Decision Support Tool for Passenger Transportation Systems Planning

José Artur L. C. Marques¹, João V. da Fonseca Neto², and Fábio N. da Silva³

¹Universidade do CEUMA, Brazil, artur.marques@ceuma.br
²Universidade Federal do Maranhão (UFMA), Brazil, jviana@ufma.dee.br
³Universidade Federal do Maranhão (UFMA), Brazil, fnelasilva@hotmail.com

ABSTRACT
The authors propose a decision support tool for passenger transportation systems planning whose mathematical model is derived from the classical problem of transport. The tool main objective is to optimize a road passenger transit system fleet, describing the possible routes between each origin / destination pair meeting the cost constraints (profitability) and quality of service. It is discussed also the complexity of the nature of the dynamic of the transportation system. The mathematical modeling can show us correctly a way to qualify and quantify the routes, what helps to understand the dynamic behavior of the population through time, different situations and ages. Simulations present the viability of implantation for real time system to estimate transportation system and support decisions.

Keywords: Public transportation, transit planning, simulation tool.

I INTRODUCTION
In order to operate a passenger transportation system, it is necessary to plan lines among the various origin/destination pairs, so that the conveyance of users among the several parts of population agglomerate is optimized. The tool proposed intends to calculate the total fleet of this kind of system in order to globally optimize (minimizing) the number of vehicles without affecting the service quality, and seeking the possible minimum profitability of the service. This approach is very useful in transit planning.

The proposed tool needs a mathematical model that will be useful in simulations to transit systems tactical planning, such as the route analysis based on characteristics related to variables model, and the bus fleet needed to meet the passengers demands.

The transit system consists of multiple elements with non-linear interactions and the inherent unpredictability of several characteristics, which makes it necessary that some states could be represented by random variables. The nature of the problem is a discrete programming model that has proximity with the continuous linear so we could adapt the model to a canonical form of a transportation problem (Goldbarg, 2005).

II THE TRANSIT SYSTEM PROBLEM
In this Section is presented a transit system description and an optimization problem stated. As urban traffic is composed by many complex characteristics we consider some variables to characterize the model and the functional to quantify each of the routes, demand, profitability and ground fleet demand to achieve the optimization goal. An interesting model for vehicle routing problem is discussed in (Madsen, 1997).

A. Transit System Description
A mathematical model is needed to available vehicle fleet optimization, which must be matrix, so that it allows the simultaneous calculation to several routes of transit, where m(Origins) and n (Destinations).

The investigation of the transit system model, it can found which are the relevant spatial dimensions and term dimensions and the demand components that are equally relevant. In this proposition, the transit system modeling must take into consideration at least the variables shown in Figure 1.

![Model's Characteristic Variables Representation](image-url)
It is shown in Figure 1 that the route $R_{ij}$ to be chosen between the origin $i$ to destination $j$, must take into consideration the total distance to be covered $D_{ij}$, the distance $d_{ij}$, from population kernel to this route and the populations $x_{w,a}$ of these kernels, where $a$ is a reference to the chosen route.

These variables are all deterministic, excepting populations $x_{w,a}$ that tends to have a random term behavior in a known sample space.

In this course, the transporter agent, the vehicle, will transport passengers from one point to another in the route, not necessarily from origin to destination and this transportation is a partial course that is extremely benefiting to the profitability of the operation, because it generates additional profits, where is possible to convey a greater number of passengers than the maximum capacity of the vehicle. In this case, a rate of passenger renewal along the route is taken into consideration, which will qualify a “going up and down” of passengers and will make it possible that the number of conveyed passengers is larger than the maximum capacity of a ride. The challenge in planning (programming) and operate this system is to make it profitable with a minimum quality service, so the route with the greatest quality that is served by the smaller fleet.

B. Problem Statements

The network models allow the solution of important real problems, and are of extraordinary practical application (Cascetta, 2009).

The nature of the problem is of discrete programming and adapted to a canonical form of the transportation problem model.

III MODELS OF THE TRANSIT PROBLEM

In this Section will be discussed the classic transportation problem and define the cost and quality functions that will be used in the optimization model thereafter proposed.

A. The Classic Transportation Problem

The transportation problem can be seen as a flow problem in which the objective is to globally minimize the costs off low through the arcs (routes/paths).

