A FUZZY LOGIC INFERENCE APPROACH FOR THE ESTIMATION OF THE PASSENGERS FLOW DEMAND

Aránzazu Berbey Alvarez, Rony Caballero George
Universidad Tecnológica de Panamá, Ciudad de Panama, Panama

Juan de Dios Sanz Bobi, Ramón Galán López
Universidad Politécnica de Madrid, Madrid, Spain

Keywords: Fuzzy logic inference, Passenger flow modelling, Maximum entropy theory, Matrix method of area trips O-D.

Abstract: This paper presents a new approach that designs the flow of passengers in mass transportation systems in presence of uncertainties. One of the techniques used for the prediction of passenger demand is the origin-destination matrices. However, this method is limited to urban areas and rarely to explicit stations. Otherwise, the gravity models based on friction functions can be another alternative; however, it is difficult to fit into practical achievements. Another solution might be the application of artificial intelligence techniques so as to include some intuitive knowledge provided by an expert to predict the flow demand of passengers' trips in explicit stations. This paper proposes to combine a matrix of origin-destination trips of travel zones, with the intuitive knowledge, applying a fuzzy logic inference approach.

1 INTRODUCTION

The passenger flow modelling has been tackled by several authors (Kiluchi et al, 1999) (Aldian et al, 2003) (Watson and Prevedouros, 2006) (Cheng et al, 2009) (Murat, 2010) (Xie et al, 2010). Some of them have applied linear programming methods considering small samples (Murat, 2010). It has also been applied the maximum entropy theory (Xie et al, 2010) and results have been compared between different methods (Watson and Prevedouros, 2006). However, it has been shown that these methods have limitations in certain scenarios. In this case, the use of fuzzy logic has proven to be a promising tool because it can integrate the railway planning experts’ experience in multiple scenarios (Aldian et al, 2003) (Cheng et al, 2009). This paper proposes to integrate the effectiveness of the methods based on origin-destination matrices, with the experience of experts in railway planning using an artificial intelligence support based on fuzzy logic.

2 AREA ANALYSIS

The matrix method of area trips O-D is one of the most used methods to design the movement of passengers. This method divides the urban environment in areas of interests that generate and attract trips. However, due to economic and practical reasons, the size of the interests’ zones is usually large in order to apply directly to urban planning models. Therefore, it is interesting in many cases to divide these macro areas into smaller areas, allowing a better analysis of passenger flow.

3 SUBZONE DESIGN

Considering the attraction vectors $D_T \in \mathbb{R}^{1 \times n}$ and the generation vectors $O_T \in \mathbb{R}^{n \times 1}$, each element of the attraction vectors $D_{ij}$ and generation $O_{ij}$ can be subdivided into $n_j$ subzones, becoming two new vectors $D \in \mathbb{R}^{1 \times n}$ and $O \in \mathbb{R}^{n \times 1}$, where $n = \sum_{j=1}^{n} n_j$. 

References:

Kiluchi et al, 1999
Aldian et al, 2003
Watson and Prevedouros, 2006
Cheng et al, 2009
Murat, 2010
Xie et al, 2010
Therefore, the macro areas matrix  $T \in \mathbb{R}^{n \times n}$ is subdivided at the same time, obtaining sub-zone matrix  $M \in \mathbb{R}^{n \times n}$. Consequently, it can be shown that every element of the attraction and the generation vectors, as well as the macro area and the subzone matrices must meet with:

\[
D_{ij} = \sum_{j=1}^{n} m_{i}(j) D_{j} \quad (1)
\]

\[
O_{II} = \sum_{i=1}^{n} n_{i}(i) O_{i} \quad (2)
\]

\[
T_{ij} = \sum_{i=1}^{n} n_{i}(i) n_{j}(j) M_{ij} \quad (3)
\]

\[
A = \sum_{j=1}^{n} M_{ij} \quad \forall j = 1 \ldots n \quad (4)
\]

\[
B = \sum_{j=1}^{n} M_{ij} \quad \forall i = 1 \ldots n \quad (5)
\]

where

\[
m_{B}(x) = \sum_{k=1}^{x} n_{k-1} \quad (6)
\]

\[
m_{E}(x) = \sum_{k=1}^{x} n_{k} \quad (7)
\]

\[
n_{0} = 0 \quad (8)
\]

To obtain a subzone model, a system is governed by the principle of maximum entropy would be applied. However, this model can be improved by incorporating information by an expert, even if this is inaccurate; this supposition is done with the aim to improve the approximation, and it means to apply the principle of maximum entropy. The information can be included using techniques based on artificial intelligence and more specifically based on fuzzy logic.

