

Urban rapid transit network capacity expansion

Ángel Marín , Patricia Jaramillo

Universidad Politécnica de Madrid, E.T.S.I. Aeronáuticos, Plaza Cardenal Cisneros, 3, Madrid 28040, Spain

Universidad Nacional de Colombia, Sede Medellín, Facultad de Minas, Crr 80 No 65-223, Medellín, Colombia

Abstract

This paper examines a multi-period capacity expansion problem for rapid transit network design. The capacity expansion is realized through the location of train alignments and stations in an urban traffic context by selecting the time periods. The model maximizes the public transportation demand using a limited budget and designing lines for each period. The location problem incorporates the user decisions about mode and route. The network capacity expansion is a long-term planning problem because the network is built over several periods, in which the data (demand, resource price, etc.) are changing like the real problem changes. This complex problem cannot be solved by branch and bound, and for this reason, a heuristic approach has been defined in order to solve it. Both methods have been experimented in test networks.

Keywords: Capacity expansion; Multiperiod; Rapid transit network design; Branch and bound; Heuristics

1. Introduction

Increasing mobility and longer journeys due to the growth of cities, traffic congestion in city centres and in entrance corridors are some of the reasons why during the last few years new lines of rail transit systems (metro, light rail, etc.) have been constructed or extended.

The transit network design problem considers at upper level the location decisions and at lower level the routing user decisions. At upper level the maximum coverage of the demand for public network is the main goal, taking a list of potential rapid transit corridors and stations and budget into account. At lower level the user traffic behaviour is considered. The way of selecting and comparing these network alternatives is performed by considering that the demand chooses path and mode depending on the network supplied, considering the traffic costs.

Bruno et al. (2002) maximize the travel coverage by public network. Bruno et al. (1999) and Laporte et al. (2005) incorporate the model data of the origin–destination matrix. The papers of Laporte et al. (2002) and

Hamacher et al. (2001) deal with the problem of locating stations on a given alignment. García and Marín (2001, 2002) study the transit network design problem using bilevel programming. They consider the multi-modal traffic assignment problem with combined mode at lower level.

Laporte et al. (in press) extends on the previous models by incorporating the station location problem, the alternative of several lines and defining the model using the maximum coverage of the public demand as an objective function and the budget constraints as side constraints. Marín (in press) is an extension of the above, where the train lines are not initially given and the lines do not have fixed origins and destinations.

The capacity expansion is a long-term planning problem and its necessary extension must be carefully studied given that the infrastructure construction is expensive and produces significant disturbance in city life, therefore it is realized in stages over several years. During this long-term planning period the problem data can change a great deal from beginning to end: construction price, inflation, travel demand, etc. These changes cannot be ignored in a real model.

The paper is organized as follows: In Section 2 the rapid transit network design capacity expansion (RTNDCE) model is discussed. Because of its large scale, it is difficult to solve the expansion model for medium-sized cities. In Section 3 we define a heuristic to reduce the computational time. In Section 4 different test networks are defined in order to implement the experiments comparing the use of classical algorithms such as branch and bound and the previous heuristic. Both methods are compared for different parameters. Finally the conclusions and the references are considered.

2. Rapid transit network design capacity expansion model

The rapid transit network design capacity expansion, RTNDCE, model is an extension of the RTND problem defined in Laporte et al. (in press) and Marín (in press). Its formulation is based on considering the decisions during certain periods of time, t , belonging to the planning set $T = \{1, \dots, |T|\}$. Each period refers to the unit of time from time point t to $(t + 1)$. It is assumed that the nodes and edges can be installed at any point in time within a certain time period. However, the costs incurred during this time are assumed to take place at the commencement of the time period in which the changes take place. RTND studies the locations of stations and alignments connecting them with a finite number of transit lines: $L = \{l = 1, \dots, |L|\}$, where the lines do not have predetermined origins and destinations and the number of lines is a variable but the maximum number of lines $|L|$ is fixed.

2.1. Data

The data required for the model are as follows:

1. For key stations, the set of potential locations is $N = \{i = 1, \dots, |N|\}$. From that the set E of feasible (bidirectional) edges linking the key stations N is defined. Therefore, we have a potential network (N, E) from which the optimum rapid transit network is selected. Let us denote by $N(i) = \{j : \exists a \in E, a = (i, j)\}$ the set of nodes adjacent to node i .
2. Each feasible edge has an associated length $d = (d_{ij})$. The length of the edges usually corresponds to the Euclidean distance between pairs of nodes (i, j) if the system is underground, and street distance if it is at grade. However, forbidden regions will increase the distance and d can also be interpreted as the generalized cost of using the arc a .
3. The demand is given by the origin/destination (o, d) pairs of nodes. The demand at pair w at period t is given by the matrix: $G^t = (g_w^t) : w = (o, d) \in W \forall t \in T$, where W is the set of ordered pair of demands.
4. Let c_{ij}^l and c_i^l be the cost of constructing a section of line l in edge ij and a station on line l at node i at period t . The upper bound to construction costs at each period t is a budget of c_{\max}^t .
5. The generalized cost of satisfying the demand of pair w through the private and the public network at period t are $u_w^{\text{pri},t}$ and $u_w^{\text{pub}}(t)$, respectively. Note that the latter cost is a variable that depends on the final topology of the public network and therefore on the edges that are selected, meanwhile $u_w^{\text{pri},t}$ are input data.

2.2. Variables

The decisions would be considered to be temporally disaggregated, which are described by

- $y_i^l(t) = 1$, if line $l \in L$ is located using the node i at period t , 0 otherwise.
- $x_{ij}^l(t) = 1$, if line $l \in L$ is defined using the edge ij at period t , 0 otherwise.
- $h_l(t) = 1$, if the line l has at least a link located at period t , 0 otherwise.
- $f_{ij}^w(t) = 1$, if the demand of the pair w uses edge ij at period t , 0 otherwise.
- $p_w(t) = 1$, if the demand w uses the public mode at period t , 0 otherwise.

