Estimation of Passenger Demand in Urban Public Transport

B. Horváth, R. Horváth, B. Gaál

Széchenyi István University, 9026 Győr, Egyetem tér 1.
Phone: 503 400, fax: 613 561
e-mail:hbalazs@sze.hu,ricsi@sze.hu, gaalb@sze.hu

Abstract: This article shows a renewed, improved, practical variant of a former theoretical matrix estimation method. The authors will demonstrate the method's theoretical and practical sides on a test network.

Keywords: public transport, travel demand, OD matrix estimation, VISUM

1. Introduction

The principle of every transport planning task is the knowledge of travel demands. It is especially in public transport, where demand and supply heavily related. As the operation of public transport financed from common sources, it is utmost importance, that the creation and development of the public transport system must be based on objectivecriteria. Estimation of transport demand has long history, the authors also have dealt with the problem in the quoted publications [8][7] earlier.

Former procedures usually oversimplified the complex issue of route selection or contained an excessively high number of non-automatable, manual steps making impossible the modelling of extensive and complex networks [16].

In the first part we show the accomplishments of this field, in the second part we specify the theoretical background of our procedure, the third part contains the embedment of the procedure into a state-of-art transport planning software. Finally we prove the method's applicability in practice one test network.

2. Previous accomplishments

The bibliography of the matrix estimation methods is abundant. Even so, only a little portion of these deals with public or multimodal transportation, the great majority is about private transportation.

In the last few decades the 4-step (or 4-stage) model has become the most recognized transport planning procedure. In this method travel demand can be determined during a four step process. The first step is trip generation. Homogeneous areas, zones have to be defined on the basis of different aspects and the departing and arriving traffic of these zones have to be given. Consequently, the sums of the rows and columns ofan OD (origin-destination) matrix have to be defined. The next step is modal split. Practically, so many OD matrices have to be done as many transport modes are actually studied. In the third step (trip distribution) the OD matrix is filled, the from-to traffic is determined. Finally, the last step is the flow assignment, when demand is assigned onto the transport network.

One of the most problematic parts of the model is the filling of the OD matrix. Detailed data are needed for the knowledge of the accurate demand. Traditionally, these data were obtained by home interviews and roadside surveys. However, this kind of data collection is expensive in terms of money and manpower and at the same time it contains internal sources of error because of sampling processes and problems in elaboration [13].

Hence, professionals mostly apply some kind of matrix estimation method. Through these methods relatively precise estimations can be made by considering traffic counts and other information. Advantages of traffic count are less expenditure (money, manpower) and suitability to follow the changes of demand in time. Also, it is developable and automatizable by means of modern technology.

Among OD matrix estimation methods there are several approaches.

Van Zuylen and Willumsen suggested the application of an entropy maximization method for un-congested networks [15]. Partially, this was the base of the most likelihood estimation [11][2], the least squares estimation [3][1] and Bayesian framework methods [11]. These approaches were generalized and enhanced by other authors [17][10][6].

Beside static estimations, which assume time-independence, dynamic estimations have evolved. By the help of these time-varying OD flows and matrices can be determined using time-varying traffic counts [4]. On the field of the usage of dynamic estimators there are two main approaches. Simultaneous estimators give estimations jointly for all OD matrices for all time slices using every obtained data set. Sequential estimators ensure computational advantages since they disassemble the whole optimization problem into more manageable pieces and give the possibility of using the estimates for an interval as a priori estimate of subsequent intervals [12].

Matrix estimation methods are tightly connected to flow assignment methods. It is possible to verify the retrievability of the counted traffic with the estimated data by assigning the OD matrix onto the transport network. When there is a significant difference between the counted and computed values, the results can be refined by subsequent iteration steps.

The abovementioned methods can be used for estimating public transport demand also. Li and Cassidy created an algorithm to estimate the OD matrix of transits by boarding and alighting data on (bus) stops [9]. The number of boarding and alighting passengers at a stop can be counted simply and solely. These data give the sums of rows and columns of the OD matrix. These sums can correspond to many OD matrices, the most appropriate must be selected out of them. To get the most likely OD matrix another seed-matrix (or base matrix, model matrix) is needed, which contains the travel preferences of passengers. (Generally, this is an older OD matrix.) In the algorithm, the authors differentiate „minor” stops like the ones in residential areas and „major” stops, the ones that serve activity centres such as large business, multimodal transit point etc. On this basis, different likelihoods can be assigned to the trips between single stops.

