Optimization of Transit Network Assignment

Alireza Naseri

Department of Civil Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran

Abstract: Urban development and the installation of new and manifold activities on a territory, with industrial, commercial or service character, cause an always increasing demand on mobility to which it is necessary to offer an efficient well-organized and well-distributed on the territory transportation system. The adjustment of old networks to new needs must occur across unitary and organic vision of the problem in order to obtain maximum economy for the resources employed and maximum functionality for transport users. Solution of the transportation network optimization problem actually requires, in most cases, very intricate and powerful computer resources and simulation, so that it is not feasible to use classical algorithms. Particularly, this paper will focus the attention on the possibility of modeling and optimizing transit networks by simulation. The transit assignment offered in EMME/2, which is based on optimal strategies, does not consider congestion effects due to limited vehicle capacity. This assignment model can be extended by taking into account the vehicle capacity by means of a volume-dependent transit time function, leading to the formulation of a transit assignment method. Transportation Planning Software can be used to solve assignment model and research show good perspectives for this approach.

Key words: Transit Network, Assignment, Flow Optimization, Simulation.

INTRODUCTION

Launching and exploiting public transport systems usually entails exorbitant costs for system provider and government. So, system provider is trying to optimize their costs. On the other hand, the purpose of launching public transportation system is creating fast and affordable transportation for all segments of the society and at the same time reducing the negative effects of transportation such as air pollution, congestion, noise pollution and so on. Usually minimizing total travel time of the users is chosen as the aim of optimization of transportation network.

In a public transportation system, parameters like travel time, convenience and others are directly under the effect of the public transport network coverage, periodicity and its service versus other traffic services and features which are related to them. The quality of the public transportation network is determined on the basis of factors like directness of the routes, service coverage, operating expenses and cost of public transportation for the users and average number of vehicle changes. Also, route selection is one of the most important factors in desirability or non-desirability of public transportation (Zhao, 2004). In transit assignment, congestion effects due to overcrowding of the vehicles are not taken into account for modeling of the route choice. It is a suitable approach in all those cases where the goal of the planning process is to provide enough capacity for all transit passengers on the routes of their choice. There are, however, situations in which it is not feasible to provide enough transit capacity to preclude congestion. In these cases, the route choice of the transit passenger is likely to be influenced by the congestion on board the vehicles, so that some travelers will switch from congested to less congested routes, even if the latter are less attractive in terms of travel time or cost (Guihaire, 2008).

The congestion is modeled by means of volume dependent cost functions, similar to the volume-delay functions used in the highway assignment. After having presented the mathematical formulation of the model, we discuss its implementation in EMME/2.

Method:

As congestion in large urban areas continues to worsen and gas prices began to rise in the recent years, the attractiveness of public transit as an alternative to private cars has also been growing. However, for a public transit system to help meet the growing travel demand and alleviate the congestion problem, it must be able to provide reasonable travel time and convenience relative to private vehicles. Travel time and convenience are affected directly by the configuration of a transit network and service frequency, although other service and traffic characteristics and pedestrian environment will also have an impact on the willingness of the public to use transit. The quality of a TN may be evaluated in terms of a number of parameters including route directness, service coverage, operator cost, transit user cost (including waiting, in-vehicle, and transfer times), and the average number of transfers required to accomplish a trip. Route directness may be measured by the additional travel time incurred to a transit user when a bus does not follow the most direct route between the user’s origin to destination.
and destination. Service coverage refers to the percentage of total estimated demand (i.e., transit trips) that may be potentially satisfied by the transit services provided, based on a given transit route network. Operator cost is the cost to a transit property to provide transit services within a given network. Transfers are a result of not being able to provide direct services between all pairs of origins and destinations (Zhao, 2003).

The transit network is assumed to be represented by a standard node/link type network. Travel time (or cost) \(c_a\) and a service frequency \(f_a\) is associated with each network link \(a\). The demand between nodes \(i\) and \(j\) is given by \(g_{ij}\).

Note that in this type of network representation, the itineraries of the transit lines are implicitly contained in the network topology. The set of nodes not only contains the physical nodes of the underlying street or rail network, but also one additional node for each transit stop of each line. Correspondingly, the links are subdivided into various classes, such as boarding, alighting, in-vehicle and walking links. Note that only boarding links imply waiting, thus have a finite frequency \(f_a\). All other links are served continuously (\(f_a = \infty\)).

