

# The Origin–Destination Matrix Development

Vasile Dragu<sup>1\*</sup>, and Eugenia Alina Roman<sup>1</sup>

<sup>1</sup>University Politehnica of Bucharest, Faculty of Transports, Splaiul Independenței, No. 313, 060042, Bucharest, Romania

**Abstract** Transport studies are conducted for a better understanding of the actual mobility and for developing transport forecasting models to predict the future transport demand and the changes in travel patterns. Transport planning involves the decision-making process for potential improvements to a community's roadway infrastructure. The first transport models used to analyze globally the transport system requirements while nowadays models were rethought as a demand – supply interaction reflecting the correlation between transport and socio-economic development. The transport forecasting methodology use a four stage structure consisting of: trip generation, trip distribution, modal split, traffic assignment. In the second stage of the model, the generated trips for each zone are distributed to all other zones based on the choice of destination. The trip pattern is represented by means of an origin-destination (O-D) matrix. The Growth Factor Model and the Gravity Model are two methods to distribute trips among destinations. The two methods for developing the O-D Matrix are presented and criticized in this paper, showing the similarities and differences between them and highlighting the implications for rigorous determination of future transport demand. A case study is done to emphasize the differences between these models and their implications in carrying out transport studies.

## 1 Introduction

The knowledge of transport demand, both for passengers and freight, it is important because the correct dimensioning of the infrastructure capacity and of necessary investments depends on the level of this knowledge [1]. It is known that traffic delays, caused by congestion, represent the most irritating problems of nowadays modern world, characterized by an increased need for mobility [2-3].

The intensive use of the passenger car in daily trips has highlighted the negative external effects and has permanently led to the life quality degradation of both those involved and those not involved in the transport process [4].

The traffic congestion is the most irritating phenomenon of the large urban agglomerations in inhabitants life.

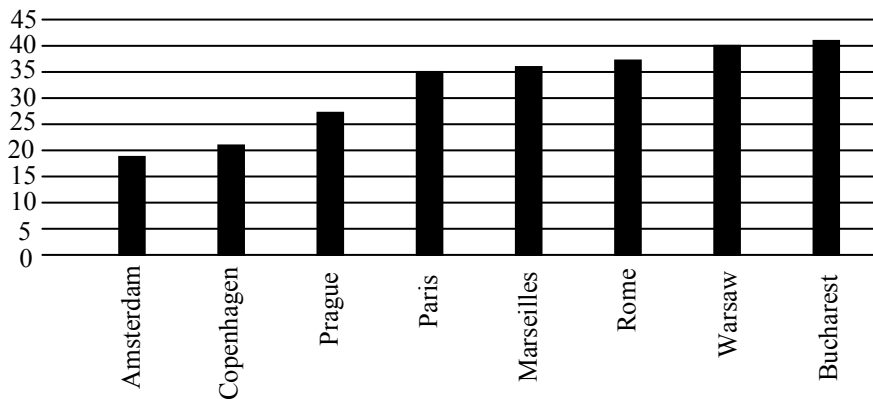
---

\* Corresponding author: [v\\_dragu@yahoo.com](mailto:v_dragu@yahoo.com)

In 10 cities in Romania the number of passenger cars exceeds 600 PC/1,000 inhabitants. Bucharest became a city with more vehicles than Amsterdam, Copenhagen, Prague or Rome [5]. The Mobility Plan proposed by Bucharest Metropolitan Transport Association (part of the civic society acting in accordance with the municipality) aims to create a complex public transport network to solve the actual congestion problems. As a benchmark, was settled to increase from 20% to at least 80% the use of the public transport. This goal has to be realized in the zone having the highest density of population in the country – over 8,000 inhab./km<sup>2</sup>.

According to the Sustainable Urban Mobility Plan 2016-2030 - Bucharest-Ilfov [6], Bucharest is the city with the largest congestion in Europe, and one of the most affected in the world. This study has determined the congestion index for several cities in the world (the congestion index is a percentage ratio of the difference between the real trip time and the time of the trip taking place outside of the peak hours relative to the time of the trip taking place outside of the peak hours) and it showed that Bucharest has a congestion index of 41%, reaching the 8th place (out of the 146 countries analyzed) in the worst ranking in the world and it is the last in Europe (fig. 1).