Obviously, a trip between two residential stops has less likelihood than a trip between stops of residential and industrial or service areas (i.e. trip between home and workplace) on a common weekday.

The algorithm suitable to assign not only the OD matrix of the passengers on the examined route, but also the likelihood of alighting in each stop. Since this latter characterize the destination of trips, it is assumable that it remains approximately the same in case of transit trips under the same circumstances. Utilizing this assumption the OD matrices of future travels can be estimated more precisely.

OD matrices of public transport can be used for example to evaluate the service quality of public transport networks and to compare them. Good example is the survey of the bus network of Porto, Portugal by Guedes et al. They made an OD matrix from data of electronic ticket validation (utilizing the fact that most of the passengers also travel back). The parameters of service quality like travel time, number of transfers, waiting time etc. easily computable in view of the matrix and the network [5].

Transport planning in the field of public transport became accentuated subsequent upon the problems caused by the intense growth of urban traffic and increase in demand for liveable environment. The large-scale headway of private transport in cities and the increasing demand for mobility forced positive discrimination of public transport upon the decision-makers to handle the problems. To substantiate decisions and strategies connected to these challenges theoretical and methodological bases are needed, therefore researches will be more frequent in this field of science expectedly.

3. Theoretical background of the matrix estimation method

The basis of the matrix estimation procedure shown in this article was developed by Prileszky [14]. This is the revised, improved version of the original method sunk into oblivion.

3.1. Theory of the matrix estimation method

As mentioned before, the basis of the transport planning is the knowledge of transport demands. 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. In case of most 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 travellers' 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.

3.2. Operation of the model

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 numbersof 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:

$$\frac{p_i}{p_i + \acute{a}t_i} = \frac{P_{korr,i}}{P_i} \tag{1}$$

where $\acute{a}t_i$ numbers of transfers in i. stop

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

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

$$A_{korr,j} = A_j \cdot \frac{a_j}{a_j + \acute{a}t_j} \tag{3}$$

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

where k number of stops in studied zone

$F_{i,n}$ boardings in n. stop of i. zone

$L_{j,m}$ alightings in m. stop of j. zone

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

$$s_i = \frac{P_{korr,i}}{p_i} \qquad (5)$$

$$o_j = \frac{A_{korr,j}}{a_j} \qquad (6)$$

where s_i i. row’s factor

o_j j. column’sfactor

However, as we mentioned, this matrix estimation method doesn’t use only stop point data. We take 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 we have 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.

$$él_{i,j} = \frac{real\ link\ load_{i,j}}{computed\ link\ load\ \ from\ assignment_{i,j}} \qquad (7)$$

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, we have to know the shortest path or k shortest paths in the given relation and we have 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
    - every factorweighted equally
    - link factors associated to paths
    - link factors associated to paths and weighted by assignment

We show the differences further on the model network.

4. Matrix estimation method in the practice

The key for the functional correctness of the matrix estimation method is the accuracy of the applied flow assignment method. To aid the future improvementof the method, we fitted the procedure to the VISUM transport planning software of the PTV AG.