The waiting time at a node depends on the set of attractive links \(\bar{A}_i^+ \subseteq A_i^+\), i.e. the set of outgoing links which are considered for travel by the travelers by boarding the first vehicle leaving on any of these links. For any given set of \(\bar{A}_i^+\), the waiting time at node \(i\) for \(\bar{A}\) (Spiess, 1984; Spiess and Florian, 1989):

\[
W(\bar{A}_i^+) = \frac{1}{\sum_{a \in \bar{A}_i^+} f_a} 
\]

(1)

Probability of leaving node \(i\) on link \(a\) is:

\[
P_a(\bar{A}_i^+) = \frac{f_a}{\sum_{a' \in \bar{A}_i^+} f_{a'}} , \quad a \in \bar{A}_i^+ 
\]

(2)

Where:

- \(I\) : Nodes
- \(\bar{A}\) : Links
- \(A_i^+\) : Outgoing links at node \(i\)
- \(A_i^-\) : Incoming links at node \(i\)
- \(\bar{A}\) : Attractive links \(\bar{A}_i^+ \subseteq A_i^+\)
- \(\bar{A}_i^+\) : Attractive outgoing links at \(i\)
- \(\bar{A}_i^-\) : Attractive incoming links at \(i\)

Given the above relations, any strategy for reaching destination \(r\) is completely defined by the corresponding subset of attractive links \(\bar{A} \subseteq A\).

The strategy for reaching a destination is the one which minimizes the total expected cost. Note that the cost of a strategy is the sum of link travel times \(c_a\) weighted by the probability of traveling on link \(a\), and the waiting time at nodes \(i\) weighted by the probability of traveling through node \(i\). It has been shown that for fixed link travel times \(c_a\), the assigning of the trips from all origins to destination \(r\) according to the optimal strategy corresponds to solving the following linear optimization problem:

\[
\text{Min} \sum_{a \in A} c_a v_a + \sum_{i \in I} w_i 
\]

(3)

S.t:

\[
v_a \leq f_a w_i , \quad a \in \bar{A}_i^+ , \quad i \in I 
\]

(4)

\[
\sum_{a^+ \in A^+_i} v_{a^+} - \sum_{a^- \in A^-_i} v_{a^-} = g_{ij} , \quad i \in I 
\]

(5)

\[
v_a \geq 0 , \quad a \in A 
\]

(6)
The variables \( w_i \) represent the total waiting time (in person minutes) at node \( i \). The problem can be solved very efficiently by means of the following label-setting type algorithm:

**Fixed Cost Optimal Strategy Transit Assignment for Trips to Destination \( r \):**

**Step 0:**

\[ u_i \leftarrow \infty, i \in I - \{r\}; \ u_r \leftarrow 0; \ a \in A. \]

**Step 1:**

\[ S \leftarrow A; \ A \leftarrow \emptyset. \]

\[ f_i \leftarrow 0, \ V_i \leftarrow 0, \ i \in I; \]

\[ f_i \leftarrow f_i + f_a(u_j + c_a) \]

\[ f_i \leftarrow f_i + f_a, \ A \leftarrow A + a; \]

**Step 2:**

\[ S \leftarrow S - \{a\}. \]

**Step 3:**

**Assignment Model:**

The modeling of congestion aboard the vehicles may be done associating congestion functions with the segments of transit lines to reflect the crowding effects. We now turn our attention to the variant of the transit assignment problem in which the link travel times \( \tau \) are no longer constants, but are continuous non-decreasing functions of the corresponding link flows \( \tau(v_a) \). Such a dependence of the link cost on the transit volume may represent an actual slowing down of the transit vehicle due to the number of passengers, but it may also be interpreted as a generalized cost which includes a “discomfort” term which increases as the vehicles get crowded. In this context, the transit assignment problem is no longer separable by destination node, since the link costs depend on the total flow of passengers. The total transit volumes are the sum of the volumes bound for each of the destinations (Wardrop, 1952).

The final model is nonlinear and leads to a transit equilibrium model by resorting to Wardrop’s user optimal principle may be stated as: For all origin destination pairs the strategies that carry flow are of minimal generalized cost and the strategies that don’t carry flow are of a cost which is larger or equal to the minimal cost. This leads to a convex cost optimization problem. The resulting model is:

\[
\begin{align*}
\text{Min} & \sum_{a \in A} \int_0^{v_a} c_a(x)dx + \sum_{i \in I} \sum_{r \in R} w_i' \\
\text{Subject to:} & \\
& v_a = \sum_{r \in R} v_a' , \quad a \in A \\
& v_a' \leq f_a w_i' , \quad a \in A^r , \quad i \in I , \quad r \in R \\
& \sum_{A'} v_a' - \sum_{A^r} v_a' = g_{ij} , \quad i \in I , r \in R \\
& v_a' \geq 0 , \quad a \in A , \quad r \in R
\end{align*}
\]

The model is solved by using an adaptation of the linear approximation method. Each sub problem requires the computation of optimal strategies for linear cost problems. An important advantage of this method is the fact that only total volumes need to be computed and stored, since the destination dependent volumes \( v_a' \) are dealt with implicitly (Spiess, 1984; Frank and Wolfe, 1956).