However, in 2012 Warsaw was at the top of the ranking, with a 42% congestion index, followed by Marseille (40%) and Palermo (39%), while Bucharest was far behind. While these cities have been able to reduce their congestion, the problem in Bucharest has become more serious.

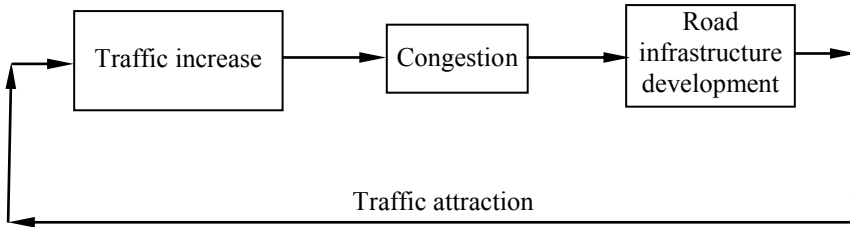


**Fig. 1.** The congestion index (source: *Sustainable Urban Mobility Plan 2016-2030 - Bucharest-Ilfov*)

Classical measures to meet the need for mobility have been focused exclusively on the development of the road infrastructure that led to the occupation of the land, degrading the natural landscape, increasing the risk of accidents and attracting new traffic, which has again led to congestion and the process is cyclical until the extension of the road infrastructure is no longer possible [7]. This describes a false pattern of congestion elimination (fig. 2.1.).

In order to ensure sustainable mobility development and, implicitly, a decrease in congestion, the traditional methods of reducing the capacity of the transport infrastructure should be associated with measures of reducing the social need for mobility. Among these, transport planning models are also important (the so-called “four-step travel demand model”). In the transport model, an important stage is trip distribution because it determines the number of trips made between the zones of the analysed city. In this stage, the traffic flows volumes cannot be determined because the modal split and the traffic assignment

need to be done first, but it is an incipient measure of the traffic volume and an initial information on possible traffic congestion.



**Fig. 2.** The false model to decrease congestion [8]

## 2 The origin-destination matrix

The origin - destination matrix is obtained in the second stage of the 4-steps transport model (trip generation, trip distribution, modal split, traffic assignment). This must be rigorously dimensioned in a transport study as it represent a transposition of human behavior into the transport process [9-10].

The matrix provides the total number of trips having the origin in  $i$  ( $G_i$ ) or the destination in  $j$  ( $A_j$ ), and the model used seeks to determine the  $G_i$  trip distribution by destinations and  $A_j$  trip distribution by origins on a certain network (table 1).

**Table 1.** The origin – destination matrix

Attraction Production	1	2 ...	j ...	n	$\sum_j n_{ij}$
1	$n_{11}$	$n_{12}$	$n_{1j}$	$n_{1n}$	$G_1$
2	$n_{21}$	$n_{22}$	$n_{2j}$	$n_{2n}$	$G_2$
⋮					
i	$n_{i1}$	$n_{i2}$	$n_{ij}$	$n_{in}$	$G_i$
⋮					
n	$n_{n1}$	$n_{n2}$	$n_{nj}$	$n_{nn}$	$G_n$
$\sum_i n_{ij}$	$A_1$	$A_2$	$A_j$	$A_n$	$\sum_{ij} n_{ij} = N$

The meanings of the notations are:

$n_{ij}$  represent the number of trips between zone  $i$  and zone  $j$ ;

$G_i$  – the number of trips produced by zone  $i$ ;

$A_j$  – the number of trips attracted by zone  $j$ ;

$n$  – the number of zones in which the city is divided;

The following requirement is fulfilled:

$$\sum_{i=1}^n G_i = \sum_{j=1}^n A_j, \tag{1}$$

which is usually called *the closing condition at the edge*.

Depending on the objective of the study and the availability of the data, the matrix deaggregation can be made depending on family type, the trip purpose and/or time of travel.

