EFFICIENT METHODS FOR THE TRAFFIC ASSIGNMENT PROBLEM

M. EL BAKRY *

ABSTRACT

The capacity restrained traffic assignment problem is defined as "assign the vehicular traffic to different links of the network, given an origin-destination matrix for the traffic required between the nodes of the network, and a nonlinear relation for the travel time on each link as function of the traffic flow in that link". The traffic is assumed to follow user's optimization according to Wardrop's principle, which states that the travel costs on all used paths for each of the origin-destination pairs are equal and less than the travel cost on the unused paths.

This paper introduces four methods representing different approaches for solving this problem, the successive incremental assignment, Leblanc's model, a modified form of Leblanc's model and the method of negative loops. These methods are explained and compared with the help of an illustrative example.

1. PROBLEM FORMULATION

The traffic area under study is represented by a directed graph (network), in which each directed arc represents a road direction (link) and each node represents an intersection. Nodes are used as places at which the traffic enters and leaves the network with prespecified amount of flow.

* Dr. Eng., Electronic Research Institute, National Research Centre, Dokki, Cairo, Egypt.
given in the form of an origin-destination matrix.

It is required to find the traffic flow in each link in the network, assuming that each driver takes the path of least cost from his origin to his destination. According to Wardrop's first principle,[1], an equilibrium state will be reached, which is referred to as "the principle of equal travel costs", and states that: For any two nodes r and s

i. if two or more paths between node r and node s are actually travelled, then the cost of travel between r and s must be the same for each of these paths, and

ii. there does not exist an alternative unused path between nodes r and s with less cost than that pathes which are travelled.

It is known that the travel time is a significant factor to the majority of travellers when they choose their routes. We shall use, as it is usual in most traffic models, the travel time as a basic measure of travel cost.

We shall use the notation \( f_{ij}(x_{ij}) \) to represent the average travel time on the link from node i to node j if the flow on that link is \( x_{ij} \).

If all the links of the network are slightly loaded, the average travel time on every link is constant independent of the flow: \( f_{ij}(x_{ij}) = f_{ij}^{(o)} = \frac{L}{V_{ij}} \)

where \( f_{ij}^{(o)} \): The travel time on the link \( ij \) at zero flow,

\( k_{ij} \): The capacity of the link \( ij \), at which the average travel time is equal to \( 2 f_{ij}^{(o)} \), and

\( n \): positive constant.

This form will be used in solving an illustrative example in the next sections. For the simplicity of analysis and without loss of the nonlinearity nature of \( f_{ij}(x_{ij}) \), we shall take \( n=2 \) in the above form.
2. SUCCESSIVE INCREMENTAL ASSIGNMENT [3]

Procedure: 1. Start with the network at zero flow. The travel time on every link $ij$ is $f_{ij}(x_{ij}) = f_{ij}^{(0)}$. Determine a shortest path between every origin-destination pair according to these travel times.
2. A certain predetermined percentage (increment) of the traffic demand (O/D matrix) is added to a shortest path for every origin-destination pair obtained in the last step.
3. Find $f_{ij}(x_{ij})$ for each link, then determine accordingly a shortest path between each origin-destination pair.
4. If all the traffic demand is assigned stop, otherwise go to step 2.

Example: Fig.1 shows a simple network

Travel time functions:

- $f_{12} = 1 \left( 1 + \left( \frac{x_{12}}{100} \right)^2 \right)$
- $f_{13} = 2 \left( 1 + \left( \frac{x_{13}}{200} \right)^2 \right)$
- $f_{14} = 5 \left( 1 + \left( \frac{x_{14}}{200} \right)^2 \right)$
- $f_{24} = 2 \left( 1 + \left( \frac{x_{24}}{100} \right)^2 \right)$
- $f_{34} = 2 \left( 1 + \left( \frac{x_{34}}{200} \right)^2 \right)$

O/D Matrix:

- From $Q_1$ to $S_1$: 50, to $S_2$: 100
- From $Q_2$ to $S_1$: 0, to $S_2$: 200

We suppose that 1% of the traffic will be added in each incrementation step of the procedure.

The traffic flow from $Q_1$ to $S_1$ is fixed by the path $V_1-V_3$ and from $Q_2$ to $S_2$ is fixed by the path $V_2-V_4$ while the traffic from $Q_1$ to $S_2$ has 3 alternative paths.