The standard structure of a quantity transportation model can, therefore, be represented by

$$\min \sum_{i=1}^{m} \sum_{j=1}^{n} C_{ij}x_{ij}$$

subject to

$$\sum_{j=1}^{n} x_{ij} = a_{i}; \forall i = 1, 2, ..., m$$

$$\sum_{i=1}^{m} x_{ij} = b_{j}; \forall j = 1, 2, ..., n$$

$$x_{ij} \geq 0$$

Where $x_{ij}$ is the number of quantities conveyed from origin $i$ to destination $j$; $C_{ij}$ the weighting factor related to the transportation from origin $i$ to destination $j$, $a_{i}$ the capacity of origin $i$; $b_{j}$ the necessity of destination $j$; $n$ the number of destinations; $m$ the number of sources (origins). It is worth to point out that the weighting factor $C_{ij}$ is the cost of the conveyance of a unity of flow (passenger) through the arc $ij$ (path). The model must provide a calculation of factors $C_{ij}$ (origin-destination weight) of the objective function and the coefficients of the constraint function variables, along with the parameters of the functions of inequality/equality of the constraints. The optimization model formulation for the minimizing of transit time will be so that it corresponds with a linear optimization constrained structure as shown in Eq. (1), where $C_{ij}$ should be defined as the weights related to the several roads and our decision variable is the fleet.

In order to facilitate the comprehension, the numbers of destinations settings were changed to $m$ and the number of destinations were inverted to $n$, according to Eq. (1).

B. The Transportation Cost and Quality Functions

In this section will be discussed the definition of quality and cost functions:

The Passenger Transportation Quality Function. The public transportation service quality is function of the users accessibility to the bus lines in origin and destination, the distance covered by the vehicle from origin to destination, the performance of each ride (cycle time), on the term established by the route planning, the arrival time interval of each line vehicle from a given point and the maximum capacity allowed in the vehicle, that is:

$$Q = f(a,d,c,h,l)$$

Where $a$ is the accessibility (distance from origin to the boarding point and from the arrival point to the final destination), $d$ is straight course of routes (ratio between the shortest distance among routes between origin and arrival and distance of the route); $c$, reliance (percentage of programmed rides that were not performed in whole orin part, with delays greater than 5 min) is the ICV; $h$, headway (service intermission on a stop) $l$, capacity (maximum capacity/capacity of the bus). The capacity must be added the renewal rate (Ir).

For practical applications of these variables, we set: $a < 250 \text{ to } 300 \text{ m}$; $d < 1, 3$; $c \leq 98\% \text{ (ICV)}$; $h \leq 15 \text{ min}$; $l \leq 100\%$.

The Passenger Transportation Cost Function. The public transportation service cost is a function of the
average operational speed (expected value) in each line, the frequency (headway and fleet) of the vehicle characteristics and the number of passengers transported per kilometer, that is:

\[
C = f(Vo, f, c, Ipk) \quad (6)
\]

where \(Vo\), commercial or operating speed, \(f\) is the frequency of service (that is related to headway and fleet), \(c\) is the vehicle characteristics (engine, number of doors, capacity etc.), \(Ipk\) is passengers per km rate (influenced by the “going up and down” on the intermediate stops). The commercial speed \(Vo\) is equal to the ratio of the covered distance on the route \((d)\) and the cycle time \((Tc)\), that is:

\[
C = f(t, d, c, i)(7)
\]

Where \(t\) is boarding and arrival time, \(d\) is the distance between stops, \(c\) is the vehicle characteristics (acceleration and braking), \(i\) road interferences (lights, traffic jams, etc.). The variables characterization is performed over passenger conveyance time, from origin \(i\) to destination \(j\), and this time is variable during the operation period, hence, there is a \(T_{ij}(t)\) as the decision variable. However, this variable is dependent on others and the variable that is in fact related to time that is independent is the amount of the fleet \(F_{ij}(t)\) that is available for the passengers during the period \(t\). Therefore, the model must incorporate all the issues relative to the cost and quality functions for better representation of the phenomena.

Hence, there are variables to be considered in the quality function: accessibility: incorporated as \(d_{w,a}\) (distance from the population kernel \(w\) to the route \(a\) in the objective function); Headway is \(h_{max}\) that is incorporated as constraint of \(T_{ij}\); capacity: used as a weight calculation of the fleet, \(c\) (multiplied by the rate \(I\)) and the cost function of commercial speed that is \(V_{o,ij,a}\) (average commercial speed on route \(a\) of origin \(i\) to destination \(j\), incorporated in the objective function); the total covered distance (between origin and destination in route \(a\)): \(D_{ij,a}\) (with impact on \(I_{pk}\) in the objective function); population conveyed on a ride: \(X_{ij}\) (with impact on \(I_{pk}\) in the objective function), or the sum of \(x_{w,a}\) passenger per km rate, minimum: \(I_{pk}\) min as profitability constraint.