Basically, two methods are used in trip distribution stage [11-12]:

1. Growth factor methods, which are:
  - a) constant factor method;
  - b) average factor method;
  - c) Detroit factor method;
  - d) FRATAR method;
  - e) FURNESS method.
2. Synthetic methods using gravity type models or opportunity models.

Growth factor methods assume that in the future the tripmaking pattern will remain the same as today but that volume of trips will increase according to the growth of the generating and attracting zones. In this method the current travel allocation matrix it is assumed to be known. Applying a growing factor to the matrix results a new matrix. The growing factor can be constant (the same value for all the matrix) or it may vary and it can be applied for the next models:

- simply restricted, only for origin,  $G$  or destination  $A$ ;
- double restricted – both for origin and destination.

This method assumes that all zones will be developed in the future in a uniform way and the existing traffic model will be preserved, but growth factors will be taken into account.

Initially, in transport studies, the constant factor method was used. This method assumes that all zones will increase in the same manner and the existing traffic pattern will be the same for the future when the growth is taken into account. The relation between the present and the future trips can be described:

$$n'_{ij} = n_{ij} \cdot E, \quad (2)$$

where  $n'_{ij}$  - represents the number of the future trips between zone  $i$  and zone  $j$ ;

$n_{ij}$  - the number of the present trips between zone  $i$  and zone  $j$ ;

$E$  - the growth factor calculated as the ratio between the number of future trips and the existing number of trips in the analysed area.

The average factor method uses growth factors for producing and attracting zones, the trip distribution for each zone being obtained as in the constant factor method.

Expressed mathematically, this can be stated to be:

$$n'_{ij} = n_{ij} \frac{(E_i + E_j)}{2}, \quad (3)$$

with the producing and the attracting factors being calculated as such:

$$E_i = \frac{G_i}{g_i} \text{ și } E_j = \frac{A_j}{a_j}, \quad (4)$$

where  $G_i$  represents the number of trips that will be generated by zone  $i$ ;

$g_i$  – number of trips currently produced by zone  $i$ ;

$A_j$  – number of trips that will be attracted by zone  $j$ ;

$a_j$  – number of trips currently attracted by zone  $j$ .

The attractiveness and the production processes are not from the outset in agreement with future forecasts and the procedure needs to be iterated using new values for  $E_i$  and  $E_j$

factors ( $E'_i = \frac{G_i}{g_i}$  and  $E'_j = \frac{A_j}{a_j}$ , where  $g'_i$  and  $a'_j$  are the total productions and attractions of zones  $i$  and  $j$ , obtained from the first distribution of trips). The process is iterated using successive values of  $E'_i$  and  $E'_j$  until the growth factor approaches unity,  $E_i = E_j \cong 1$  and the successive values of  $n'_{ij}$  and  $n_{ij}$  are within 1 to 5 per cent, depending upon the accuracy required in the trip distribution.

For the origin destination matrix estimation using the Detroid method the mathematical relation is: [13]:

$$n'_{ij} = n_{ij} \frac{E_i E_j}{E}, \tag{5}$$

where  $E_i$  și  $E_j$  have the same meanings like before and  $E$  is the total productions factor, which is determined with the following relation:

$$E = \frac{\sum_{i=1}^n G_i}{\sum_{i=1}^n g_i}, \tag{6}$$

It is considered that there are no trips outside the city or trips with origin in other zones than those of the city. This can be written like this:

$$g_i = \sum_{j=1}^n n_{ij}; \quad a_j = \sum_{i=1}^n n_{ij}, \tag{7}$$

$$G_i = \sum_{j=1}^n n'_{ij}; \quad A_j = \sum_{i=1}^n n'_{ij}, \tag{8}$$

$$\sum_{i=1}^n g_i = \sum_{i=1}^n \sum_{j=1}^n n_{ij}, \tag{9}$$

$$\sum_{i=1}^n G_i = \sum_{i=1}^n \sum_{j=1}^n n'_{ij}. \tag{10}$$

The number of trips distributed between the areas of the city in the future situation is determined with the relation (5), iterative corrections are made until the production ( $E_i$ ), attraction ( $E_j$ ) and total factors ( $E$ ) are close enough to 1. As a rule, a tolerance of  $\pm 0.05$  relative to the unit is allowed, this could vary depending on the degree of precision required by the decider.