At zero flow the traffic from $Q_1$ to $S_2$ will take the path $V_1-V_2-V_4$ with shortest time = 3 min. The assignment of the traffic from $Q_1$ to $S_2$ to the path $V_1-V_2-V_4$ will continue in each incrementation and thus the travel time on the links $V_1-V_2$ and $V_2-V_4$ will continue to increase until a turning point is reached at which $f_{1,2}(x_{1,2}) + f_{2,4}(x_{2,4}) = f_{1,3}(x_{1,3}) + 2$. Then the traffic from $Q_1$ to $S_2$ will change to the path $V_1-V_3-V_4$. Let this turning point be reached at a percentage "100a" of the total demand, $0 < a < 1$, then we have:

$1(1+(100a/100)^2) + 2(1+(300a/100)^2) = 2 + 2(1+(50a/200)^2)$

which gives $a = 0.2301$.
Starting from $a = 0.24$, $f_{1,3} + f_{3,4}$ will be less than $f_{1,2} + f_{2,4}$ since the additional traffic on the link $V_1-V_3$ due to traffic from $Q_1$ to $S_1$ is less than the additional traffic on the link $V_2-V_4$ due to the traffic from $Q_2$ to $S_2$.

The assignment of the traffic from $Q_1$ to $S_2$ to the path $V_1-V_3-V_4$ will continue until $f_{1,2}(x_{1,2}) + f_{3,4}(x_{3,4}) = 5$ then it will change to the path $V_1-V_4$. The value of the percentage "100a" at which the second turn occurs is given by

$$2 \left[ 1 + (50a + 100a - 23/200)^2 \right] + 2 \left[ 1 + (100a - 23/200)^2 \right] = 5$$

which gives $a = 0.976$

Thus starting from $a = 0.98$ till $1$ the traffic from $Q_1$ to $S_2$ will take the path $V_1-V_4$.

We note that in the final assignment, the 3 alternative paths for the traffic from $Q_1$ to $S_2$ are all used, while the travel time on these paths are completely different: for the path $V_1-V_2-V_4$, $f_{1,2} + f_{2,4} = 13$ min,

for the path $V_1-V_3-V_4$, $f_{1,3} + f_{3,4} = 5.04$ min,

and for the path $V_1-V_4$, $f_{1,4} = 5$ min

We conclude that this method may not realize the principle of equal travel costs, whatever the number of increments used.

3. LEBLAN'S MODEL [4]

Applying the principle of equal travel costs to a traffic network is according to Beckmann[5], equivalent to solve the nonlinear convex optimization problem: minimize $Z = \sum_{\text{all links } ij} x_{ij} f_{ij}(t_{ij}) \ dt_{ij}$.

subject to the continuity of flow constraints at every node of the network. Leblan's model is based on this concept and may be summarized in the following procedure:

1. Start with the network at zero flow, determine a shortest path between every origin-destination pair according to the travel times $f_{ij}^{(0)}(x_{ij})$.

2. Assign the whole traffic demand (O/D matrix) to the shortest paths found in step 1. This gives the zero assignment vector $x^{(0)}$.

Set iteration counter $n = 1$.

3. Determine the new shortest paths according to the last assignment $x^{(n-1)}$.

4. Corresponding to the new shortest paths find the traffic assignment $y^{(n)}$.
5. Apply a one-dimensional search technique to find the value of $x^{(n)}$:

$$x^{(n)} = (1-a) x^{(n-1)} + a y^{(n)}, \quad 0 \leq a \leq 1$$

that minimize

$$Z = \sum_{\text{all links } ij} \int_0^x f_{ij}(t) \, dt.$$ 

6. Test the stopping criterion. If it fails set $n=n+1$ and go to step 3.

The stopping criterion may be:

- Fixed number of iterations, or
- An upper limit for the max. difference between two successively obtained values of travel times $f_{ij}$ or values of flow $x_{ij}$, for all links $ij$.

Solution of the example of fig 1:

The flow vectors $x^{(n)}$ and $y^{(n)}$ will be written as a raw vector in the form $(x_{1,2}', x_{1,3}', x_{1,4}', x_{2,4}', x_{3,4}')$ and the travel time vector $f$ in the form $(f_{1,2}', f_{1,3}', f_{1,4}', f_{2,4}', f_{3,4}')$.