Therefore, \(I_{pk}\) min and \(h_{max}\) will be parameters. There are objective function variables for the system varying as: \(i\) origins varying up to \(n\); \(i = 1, 2, \ldots, n; j\) destinations varying up to \(m\), \(j = 1, 2, \ldots, m; a\) routes varying up to \(b, a = 1, 2, \ldots, b; w\) population kernels on route \(a\), varying up to \(k, w = 1, 2, \ldots, k\).

It is worth to point out that the model will adopt the premise that roads (routes) are not congested, therefore, variables of flow and density of vehicles on routes, that are interdependent, will not be explicitly represented, although average speed variable is influenced by these variables.

C. Passenger Transportation System Optimization Model

Two constraints showed are the most significant in model formulation:

1) Headway: passenger waiting time between two buses of the same line, at the same point. A maximum value \(h_{max}\) must be adopted for the system under study, as a quality constraint.

2) Profitability: the ratio cost/benefit to the line operating company must have a minimum value that, here, will be represented by a representative rate of passengers transported by the fleet on a given route, considering all the population kernels along it. This parameter will be assessed by the passenger per km rate, \(I_{pk}\).

Another important quality constraint is capacity (capacity of the bus) that will be incorporated in the objective function as a weighing term \(1/C_{vs}\), multiplied by the renewal rate \(f_{r}\).

The optimization model is centered in the decision variable related to the choice of the fleet amount that has to be designated for each line on each route in the several schedules during the day, so that the transit total time \(T_{ij,a}\) is minimized from \(n\) origins to \(m\) destinations through \(b\) routes and with a well-balanced profitability among the several lines of the whole system.

The proposed optimization model has the following form:

\[
\min_{F_{ij,a}} \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{a=1}^{b} T_{ij,a}(F_{ij,a}(t)) = \frac{R_{ij,a}(t)}{F_{ij,a}(t)} \quad (8)
\]

subject to

\[
\frac{1}{b \cdot m} \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{a=1}^{b} T_{ij,a}(t) \leq headway_{max} \quad (9)
\]

\[
\frac{1}{b \cdot c_r} \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{a=1}^{b} D_{ij,a} \geq I_{pk \ min} \quad (10)
\]

\[
\frac{1}{F_{ij,a}} \geq 0 \quad (11)
\]

Where \(R_{ij,a}(t)\), the distance of the route \(a\) (course on the routes) between origin \(i\) and destination \(j\); \(F_{ij,a}(t)\), the fleet(amount of buses) used on each route.

The total system time is directly proportional to the chosen routes, and inversely proportional to the fleet. Hence, the fleet of every route will be the inverse of the decision variables values.

This time is subjected to the constraint of the adopted head way limit. In the formulation, it is inserted the weighted by the inverse of number \(b\) of routes and \(m\) destinations.
Therefore, the average time (headway) is compared to the maximum adopted time. 

In the second constraint, related to profitability, weighing is done again by the number of routes \( b \) and by a coefficient of profitability \( cr \), related to the amount of paying passengers on a ride from origin \( i \) to destination \( j \) that is also influenced by the passengers that go up and down along the route. Hence, the whole system has to achieve a minimum profitability value. The variable \( D_{ij} \) is the distance covered on route \( ij \).

A third constraint has to be added: the non-negativity of the decision variable \( F_{ij}(t) \), apart from the physical obvious meaning, this constraint directs adequately the search for the solution.

Thus, it is defined the metrics that will be used for the calculation of the employed weighting in the qualification of routes \( R_{ij} \), with origin \( i \) and destination \( j \) going through route \( a \):

\[
R_{ij}(D_{ij},(t),V_{ij},a(t),x_{w,a}(t),d_{w,a}) = \frac{1}{\sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{a=1}^{b} \frac{x_{w,a}(t) \cdot V_{ij,a}(t)}{Cv \cdot I_{r} \cdot k \cdot \sum_{a=1}^{b} d_{w,a} \cdot D_{ij,a}(t)}} (12)
\]

where the known values of \( F_{ij}(Cv,X_{ij}(t)) \) are given by

\[
F_{ij}(Cv,X_{ij}(t)) = \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{X_{ij}(t)}{Cv} (13)
\]

It means that the route \( R_{ij} \) is function of the covered distance on the operating speed \( V_{ij,a} \), possible on route \( a \), the distance \( d_{w,a} \) from de population kernels to this route and by the factor \( x_{w,a} \) from the populations of these kernels.

Thus, the route \( R_{ij} \) becomes proportional to the served population \( x_{w,a} \) is weighted by the total \( k \) of population kernels to obtain the average, by the operating speed \( V_{ij} \) and inversely proportional to the distances of these population kernels to the route, and the total covered distance on route \( D_{ij} \).