The Fratar Method [14] makes the assumption that the existing trips between zones  $i$  and  $j$  will increase proportional to the attraction factor of the destination zone  $j$  and to the production factor of the origin zone  $i$ :

$$n'_{ij} = K_i \cdot E_j \cdot n_{ij}, \tag{11}$$

with the restriction:

$$G_i = \sum_{j=1}^n n'_{ij} = K_i \sum_{j=1}^n E_j n_{ij} = E_i g_i, \quad (12)$$

from where the expression for  $K_i$  is obtained:

$$K_i = \frac{E_i g_i}{\sum_{j=1}^n E_j n_{ij}}. \quad (13)$$

The number of trips generated in the future situation is:

$$n'_{ij} = n_{ij} E_i E_j \frac{g_i}{\sum_{j=1}^n n_{ij} E_j}. \quad (14)$$

In the Furness method [15] the productions of the flows from a zone are first balanced and then the attractions to a zone are balanced. The algorithm is iterative and can be defined as follows:

$$n_{ij}^{(1)} = n_{ij} \cdot \frac{G_i}{g_i}, \quad (15)$$

$$n_{ij}^{(2)} = n_{ij}^{(1)} \cdot \frac{A_j}{\sum_{i=1}^n n_{ij}^{(1)}}, \quad (16)$$

$$n_{ij}^{(3)} = n_{ij}^{(2)} \cdot \frac{G_i}{\sum_{j=1}^n n_{ij}^{(2)}}, \quad (17)$$

where  $\sum_{i=1}^n n_{ij}^{(1)}$  is the sum of attracted trips in zone  $i$ , in the first iteration;

$\sum_{j=1}^n n_{ij}^{(2)}$  - sum of trips generated by the zone  $j$ , in the second iteration.

The use of the gravity models, also called synthetic models, leads to a trip distribution made according to socio-economic parameters of the analysed zones. For this reason, with synthetic models, it is possible to forecast future trips also in cases where there are no trips between zones, or even if future zones will appear in the study area.

Information about the existing situation are analysed and introduced in the synthetic models, in order to obtain a relation between production, attraction, trip distribution and the impedance function. The impedance function or the deterrence function shows the difficulties in carrying out trips and it is expressed in distance, trip time, travel cost, or the generalized cost. The principle of these models is to consider that the number of trips  $n_{ij}$  between two zones  $i$  and  $j$  of the study area, is proportional to the population of the origin zone  $i$  (or other generating element), noted with  $g_i$ , and to the number of jobs in the destination zone  $j$  (or other element that attract travel) noted with  $a_j$ , and also to the impedance function.

In general, the gravity model can be written as it follows:

$$n_{ij} = g_i \cdot a_j \cdot f(Z_{ij}), \tag{18}$$

where:  $g_i = \sum_{j=1}^n n_{ij}$  represent the number of trips produced by zone  $i$ ;

$a_j = \sum_{i=1}^n n_{ij}$  - number of trips attracted by zone  $j$ ;

$f(Z_{ij})$  - the impedance function of the trips.

For the impedance function the following formulations are used [16-17]:

$$a) f_1(Z_{ij}) = \alpha_1 d_{ij}^{-\beta_1} \tag{19}$$

$$b) f_2(Z_{ij}) = \alpha_2 e^{-\beta_2 d_{ij}} \tag{20}$$

where  $d_{ij}$  is the distance from zone  $i$  to zone  $j$  (in some models the duration or the generalized cost of trips from  $i$  to  $j$  may be used);

$\alpha_1, \alpha_2$  - the adjustment parameters;

$\beta_1, \beta_2$  - parameters that characterize the difficulties of traveling between zones  $i$  and  $j$ .