At zero flow the shortest path from $Q_1$ to $S_2$ is $V_1-V_2-V_4$, hence the initial traffic assignment is given by

$$x^{(0)} = (100, 50, 0, 300, 0) \quad \text{and} \quad f^{(0)} = (2, 2.125, 5, 20, 2).$$

1st Iteration: According to the above values, the travel time over the path $V_1-V_2-V_4 = 22$ min. A new shortest path from $Q_1$ to $S_2$ is given by $V_1-V_3-V_4 = 4$ min. The corresponding total assignment will be

$$x^{(1)} = (0, 150, 0, 200, 100)$$

Apply the search technique

$$x^{(1)} = (1-a) x^{(0)} + a y^{(1)}$$

$$= (100-100a, 50 + 100a, 0, 300-100a, 100a)$$

Substituting in the expression for $Z$, we obtain

$$Z = \sum_{\text{all links } ij} f_{ij}^{(0)} (x_{ij} + x_{ij}^3/3k_{ij})$$

$$= 100 + 100a + 100/6 a^3 + 100/3 (1-a)^3 + 100/6 (.5+a)^3 + 200/3 (3-a)^3.$$ 

To minimize $Z$ we solve $dz/da = 0$, and obtain $a = 8.5 \pm 6.06/2$, i.e. no turning point lies bet $0,1$. The min. value of $Z$ lies on the boundary, actually at $a = 1$

$$x^{(1)} = y^{(1)} = (0, 150, 0, 200, 100) \quad \text{and} \quad f^{(1)} = (1, 3.125, 5, 10, 2.5)$$

2nd Iteration: The travel time over the path $V_1-V_2-V_4 = 1+10=11$ min,

over the path $V_1-V_3-V_4 = 3.125+2.5 = 5.625$ min, and over the path $V_1-V_4 = 5$ min. Thus the new assignment from $Q_1$ to $S_2$ takes the path
We have already after two iterations an accurate solution which satisfy the principle of equal travel costs.

We have used this model to solve large scale problems and it gives almost accurate results in small number of iterations. A Pascal-program for this model is available in Ref. [6].

4. A MODIFIED FORM OF LEBLANC'S MODEL [9]

Leblanc's model is simplified by excluding the one dimensional search technique and taking $a = \frac{1}{n}$ at the $n$th iteration, which leads to

$x(1) = y(1)$, $x(2) = \frac{1}{2} x(1) + \frac{1}{2} y(2)$

...and in general

$x(k) = \frac{1}{k} (y(1) + y(2) + \ldots + y(k))$

The proof of the convergence of this method is given by Powell [7], but actually the convergence is much slower than that of the Leblanc's model.

Solution of the example of fig 1

Using the same notations for the flow vector and travel time vector and following the same steps as in section 3, we obtain

$x(0) = (100, 50, 0, 300, 0)$, $f(0) = (2, 2.125, 5, 20, 2)$

1st Iteration : Shortest path from $Q_1$ to $S_2$, $V_1-V_3-V_4$

$x(1) = (0, 150, 0, 200, 100)$, $y(1) = x(1)$ and $f(1) = (1, 3.125, 5, 10, 25)$

2nd Iteration : Shortest path from $Q_1$ to $S_2$, $V_1-V_4$

$x(2) = (0, 50, 100, 200, 0)$, $x(2) = \frac{1}{2} (y(1) + y(2))$, we obtain

$x(2) = (0, 100, 50, 200, 50)$ and $f(2) = (1, 2.5, 5.3125, 10, 2.125)$
3rd Iteration: Shortest path from $O_1$ to $S_2$: $V_1-V_3-V_4$

$y^{(3)} = (0, 150, 0, 200, 100)$, $x^{(3)} = \frac{1}{3} (y^{(1)} + y^{(2)} + y^{(3)})$, thus

$x^{(3)} = (0, 116.7, 33.3, 200, 66.7)$ and $f^{(3)} = (1, 2.681, 5.139, 10, 2.222)$

4th Iteration: Same shortest paths as in iteration 3, i.e. $y^{(4)} = y^{(3)}$

$x^{(4)} = \frac{1}{4} (y^{(1)} + y^{(2)} + y^{(3)} + y^{(4)}) = (0, 125, 25, 200, 75)$ and $f^{(4)} = (1, 2.78, 5.078, 10, 2.28)$

The used paths from $O_1$ to $S_2$ are $V_1-V_4$ with time 5.078 min and $V_1-V_3-V_4$ with time $2.78 + 2.28 = 5.06$ min. This solution is very close to the exact one. However if we try to obtain more accurate solution a very large number of iterations have to be performed.