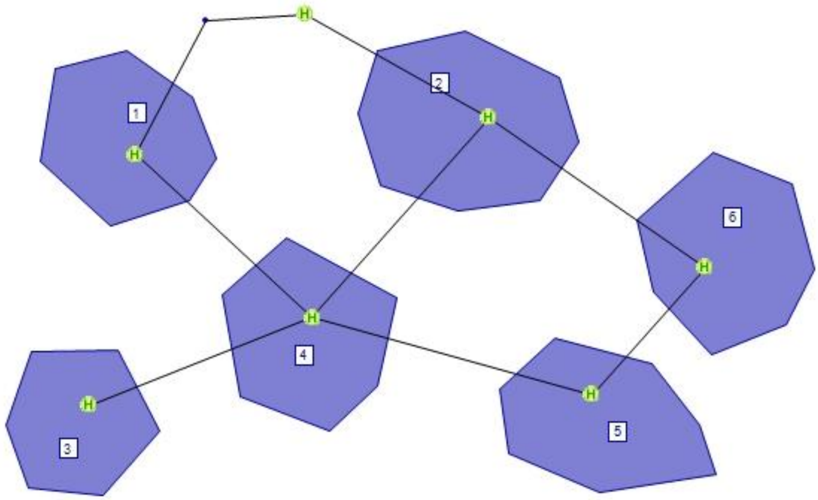

Figure 1.The test network

Applying this software was obvious, as the VISUM is one of the market leading productsin this field and we use it for ten years. Another asset of the program is the cooperation with numerous program languages through Windows COM protocol. This possibility helped the development of the method. In the first phase of the development we used the Python script based language. The third reason was that the VISUM can be used as [transport] data bank for containing different network data.

The matrix estimation procedure is the following:

The matrix estimation method reads out the model data built up previously in VISUM through the COM interface then process it. The readout contains the results of the base assignment, executed previously.

After the former computing method the Python script modifies the OD matrix and writes it back into the transport model. Then another monitoring assignment and assignment analysis are executed (which is part of the VISUM).

The analysis shows how accurate the results - we get from the OD matrix – transport network – flow assignment method trio - are. In case of necessity the whole process can be repeated. According to our tests so far, about 30-50 iterations are needed to calculate the right matrix in cases of minor networks.

5. Practical example, application verification

We made a model network to verify the functioning of the method. Out of the former four approaches we tried three pieces. We didn't deal with the shortest path approach only, assuming that it doesn't give adequately realistic results.

5.1. Application results: k shortest paths, every factorweighted equally

In this approach we took into consideration every factor belonging to certain relations equally. So the row and column factors and every link factor concerned with this relation were taken into consideration with simple averaging.

After implementing the iteration method 50 times according to the results of the assignment analysis (Fig. 2.) the measured and calculated values were bordered on each other, as the correlation factor was R^2=0,97, while the relative error was only 6 percent.

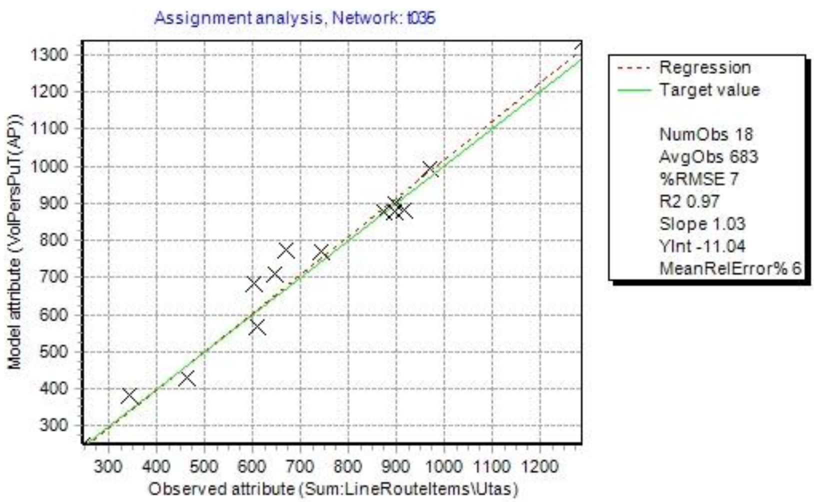

Figure 2. Assignment analysis in case of simple mean, after 50 iterations

At the present state of the research, this method gives the best results. However, according to our assumption this is the result of a calculation error. We will specify this error later.

5.2. Application results: k shortest paths, link factors assigned to path

In this approach we took into account every one of the links assigned to the examined path. Thus, we got a link factor mean for every single path. Averaging these with the row and column factors again, we got the characteristic factor of the relation.

We have executed this method with 50 iterations too. The results can be seen in Fig. 3. The analysis shows slightly worse results compared to the first case.

The cause of the deterioration in our opinion is the averaging of the link factors of every path (by path) thus the links concerned with many different paths fell into the calculations multiple times, disfiguring the correctness of the factor relevant to single relations.