The fleet \( F_{ij} \) is the inverse function of the capacity of vehicles \( C_v \) and of the population to be served (conveyed) on each ride \( X_{ij} \).

Replacing the Eq. (12) in the optimization structure of Eq. (13) there will be the following model

\[
\min_{F_{ij},a} \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{a=1}^{b} T_{ij,a} (F_{ij,a}(t)) = \frac{1}{\sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{a=1}^{b} \frac{x_{w,a}(t) \cdot V_{ij,a}(t)}{Cv \cdot I_{r} \cdot k \cdot \sum_{a=1}^{b} d_{w,a} \cdot D_{ij,a}(t) \cdot F_{ij,a}(t)}} (14)
\]

subject to

\[
\frac{1}{b \cdot m} \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{a=1}^{b} T_{ij,a}(t) \leq \text{headway}_\text{max} (15)
\]

IV THE CLASSICAL SOLUTION APPROACH

In the sense of linear programming technics the solution technic is called classical solution approach, based on the algebraic solution of the linear system \( Ax = b \), due to its counterpart approach that is bio-inspired solutions, such as genetic algorithms and neural networks and others (De Carvalho,1992).

The suggested model refers us to integer programming situation of which methods of solution are heuristic. The approach of this study will be of an approximation using the method of continuous linear programming. Considering that the results are real numbers to the decision variables, and that integers are expected in the solution, it must be verified if the results are within the expected, with the due approximations.

There are several methods to solve continuous linear programming problems. The Simplex method is the classical method and the best-known, having an excellent performance in the simulation convergence. Among other solution methods, one can point out the interior-point methods.

In order for the optimization method to be applied with the Simplex method (Simplex Table) it is necessary to transform the constraints inequalities into equalities, adding the residual variables. However, in case one adopts the simulation in the software MATLAB (Optimization Toolbox), there is no need for that adapting, it is only necessary to put all constraints as inequalities with the operator \( \leq \) (less than or equal to). In order to do that, the second inequality of the suggested model has to have the signs of its terms inverted.

Initially, the terms on the right of the inequalities will be named by the classical form, \( b_i \), with \( i \) varying from 1 ton:

\[
\frac{1}{b \cdot m} \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{a=1}^{b} T_{ij,a}(t) \leq b_1 (18)
\]

\[
\frac{1}{b \cdot m} \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{a=1}^{b} D_{ij,a} \geq b_2 (19)
\]

Where \( b_i \) is the maximum headway and \( b_2 \) is the minimum \( I_{ij} \). The signs of the second inequality were inverted. Thus, the suggested model takes the following form:
\[
\min \sum_{ij} \sum_{j} \sum_{a} \frac{1}{C_i \cdot Ir \cdot k} \sum_{w} x_{w,a}(t) \cdot V_{ij,a}(t)
\]

subject to

\[
\sum_{ij} \sum_{j} \sum_{a} T_{ij,a}(F_{ij,a}(t)) \leq b_i
\]

\[
-\frac{1}{b \cdot cr} \sum_{ij} \sum_{j} \sum_{a} D_{ij,a} \leq -b_2
\]

\[
\frac{1}{F_{ij,a}} \geq 0
\]

In order to solve the model, it has to be applied with variables values that will result in decision variable coefficients values that will give the canonical form of the transportation problem. The decision variables values, that will be as many as the existent origin-destination pairs, will turn-out to be the inverse of the values that we wish to obtain: the optimal fleet to each origin-destination route with a established route qualification.

V COMPUTATIONAL EXPERIMENTS

The proposed tool suggests simulations performed in Optimization Toolbox MATLAB software, considered as a case of linear programming.

A simple case, compatible with real urban transit system data, of which results allows one to make conclusions about this tool.

On validity first stage, it is necessary to insert test case data, in order to obtain an equation system in the optimization problem canonical form with transportation constraints.

The Linprog Solver in MATLAB Toolbox, allows simulations through three different algorithms: Large Scale, based on the Y. Zhang LIPSOL (Linear Programming Interior-Point Solvers) method, which is a variant of the method of a primal-dual interior-point method, and a more adequate for the great dimension variables space; Medium Scale, that is a variant of sequential quadratic programming, with the quadratic term equal to zero; and the simplex method.

The simulations were carried out using all of them in order to evaluate the tool performance, and verify which method performs the best to our search.