Find any feasible solution \((v^*, w^*)\) where \( v^* \) denotes the vector of total flows \( v_a' \), and scalar \( w^* \) denotes the corresponding total waiting \( \sum_{i \in I, r \in R} w_i' , k \leftarrow 0. \)
Step0:

\[ k \leftarrow k + 1 \]

Compute \((\hat{v}, \hat{w})\) by solving the fixed cost transit assignment with costs \(c_a = c_a(v_a^{k-1})\) for each destination \(r \in R\).

Step1:

Find \(\lambda^*\) that minimizes the objective function on the line segment

\[ (1 - \lambda) (v^{k-1}, w^{k-1}) + \lambda (\hat{v}, \hat{w}), 0 \leq \lambda \leq 1. \]

Step2:

Set \((v^k, w^k) \leftarrow (1 - \lambda^*) (v^{k-1}, w^{k-1}) + \lambda^* (\hat{v}, \hat{w})\).

If \(\sum_{a \in A} c_a(v_a^{k-1}) (v_a^{k-1} - \hat{v}^{k-1}) + w^k - \hat{w} < \varepsilon\) then STOP.

Step3:

Otherwise go to STEP 1.

The minimization in Step 2 is best implemented not by actual minimization, but by annulling the derivative, i.e. solving the equation

\[
\sum_{a \in A} c_a(v_a^{k-1} + \lambda (\hat{v} - v_a^{k-1}))(v_a^{k-1} - \hat{v}^{k-1}) + \lambda (\hat{w} - w^{k-1}) = 0
\]

Note that the stopping criterion used in Step 3 of the above algorithm corresponds to the absolute gap, which is an upper bound for the difference between the objective function at the current solution and at the true optimum.

Preparation of Simulation:

There is a need to model the limited capacity of transit lines and the increased waiting times as the flows reach the capacity of the vehicle. As a transit segment becomes congested, the comfort level decreases and the waiting times increase. The effective frequency of a line is line with infinite capacity and Poisson arrivals which yields the waiting time obtained by the adjusted headway. Now we describe how the transit assignment can be simulated in Transportation Planning Software (Spiess, 1984; Inro, 1992). An important feature of EMME/2 is its modularity. The various functionalities used in the transportation planning practice are implemented as a set of independent basic tools, all acting on a common data bank. These can easily be used individually or in combination to form more complex models. A powerful macro language is provided within the EMME/2 system, which allows the user to implement the various steps of the model and to automate the procedure.

To implement the equilibrium transit assignment discussed in the previous section, the following basic EMME/2 tools are used:

1. Fixed cost transit assignment: This is the standard EMME/2 transit assignment model. It implements the optimal strategy assignment described earlier. Of course, in EMME/2 the transit network is expressed by explicit transit line itineraries, so that the user need not be concerned with the “exploded” network representation used here for the mathematical formulation of the model. The travel cost on each transit line segment is given by applying the corresponding user definable travel time function. The assignment can be fine tuned by various parameters and weights, which are not discussed here.

2. Network Calculator: This is a very general tool to evaluate algebraic expressions combining any kind of network information. Automatic conversion between the different element levels (node, link, transit line, transit segment) is provided.

3. Matrix Calculator: Similar to the Network Calculator, this tool allows the evaluation of expressions containing any kind of matrix information. Automatic conversion between the different matrix formats (full matrix, origin vector, destination vector, scalar value) is provided.

The travel cost function \(c_a(v_a)\) is given by a fixed travel time \(c^0_a\) and a volume dependent congestion function \(d_a(v_a)\) in the form

\[
c_a(v_a) = c^0_a (1 + d_a(v_a))
\]

The congestion function can be any non-decreasing function with \(d(0) = 0\). It models the discomfort of traveling on a segment at a volume \(v_a\). Since the function \(d(v_a)\) is specified as a network calculator expression, it can access any other attribute of the transit line as well, such as: headway, seated and total vehicle capacity, user attributes. By default, BPR-type and conical congestion functions are provided (Spiess, 1990), but the macro allows easy integration of other functional forms that might be required for particular applications.
During the assignment steps, the user defined segment attribute US1 will contain the value of the congestion function \( d_a(v_a) \).