5. METHOD OF NEGATIVE LOOPS

Sahling,[8], proposed a method for the traffic assignment problem. It depends on the use of incremental networks, which is obtained from the traffic network by adding to the used links another links in the opposite direction with negative values of travel times used on that links. Fig.2b shows an incremental network for the network of fig.2a.

The method is based on the fact that if the principle of equal travel times is not reserved in the network, then loops with net negative sum of travel times (negative loops) exist. E.g. in fig.2b two negative loops could be found $V_1-V_4-V_2-V_1$ and $V_1-V_3-V_4-V_2-V_1$ with sums -17 and -17.875 respectively.

To eliminate a negative loop, we add a flow $w$ to each of the positive links of the loop and subtract the same flow $w$ from each of the negative links. To determine $w$ we have to solve the equation:

$$\sum_{uv \in \text{all + links}} f_{uv}(x_{uv} + w) = \sum_{uv \in \text{all - links}} f(x_{uv} - w)$$

The procedure of the method may be summarized as follows:-

1. Start with the network at zero flow. The travel time on every link $ij$ is $f_{ij}(x_{ij}) = f_{ij}^{(0)}$. Determine a shortest path between every origin-destination pair according to these travel times.

2. Assign the whole traffic demand to the network according to the shortest paths found in the last step.

3. Draw the incremental network and find out the negative loops. If these is no negative loops or there exist negative loops with very small...
absolute value of negative sum (beyond a given accuracy limit) stop, we have obtained the required solution.

4. Take a negative loop with least absolute value of -ve sum. Find w by solving the above equation, then obtain \( w = \min(w, \min(x)) \) for -ve links.

The value of \( w \) represents the max. flow that can be subtracted from the negative links without leading to a net negative flow. A new traffic assignment is obtained by correcting the flow on links of the negative loop with \( w \).

5. Calculate the new travel times over the network and the corresponding shortest paths. Go to step 2.

Solution of the example of fig 1

Fig 2.a shows the assignment according to the shortest paths at zero flow. From fig 2.b, we find that the incremental network contains two negative loops. We consider the loop with least absolute value of sum of flow, i.e., the loop \( V_1-V_4-V_2-V_1 \). The value of \( w \) is obtained by solving the equation:

\[
5(1+(w/200)^2) = (1+(100-w/100)^2) + 2(1+(300-w/100)^2),
\]

which gives \( w = 149 \). \( w = \min(w, \min(100, 300)) = 100 \). Adding a flow 100 to the link \( V_1-V_4 \) and subtracting the same flow from each of the links \( V_1-V_2 \) and \( V_2-V_4 \), we obtain the traffic assignment of fig 2.c. Its incremental network, fig 2.d, contains only one negative loop \( V_1-V_3-V_4-V_1 \). The flow \( w \) that corrects the loop assignment is given by

\[
5(1+(100-w/200)^2) = 2(1+(50+w/200)^2) + 2(1+(w/200)^2),
\]

which gives \( w = 75.6 \). \( w = \min(75.6, 100) = 75.6 \). The resulting total traffic assignment is shown in fig 2.e, which is the same solution obtained by Leblanc's model in two iterations.

The method of negative loops is an intelligent graphical method that applies directly the principle of equal travel times without the need of nonlinear programming formulation. An ALGOL-program for this method is given in [8].

6. CONCLUSION

We have considered four models for solving the capacity restrained traffic assignment problem that represent different approaches for the solution. These models are briefly explained with the help of an illustrative example. The successive incremental assignment method is a heuristic method that may lead to fault solutions. The model of Leblanc as well as the
Fig. 2 Solution of example by negative loops
method of negative loops of sahling leads to an exact solution in a small number of iterations. The modified form of Leblanc is much simpler but it converges very slowly to the exact solution.

7. REFERENCES