A. Test Case

The tested example considers: two origins, two destinations, two routes to each origin-destination pair, two population kernels on each route; and 100-passenger capacity vehicles, and a renewal rate of 1.5, that are \( b_n = 2 \); (origin), \( m = 2 \); (destination), \( h = 2 \); (routes), \( k = 2 \); (population kernels) \( Cv = 100 \). \( (capacity) \), \( Ir = 1.5 \) (turnover). These parameters: \( l_{pd} = 2.5 \) passengers/km, \( h = 15 \) min; \( cr = 1 \) (all paying passengers).

Moreover, the calculation will be carried out for a specific time (rush hour), that is, the function will not vary through time.

Objective Function \( f(x) \) Calculation. The objective function is written in following form:

\[
\begin{align*}
& f(x) = 0.56 \cdot x_1 + 0.88 \cdot x_2 + 0.896 x_3 + 0.96 x_4 \\
& + 1.18 \cdot x_5 + 0.93 \cdot x_6 + 0.65 \cdot x_7 + 0.93 \cdot x_8
\end{align*}
\]

Constraint Function \( g_1(x) \) Calculation. The restriction \( g_1(x) \) is, in fact, the objective function, that calculates the total time, pondered by the number of routes, \( b \), and the number of destinations, \( m \). If the parameter \( h = 15 \) min (0.25h) is considered,

\[
g_1(x) = 0.14 x_1 + 0.22 x_2 + 0.22 x_3 + 0.242 x_4 + 0.295 x_5 + 0.233 x_6 + 0.164 x_7 + 0.233 x_8
\]

Constraint Function \( g_2(x) \) Calculation. The constraint function \( g_2(x) \) is given by

\[
g_2(x) = -20 x_1 - 20 x_2 - 23 x_3 - 22 x_4 - 18 x_5 - 17.5 x_6 - 30 x_7 - 18 x_8
\]

B. Solutions

The results found, according the solution feasible space to the bounds, are shown in Table I.
Stopping criteria default values, 10 times the number of variables, and tolerance function among interactions, $1 \times 10^6$ were used. Simulations with the three available algorithms on the Solver Linprog were carried out, and the results were absolutely the same, but the Simplex algorithm had the best performance when it came to the number of necessary interactions to reach the solution (one to eight interactions), which shows its robustness and efficiency. Considering that the number of decision variables is small, in this test case, that must have been determinant for the Simplex method to have had the best performance.

The variable values that are searched are the inverse of decision variable values, thus the values in Table 1 are inverted so that the values of Table 2 are found.

<table>
<thead>
<tr>
<th>Variable ($F_i$)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_1$</td>
<td>10</td>
</tr>
<tr>
<td>$F_2$</td>
<td>10</td>
</tr>
<tr>
<td>$F_3$</td>
<td>10</td>
</tr>
<tr>
<td>$F_4$</td>
<td>12.5 (13)</td>
</tr>
<tr>
<td>$F_5$</td>
<td>20</td>
</tr>
<tr>
<td>$F_6$</td>
<td>14.2 (14)</td>
</tr>
<tr>
<td>$F_7$</td>
<td>17.7 (18)</td>
</tr>
<tr>
<td>$F_8$</td>
<td>14.2 (14)</td>
</tr>
</tbody>
</table>

A variables sensitivity analysis suggests that:

1) The tool has a good performance for great values of distances $D_{ij,a}$ and $d_{w,a}$ and small values of $x_{w,a}$ and $V_0$ that result together in objective function coefficients with approximate values that are less than 1;

2) In the problem minimization course time, it is evident that the quality function will have the greatest influence in the results, and, consequently, with higher values of fleet;

3) The variations of quality constraint patterns (headway) did not influence the result, and the profitability constraint patterns variations (cost) only influence the result if it is increased in great proportions.

The results found allow to state that the tool tends to easily measure the fleet, and with values that are greater than the used in practice, more than that can be minimized by the use of a more severe constraint related to cost (profitability), such as the minimum capacity of the vehicle that is able to make the operation profitable.

VI CONCLUSION

One of most important of transit managers challenges is to decide properly about public transportation fleet size. The proposed tool, using the well-known MATLAB software had presented satisfactory performance in simulations. The sensitivities analysis established that the tool performance is satisfactory always for great distance values in the origin-destination pairs, in urban center scale, as well as to population values served along the routes. The tool tends to have an outstanding characteristic, an increase in value of the service quality function, because it intended to minimize the vehicle’s course time (ride), although there was a profitability constraint, with that generating results for the decision variables (fleet) higher than the ones that are common in real transit systems, that maybe improved with a cost constraint based in the minimum capacity of the vehicles.

ACKNOWLEDGMENT

The authors would like to thank PPGEE/UFMA, University of CEUMA, FAPEMA, CNPQ, CAPES and São Luís City Council for the support to develop this research.

REFERENCES


