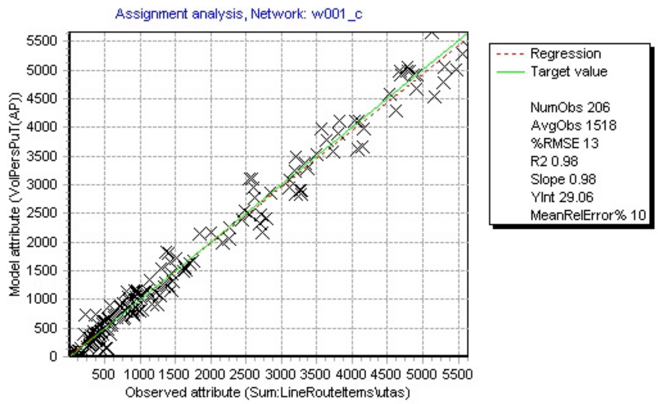

Fig. 4. Connection between observed and modelled values at the Dunaújváros network with method c.

4. Experiences of the test runs

I expected (before both set of test runs was finished) a clear picture on which method is the clear winner and what is the suggested sample size to have a reliable assignment result. Due to the fact that the first expectation is not fulfilled I am not able to answer to the second question. It is interesting that at first look there is no real connection between network characteristics like number of stops or number of lines between the tightness of the connection between observed and modelled values.

In the case of Eger method b is much better while at Dunaújváros there is no clear winner, maybe method c where the slope is 0,98 and relative error is only 10%. Based on these two cities I am not able to define which method is better or What kind of circumstances are needed to define in advance which method will be better. To decide between the two methods I have to perform more and more test runs. The bottleneck is here the lack of passenger counting data. There is a full set of passenger data for the Hungarian city Győr. The logical step would be to have test run on the public transport network of Győr. The only problem is that the city and network is approximately three times bigger than the networks studied till now. As an effect of this bigger size the runtime of the estimation method is much higher. At the first few experiences show an estimated runtime for Győr 50-60 hours. Therefore we have to go back to the previous stage: software implementation. It is important to find the points at the software code where it is possible to speed up the process.

5. Conclusions

The case of both cities showed the matrix estimation method is working. It can produce reliable OD matrix for planning purposes, although it is not clear which calculation method is better in a given case. The tasks of the next period are to let run estimation in more and more cities to collect more experiences with the connection between city and method.

References

Csiszár, Cs., 2013. Model of Multimodal Mobility Coordination and Guiding System. International, Journal of Engineering and Innovative Technology (IJEIT) Vol.3, Issue 6, pp. 125-132. (ISSN 2277-3754)

Feldman, O., 2012. The GEH measures and quality of the highway assignment models. Association for European Transport and Contributors

Horváth, B., Horváth, G., 1998. Methodology of Public Transport Planning. European Society of Transport Institutes (ESTI) Young researcher's conference, Paris

Horváth, B., 2012. A Simple Method to Forecast Travel Demand in Urban Public Transport. Acta Polytechnica Hungarica Vol. 9. No. 4., 165-176

Horváth, B. 2013. Közforgalmú közlekedési rendszerek modellezésének kérdései az utazási igények oldaláról (habilitation thesis in Hungarian) Széchenyi István University, Győr, Hungary

Horváth, B., Horváth, R., Gaál, B., 2014. A new iterative method to estimate origin-destination matrix in urban public transport. Transport Research Arena 2014 Paris, France

Horváth, B., Horváth, R., 2014. The application of a new iterative OD matrix estimation for urban public transport. The 14th International Conference Reliability and Statistics in Transportation and Communication (RelStat'14), Riga, Latvia (ISBN: 978-9984-818-70-2)

Polgár, J., Tettamanti, T., Varga, I., 2013. Passenger number dependent traffic control in signalized intersections. Periodica Polytechnica Civil Engineering. 57/2 pp. 201-210.

Prileszky, I., 1995. Közlekedéstervezés (Transport planning a student book in Hungarian), Széchenyi István University, Győr, Hungary

Winkler, Á., 2013. Utazói döntések modellezése a városi közforgalmú közlekedésben. (PhD thesis in Hungarian) Széchenyi István University, Győr, Hungary