At the same time, it is worth considering whether the whole approach is incorrect or only the multiple accounting of links in simple model networks caused the errors.

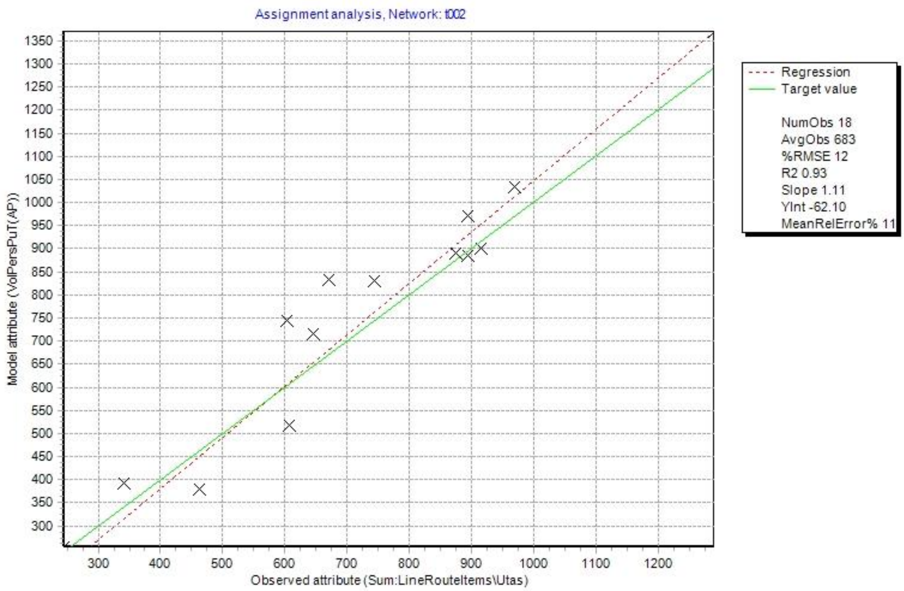

Figure 3.Result of assignment analysis in case of means of paths without weighting

5.3. Application results: k shortest paths, link factors assigned to path and weighted by result of assignment

This third approach practically the same as the latter. The only difference is that the means of the link factors assigned to single paths are weighted by the assignment proportion calculated in the course of the flow assignment. This means that the most preferential path gets the highest factor, while an inferior path in the given relation has lower factor.

We have also executed this method with 50 iterations. The results are in Fig. 4.

This approach conducted the worst results, what is the opposite of prior anticipations. Presumably this is connected to the formerly discussed problem of the links taken into consideration multiple times.

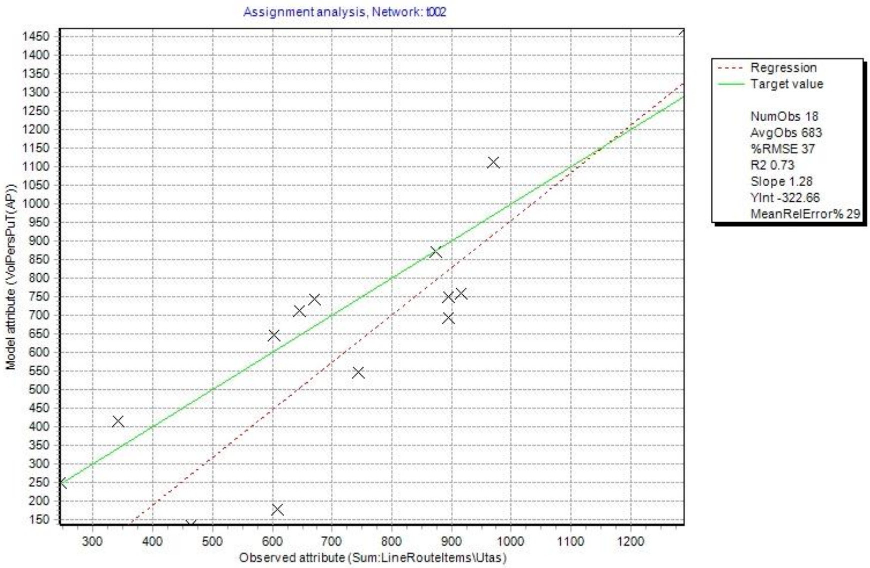

Figure 4. Result of assignment analysis in case of means of paths with weighting

6. Conclusion

The first tests with the evolved matrix correction procedure show that this method is capable to correct a model matrix by the help of cross-section methods. However, uncertainties emerged during the tests pointed out the necessity of improving the application of factors to improve the correctness of the result.