### 3 Case study

In a study area divided in 4 zones the following travel patterns were settled:

- number of trips produced and attracted by each zone (table 2);

**Table 2.** Number of trips produced and attracted

Trips \ Zone	1	2	3	4
Production ( $g_i$ )	200	400	100	200
Attraction ( $a_j$ )	300	200	200	200

- the calibration parameter obtained from traffic survey,  $\beta = 2$  ;
- the distances matrix ( $d_{ij}$ ) between the zones of the city are presented in table 3 (the distances are measured in kilometers).

This case study should determine the following elements:

1. the origin destination matrix using the simple restricted gravity model;
2. the best solution for trip distribution by successive iterations on the origin destination matrix.

**Table 3.** The distances matrix

$d_{ij}$	1	2	3	4
1	3	5	7	4
2	5	4	8	5
3	7	8	3	6
4	4	5	6	2

By using the simple restricted gravity model, the elements of the O-D matrix are assessed using the relation [16]:

$$n_{ij} = \frac{g_i \cdot a_j \cdot d_{ij}^{-\beta}}{\sum_{j=1}^n a_j \cdot d_{ij}^{-\beta}}, \tag{21}$$

Table 1 shows the trip distribution and the total number of attracted ( $a_j^{(1)}$ ) and produced trips ( $g_i$ ) from each zones.

**Table 4.** The origin – destination matrix

Origin \ Destination	1	2	3	4	$g_i$
1	115	28	14	43	200
2	135	140	35	90	400
3	17	8	60	15	100
4	46	19	14	121	200
$a_j^{(1)}$	313	195	123	269	-
$a_j$	300	200	200	200	-

In terms of trips number, it is agreed to approximate the results of the calculations to integer values.

As can be seen from the distribution matrix, the number of attracted trips is different from the attractiveness of the areas (zone 1 attracts 313 trips, and its attraction power is only 300, zone 2 attracts 195 and the attraction power is 200 and so on).

To ensure the convergence of the calculation, in the next iteration the  $a_j$  values will be corrected according to the relation:

$$a_j^{(m+1)} = a_j \frac{a_j}{a_j^{(m)}}, \tag{22}$$

With the corrected values of the attracted trips, the first iteration of the assignment matrix is calculated, with the relation:

$$n_{ij}^{(1)} = \frac{g_i \cdot a_j^{(1,1)} \cdot d_{ij}^{-\beta}}{\sum_{j=1}^4 a_j^{(1,1)} \cdot d_{ij}^{-\beta}}. \tag{23}$$

The distribution matrix obtained after the first and the second iterations is shown in the table 5.

**Table 5.** The distribution matrix obtained after the first and the second iterations

Destination \ Origin	1	2	3	4	$g_i$	Destination \ Origin	1	2	3	4	$g_i$
1	114	29	24	33	200	1	115	29	21	35	200
2	130	145	57	68	400	2	132	144	53	71	400
3	12	6	74	8	100	3	13	7	71	9	100

4	50	22	25	103	200	4	48	22	24	106	200
$a_j^{(2)}$	306	202	180	212	-	$a_j^{(3)}$	308	202	169	221	-
$a_j$	300	200	200	200	-	$a_j$	300	200	200	200	-

As you can see, the values  $a_j^{(3)}$  are much higher than previous values  $a_j^{(2)}$ , which leads to the conclusion that attracted trips have changed their distribution and the best solution of the problem was obtained at the previous step.

To support the assertions, the following two iterations of the distribution matrix will be presented below (table 6).

**Table 6.** The origin destination matrix after the third and the fourth iterations

Destination Origin	1	2	3	4	$g_i$	Destination Origin	1	2	3	4	$g_i$
1	115	28	17	40	200	1	114	29	21	36	200
2	134	141	42	83	400	2	131	145	51	73	400
3	15	8	65	12	100	3	13	7	70	10	100
4	47	20	17	116	200	4	49	22	21	108	200
$a_j^{(4)}$	311	197	141	251	-	$a_j^{(5)}$	307	203	163	227	-
$a_j$	300	200	200	200	-	$a_j$	300	200	200	200	-

It is noticed that the  $a_j^{(k)}$  values varies around  $a_j$  values, but the differences between them do not decrease more than in the first iteration.

In conclusion, the origin destination matrix was obtained after only one iteration.