We have to apply the following steps to set up the subzone model:

a)  $T_{II}$,  $O_{II}$ and  $D_{II}$ have little uncertainty within the same planning horizon.

b) Every element of the vector  $O_{i}$ and  $D_{j}$ corresponding to a subzone  $i$ or  $j$, that is part of the macro area  $I$ or  $J$ respectively, can be represented as a function of  $O_{ii}$ or  $D_{jj}$ vectors of the macro area in which it is contained, and by a potential function with exponents  $C_{DI}$ or  $C_{DI}$.

c) Every element of the trip matrix  $M_{ij}$ corresponding to the subzones  $i, j$ and that is part of the macro area  $I, J$ respectively, can be represented as a function of  $T_{IL}$ and by a potential function with exponent  $C_{MI}$.

d) $C_{DI}$,  $C_{OI}$ and  $C_{MI}$ come from a fuzzy inference engine based on the experience of an expert.

It is important to note that the exponents  $C_{DI}$ and  $C_{OI}$ are related to the relative level of importance of the station in a given planning horizon. If these exponents are zero, the estimation becomes the estimation by maximum entropy, whereas if you have a negative or positive number it corresponds to a station with low or high demand respectively.

On the other hand the exponent  $C_{MI}$ establishes the relative level of importance of the passenger flow between two stations. Here, the estimation using the maximum entropy also corresponds for a value equal to zero, and a negative or positive number corresponds to a flow of low or high demand respectively.

\[
D_{j} = \frac{1}{n_{j}} \sum_{k=1}^{n_{j}} D_{ij} \quad (9)
\]

\[
O_{i} = \frac{1}{n_{i}} \sum_{k=1}^{n_{i}} O_{ii} \quad (10)
\]

\[
M_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} T_{ij} \quad (11)
\]

4 FUZZY INFERENCE ENGINE

The exponents  $C_{DI}$,  $C_{OI}$ and  $C_{MI}$ can be designed using the help of fuzzy logic. In this particular case it seems reasonable to use triangular membership functions (Figure 1). These exponents can be estimated not only considering the proximity of the station by major population centres, such as: residential areas, industrial estates, hospitals, schools and shopping centres, but also with the presence of transport interchanges.
5 PROPOSED ALGORITHM

a) Estimate the exponents $C_{Dj}$, $C_{Oi}$ and $C_{Mij}$ basing on the inaccurate available information.

b) Estimate the trip attraction and generation vectors $D$ and $O$, for the stations using,

$$D_j = \frac{2^{\frac{1}{n_j^2}} n_j^2}{2^{\frac{1}{n_j^2}} n_j^2} D_{ij} \quad (12)$$

$$O_i = \frac{2^{\frac{1}{n_i^2}} n_i^2}{2^{\frac{1}{n_i^2}} n_i^2} O_{ji} \quad (13)$$

c) Estimate the trip matrix $M_{ij}$, using,

$$M_{ij} = \frac{2^{\frac{1}{n_i^2}} n_i^2}{2^{\frac{1}{n_j^2}} n_j^2} T_{ij} \quad (14)$$

d) Thus it is checked if $O_i = \sum_{j=1}^{n_i} M_{ij} \forall i = 1 \cdots n_i$. Consequently, if $\left| \sum_{j=1}^{n_i} M_{ij} - O_i \right| < \varepsilon$ is met for a given value, it is accepted. Meanwhile, if not, the values of the matrix elements are corrected with

$$M_{ij} = \frac{O_i M_{ij}}{\sum_{j=1}^{n_i} M_{ij}} \quad (15)$$

e) It is checked if $D_j = \sum_{i=1}^{n_i} M_{ij} \forall j = 1 \cdots n_j$. Consequently, if $\left| \sum_{i=1}^{n_i} M_{ij} - D_j \right| < \varepsilon$ is met, it is accepted. Meanwhile, if not, the values of the matrix elements are corrected with

$$M_{ij} = \frac{D_j M_{ij}}{\sum_{i=1}^{n_i} M_{ij}} \quad (16)$$

f) It is checked if $T_{ij} = \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} M_{ij}$ for all $i, j, I, J$. Consequently if, $\left| \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} M_{ij} - T_{ij} \right| < \varepsilon$ is met, it is accepted. Meanwhile, if not, the values of the matrix elements are corrected with

$$M_{ij} = \frac{T_{ij} M_{ij}}{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} M_{ij}} \quad (17)$$

g) Go back to step d, if tolerances are not met.