Another observation was that the VISUM program is suitable to the actuation of the method, however the Python as script language has inadequate speed, thus the authors couldn't test the method on real networks, as the computing demand is approximately 300000 times higher in the case of Győr than of the model network.

To eliminate this problem, the next step of the development is the transcription of the program code into C or Pascal environment.

7. Acknowledgement

„TÁMOP-4.2.2.C-11/1/KONV-2012-0012: "Smarter Transport" - IT for co-operative transport system - The Project is supported by the Hungarian Government and co-financed by the European Social Fund"

References

[1] Bell, M.G.H.: *The estimation of origin-destination matrices by constrained generalized least squares*, Transportation Research Part B 25, pp. 13-22, 1991

[2] Bell, M.G.H.: *The estimation of origin-destination matrix from traffic counts, Transportation Science*, vol. 10, pp. 198-217, 1983

[3] Cascetta, E.: *Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator,* Transportation Research Part B 18, pp. 289-299, 1985
[4] Cascetta, E., Inaudi, D., Marquis, G.: *Dynamic estimators of origin–destination matrices using traffic counts,* Transportation Science vol. pp. 363–373 1993
[5] Guedes, M., Oliveira, N., Santiago, S., Smirnov, G.: *On the evaluation of a public transportation network quality: Criteria validation methodology,* Research in Transportation Economics, vol. 36, pp. 39-44, 2012
[6] Hazelton, M. L.: *Estimation of origin-destination matrices from link flows on uncongested networks* ,Transportation Research Part B 34, pp. 549-566, 2000
[7] Horváth, B.: *A Simple Method to Forecast Travel Demand in Urban Public Transport,* Acta Polytechnica Hungarica , vol. 9, no. 4., pp. 165-176, 2012
[8] Horváth, B., Horváth, G.: *Methodology of Public Transport Planning*, European Society of Transport Institutes (ESTI) Young researcher's conference, Paris, (2-3, October, 1998
[9] Li, Y., Cassidy M.J.: *A generalized and efficient algorithm for estimating transit route ODs from passenger counts,* Transportation Research Part B 41, pp. 114-125, 2007
[10] Lo, H., Zhang, N., Lam, H.: *Estimation of an origin-destination matrix with random link choice proportion: a statistical approach*, Transportation Research Part B 30, pp. 309-324, 1996
[11] Maher, M.J.: *Inferences on trip matrices from observations on link volumes: a Bayesian approach,* Transportation Research Part B 20, pp. 435-447, 1983
[12] Marzano, V., Papola, A., Simonelli F.: *Limits and perspectives of effective O-D matrix correction using traffic counts,* Transportation Research Part C 17, pp. 120-132, 2009
[13] Perrakis, K., Karlis, D., Cools, M., Janssens, D., Vanhoof, K., Wets, G.: *A Bayesian approach for modelling origin-destination matrices,* Transportation Research Part A 46, pp. 200-212, 2012
[14] Prileszky, I.: *Közlekedéstervezés*, Széchenyi István College, Győr, 1995
[15] Van Zuylen, H.J. , Willumsen, L.G.: *The most likely trip matrix estimated from traffic counts, Transportation Research Part B 14B* pp. 281-293, 1980
[16] Winkler, Á.: *Experiments to Discover Passenger Preferences, in Public Transport Mobilita'11: Sustainable Mobility,* Bratislava, 26-27.05.2011, Slovenská Techniká Universita, pp. 136-141, 2011
[17] Yang, H., Sasaki, T., Iida, Y., Asakura, Y.: *Estimations of origin-destination matrices from link traffic counts on congested networks,* Transportation Research Part B 26, pp. 417-434, 1992