## 4 Conclusion

The growth factors method is a easy to apply method with a simple iterative process, convenient in short-term analysis when there are relatively stable economic structures.

The disadvantages of this method are:

- the origin destination matrix of the present situation is assumed to be known;
- it requires the same data as other more sophisticated models, but provides much less information;
- the relation between the activity system and the transport system are not highlighted;
- the random trip distribution irregularities are multiplied by the use of the growth factor;
- the method assumes that accessibility and cost of the trip remain unchanged between the forecasting and the base period;
- between the two analytical moments the changes in the structure of the population and in the behavior of transport system users are neglected.

The Fratar method eliminates some disadvantages of the constant, average, and Detroit factors method. The method considers that in the future ( $n_{ij}$ ) trips will increase proportionally with  $E_i$  and  $E_j$ . This approach leads to a distorted result as the number of trips produced in zone  $i$  will be higher than forecasts. In order to overcome this shortcoming, a factor is added, which is given by the ratio between the sum of all trips

produced by zone  $i$  and the sum of the growth factor products for the attraction and the number of trips made between zone  $i$  and zone  $j$ .

The emergence of synthetic / synthesis models has added value to patterns of origin-destination matrix determination. Using these models, it is possible to make forecasts of the trips also in cases where there are no trips exchanges between zones or even of trips between zones that currently do not exist. However, a number of disadvantages of these models can be identified, as:

- More and more complex calculations, as the size of the analysis area increases;
- Difficult determination of the parameters that occur in the model and it is not obvious that the parameters remain constant over time.

Synthetic models formalize the closest to reality behavior of users who tend to maximize their individual benefits by making shorter and therefore economically and financially comfortable trips.

Estimating an origin-destination matrix that matches as accurately as possible to the traveling behavior it will ultimately lead to the development of a transport model that responds to the highest degree of demands imposed by sustainable development.

## References

1. V.A. Profillidis, G.N. Botzoris *Modeling of Transport Demand – Analyzing, Calculating and Forecasting Transport*, Elsevier Eds. (2018)
2. E. Holden, K. Linnerud, D. Banister *Transp. Res. Part A* **54**, pp. 67–77 (2013)
3. S. Raicu *Transport Systems*, AGIR Eds., Bucharest (2007)
4. D. Banister *Transp. Policy*, **Vol. 15**, pp. 73-80 (2008)
5. \*\*\* *Univers Ingineresc* nr. 17/1-15 septembrie (2015)
6. \*\*\* *Sustainable Urban Mobility Plan 2016-2030 - Bucharest-IIfov* (2015)
7. V. Dragu, E. Rosca, A. Rusca *Proceedings of the European Automotive Congress EAEC-ESFA 2015*, Bucharest, Romania, **Vol. II**, Springer International Publishing Switzerland, pp. 419-429 (2015)
8. V. Dragu, V.C. Roman, E.A. Roman *Bull. AGIR*, **No. 2**, pp.78-83, AGIR Eds. (2014)
9. A. Rasouli, A. Chegenizadeh, H. Nikraz *Electr. Journ. of Geotech. Eng.* **Vol. 22**, Bund 14, pp.5453-5464 (2017)
10. \*\*\* *Australian Transport Assessment and Planning Guidelines(ATAP)- T1 Travel Demand Modeling*, Transport and Infrastructure Council (2016)
11. V. Dragu *Urban and suburban passenger traffic*, BREN Eds., Bucharest (2001)
12. M. Popa *Transport Economics*, Politehnica Press Eds. , Bucharest, (2009)
13. N. Oppenheim *Urban Travel Demand Modeling*, John Wiley & Sons, New York (1994)
14. T.J. Fratar *Highway Research Board Proceedings*, pp. 376-385 (1954)
15. K.P. Furness *Traffic Eng. Ctrl.* , **Vol 7**, pp. 448-460 (1965)
16. R.J. Salter *Highway traffic analysis and design*, Macmillan Education LTD, London (1989)
17. M. Drlićiak, J. Čelko *Transp. Res. Procedia* **Vol. 14**, pp. 1733-1742 (2016)