In terms of the so defined congestion function \( d_a(v_a) \), the objective function of the equilibrium assignment separates in a (linear) travel time part \( T \) and a (non-linear) congestion part

\[
\min \sum_{a \in A} c_{ij}^0 v_{ij} + \sum_{i \in I, j \in R} w_{ij}^R + \sum_{a \in A} \int d_a(x) dx
\]  

(14)

The derivative of the objective function with respect to \( \lambda \), used to compute the optimal step length \( \lambda^k \), is:

\[
\sum_{a \in A} c_{ij}^0 \frac{d_a(v_{ij}^{k-1} + \lambda(\hat{v}_a - v_{ij}^{k-1})/(v_{ij}^{k-1} - \hat{v}_a))}{\lambda} + \lambda(\hat{T} - T^{k-1})
\]

(15)

**The CONGTRAS Macro:**

The solution algorithm for problem was embedded in an equilibrium transit assignment macro CONGTRAS (Congested Transit Assignment) macro. This macro computes the solution of the nonlinear cost transit assignment model with an efficient implementation of the linear approximation algorithm and by computing the necessary line search with a secant method.

**Table 1: EMME/2 equilibrium transit assignment macro.**

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
<th>Module:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Initialize congestion costs US1 to 0.</td>
<td>2.41</td>
</tr>
<tr>
<td>2</td>
<td>Compute total number of transit trips and initialize iteration counter ( k \leftarrow 0 ).</td>
<td>3.21</td>
</tr>
<tr>
<td>3</td>
<td>Perform uncongested fixed cost assignment to obtain ( v^0 ) and ( T^0 ).</td>
<td>5.11/5.31</td>
</tr>
<tr>
<td>4</td>
<td>Compute congestion cost.</td>
<td>2.41</td>
</tr>
<tr>
<td>5</td>
<td>Increment iteration counter ( k \leftarrow k + 1 ).</td>
<td>3.21</td>
</tr>
<tr>
<td>6</td>
<td>Compute new segment congestion costs ( c_{ij}^0 d_a(v_{ij}^{k-1}) ) into US1.</td>
<td>2.41</td>
</tr>
<tr>
<td>7</td>
<td>Perform fixed cost transit assignment with new congestion costs to obtain ( \hat{v} ) and ( \hat{T} ).</td>
<td>5.11/5.31</td>
</tr>
<tr>
<td>8</td>
<td>Compute stopping criterion GAP.</td>
<td>2.41</td>
</tr>
<tr>
<td>9</td>
<td>Perform line search for obtaining optimal step length ( \lambda^k ). This is implemented using the secant method to annul.</td>
<td>2.41</td>
</tr>
<tr>
<td>10</td>
<td>Update transit volumes ( v^k = v^{k-1} + \lambda^k (\hat{v}) ).</td>
<td>2.41</td>
</tr>
<tr>
<td>11</td>
<td>Update total travel time ( T^k = T^{k-1} + \lambda^k (\hat{T}) ).</td>
<td>3.21</td>
</tr>
<tr>
<td>12</td>
<td>Test normalized gap stopping criterion. If ( \text{GAP} &lt; \varepsilon ) then STOP, else continue with step 5.</td>
<td>2.41</td>
</tr>
</tbody>
</table>

For technical reasons, this step is implemented in the macro not in module 2.41, but using low level data manipulations in module 1.11.

In its current form, the CONGTRAS macro only considers crowding within the transit vehicles. But of course, as can be seen from the model formulation in the previous section, it is also possible to include congestion discomfort outside the transit vehicles, such as crowding on the platforms and on the pedestrian walk link - as long as the discomfort function is symmetric (i.e. the same travelers that are causing the congestion are also suffering the effects of it).

**Conclusion:**

The model may be used to compute approximate transit network equilibrium solutions when the vehicle capacities must be respected and the increased waiting times at stops are relevant. The model may be used even if relative gap can’t be computed with the standard approaches. The modeling congestion should be done using asymmetric congestion functions, e.g. as the perceived frequency of a line for a boarding passenger depending on the number of passengers already on board, or the dwell time of a line at a node depending on the number of boarding and alighting passengers. While it is indisputable that such phenomena occur in reality, including them into assignment models as the one described here unfortunately leads to models with non-unique solutions. Since the uniqueness of the solution is a primordial requirement for any assignment model, such asymmetric models, even those for which convergent algorithms are available, are of very limited practical use.

**REFERENCES**