6 RESULTS FOR SIMULATED DATA

The following case illustrates how the proposed algorithm works. Let’s consider the railway network topology shown in Figure 2. In this case it is known from previous studies that the mobility preferences among three regions so-called A, B and C, but without giving any specific details concerning the connection preferences between specific stations. The region A has the stations E1, E2 and E3; the region B has the stations E4, E5, E6 and E7; and the region C has only two stations E8 and E9. O-D matrix between these regions is known and it is described below:

$$M_{ij} = \sum_{j=1}^{n_i} M_{ij} \forall i = 1 \cdots n_i$$
Table 1: O-D matrix between the regions.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>9397</td>
<td>5282</td>
<td>5213</td>
</tr>
<tr>
<td>B</td>
<td>25118</td>
<td>8272</td>
<td>5065</td>
</tr>
<tr>
<td>C</td>
<td>22570</td>
<td>8732</td>
<td>1134</td>
</tr>
<tr>
<td></td>
<td>57085</td>
<td>22286</td>
<td>11412</td>
</tr>
</tbody>
</table>

Table 2: Original test matrix.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>2945.4</th>
<th>1142.1</th>
<th>1667.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2385</td>
<td>1555</td>
<td>920</td>
<td>543</td>
</tr>
<tr>
<td>B</td>
<td>3407</td>
<td>460</td>
<td>485</td>
<td>3247</td>
</tr>
<tr>
<td>C</td>
<td>2194</td>
<td>302</td>
<td>267</td>
<td>901</td>
</tr>
</tbody>
</table>

We can see that the results for the dominant values of the matrix are approximated; obviously the errors are due to the uncertain information coming from the expert, but still much lower than using the approximation for maximum entropy.

Table 4: Resulting O-D matrix using the maximum entropy method.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>2945.4</th>
<th>1142.1</th>
<th>1667.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1881.2</td>
<td>12284</td>
<td>4262</td>
<td>3534</td>
</tr>
<tr>
<td>B</td>
<td>2187.5</td>
<td>2274.9</td>
<td>4772</td>
<td>2503.9</td>
</tr>
<tr>
<td>C</td>
<td>2153.9</td>
<td>219.5</td>
<td>407.2</td>
<td>407.2</td>
</tr>
</tbody>
</table>

After applying the proposed algorithm, we obtain the following matrix OD.

Table 2: Resulting O-D matrix using the proposed algorithm.

Table 3: Resulting O-D matrix using the proposed algorithm.

Table 5: Percentage errors using the proposed algorithm.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>2945.4</th>
<th>1142.1</th>
<th>1667.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1881.2</td>
<td>12284</td>
<td>4262</td>
<td>3534</td>
</tr>
<tr>
<td>B</td>
<td>2187.5</td>
<td>2274.9</td>
<td>4772</td>
<td>2503.9</td>
</tr>
<tr>
<td>C</td>
<td>2153.9</td>
<td>219.5</td>
<td>407.2</td>
<td>407.2</td>
</tr>
</tbody>
</table>

Table 6: Percentage errors using the maximum entropy method.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>2945.4</th>
<th>1142.1</th>
<th>1667.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1881.2</td>
<td>12284</td>
<td>4262</td>
<td>3534</td>
</tr>
<tr>
<td>B</td>
<td>2187.5</td>
<td>2274.9</td>
<td>4772</td>
<td>2503.9</td>
</tr>
<tr>
<td>C</td>
<td>2153.9</td>
<td>219.5</td>
<td>407.2</td>
<td>407.2</td>
</tr>
</tbody>
</table>

We can see that the results for the dominant values of the matrix are approximated; obviously the errors are due to the uncertain information coming from the expert, but still much lower than using the approximation for maximum entropy.

Table 4: Resulting O-D matrix using the maximum entropy method.
7 CONCLUSIONS

In this paper the potential and effectiveness of this new methodology has been proven. Clearly, the obtained matrix has an error of 10% compared to the original matrix. However, these results are reasonable considering the level of uncertainty coming from the information provided by the expert. As a guideline for future research, it is interesting to improve the fuzzy inference engine, since the used in this test was quite simple. Here, other sources of information can be incorporated without limiting the information just to one expert.

ACKNOWLEDGEMENTS

The authors of this paper want to express their gratitude to the National Secretary of Science and Technology (SENACYT) of the government of the Republic of Panama for funding this study through the R & D project “Methodologies and Performance Indicators of Railway Transport Systems” adjudged by a call for the promotion of R&D activities (MDEPRB09-001). Additionally, they want to thank the Research Centre on Railway Technologies (CITEF) of Universidad Politécnica de Madrid (Spain).

REFERENCES