2.3. Objective function

The RTND objective function $z = \sum_{t \in T} z(t)$ is the combination of the objective function at each period t . The objective function is defined by the following adimensionalized terms:

- To maximize the public demand (pubd) covering:

$$z_{\text{pubd}}(t) = \frac{\sum_{w \in W} g_w^t p_w(t)}{\sum_{w \in W} g_w^t}. \quad (1)$$

- To minimize the routing costs (rc): $z_{\text{rc}}(t) = \frac{\sum_{w \in W} [\mu^t u_w^{\text{pri},t} (1 - p_w(t)) + u_w^{\text{pub}}(t)]}{\sum_{w \in W} [\mu^t u_w^{\text{pri},t}]}$, where the public cost is defined by

$$u_w^{\text{pub}}(t) = \frac{1}{\lambda^t} \sum_{ij \in A} d_{ij} f_{ij}^w(t) \quad \forall w \in W \quad \forall t \in T, \quad (2)$$

where λ^t is the average velocity of the vehicles at transit edges and μ^t is a factor that represents the congestion.

- To minimize the location costs (lc): $z_{\text{lc}}(t) = \frac{z_{\text{loc}}(t)}{c_{\text{max}}^t} \quad \forall t \in T$, where the terms $z_{\text{loc}}(t)$ are defined by

$$z_{\text{loc}}(t) = \sum_{l \in L} \left(\sum_{(i,j) \in A; i < j} c_{ij}^l (x_{ij}^l(t) - x_{ij}^l(t-1)) + \sum_{i \in N} c_i^l (y_i^l(t) - y_i^l(t-1)) \right) \quad \forall t \in T, \quad t > 1, \quad (3)$$

$$z_{\text{loc}}(1) = \sum_{l \in L} \left(\sum_{(i,j) \in A; i < j} c_{ij}^l (x_{ij}^l(1)) + \sum_{i \in N} c_i^l (y_i^l(1)) \right). \quad (4)$$

- To minimize the number of private demand (prid):

$$z_{\text{prid}}(t) = \frac{\sum_{w \in W} (1 - p_w(t))}{|W|}. \quad (5)$$

RTND at each period t minimize the objective function defined by

$$z(t) = -\eta z_{\text{pubd}}(t) + \left(\frac{1-\eta}{4} \right) z_{\text{rc}}(t) + \left(\frac{1-\eta}{4} \right) z_{\text{lc}}(t) + \left(\frac{1-\eta}{2} \right) z_{\text{prid}}(t), \quad (6)$$

where η is a number close to 1, because the public trip covering, $z_{\text{pubd}}(t)$ is the main component of the objective function. The other terms are included to adequately simulate the routing user behaviour, and that the location of any facility is not free of cost.

2.4. Constraints

- Construction cost constraints (CCC). The CCC bounds the construction cost at each period as follow:

$$\text{CCC}(t) : \quad z_{\text{loc}}(t) \leq c_{\text{max}}^t \quad \forall t \in T. \quad (7)$$

- Routing demand constraints (RDC). The multi-commodity flow conservation at each node is assumed, for each demand and for each period as follow:

$$\text{RDC}(i, w, t) : \sum_{k \in N(i)} f_{ki}^w(t) - \sum_{j \in N(i)} f_{ij}^w(t) = i_w(t) \quad \forall i \in N \quad \forall w \in W \quad \forall t \in T, \quad (8)$$

where

$$i_w(t) = \begin{cases} -1 & \text{if } i = o : \forall (o, d) = w \in W \\ 1 & \text{if } i = d : \forall (o, d) = w \in W \\ 0, & \text{otherwise} \end{cases} \quad \forall i \in N \quad \forall w \in W \quad \forall t \in T.$$

In all the constraints the disaggregated level is shown between parentheses.

- The Line location constraints (LLC(l, t)) are as follow:

$$x_{ij}^l(t) \leq y_i^l(t) \quad \forall (i, j) \in A, \quad i < j \quad \forall l \in L \quad \forall t \in T, \quad (9)$$

$$x_{ij}^l(t) \leq y_j^l(t) \quad \forall (i, j) \in A, \quad i < j \quad \forall l \in L \quad \forall t \in T, \quad (10)$$

$$\sum_{j \in N(i), i < j} x_{ij}^l(t) + \sum_{j \in N(i), j < i} x_{ji}^l(t) \leq 2 \quad \forall i \in N \quad \forall l \in L \quad \forall t \in T, \quad (11)$$

$$h_l(t) + \sum_{\substack{(i,j) \in A \\ i < j}} x_{ij}^l(t) = \sum_{i \in N} y_i^l(t) \quad \forall l \in L \quad \forall t \in T, \quad (12)$$

$$\sum_{\substack{(i,j) \in A \\ i < j}} x_{ij}^l(t) \leq M_1 h_l(t) \quad \forall l \in L \quad \forall t \in T, \quad (13)$$

$$\sum_{\substack{(i,j) \in A \\ i < j}} x_{ij}^l(t) \geq h_l(t) \quad \forall l \in L \quad \forall t \in T, \quad (14)$$

$$\sum_{i \in B} \sum_{j \in B} x_{ij}^l(t) \leq |B| - 1 \quad \forall B \subset N, \quad |B| \geq 2 \quad \forall l \in L, \quad t = T. \quad (15)$$

LLC(l, t) is separable by lines l and periods t . The constraints (9) and (10) ensure that the links are not located if their origin and destination nodes are not previously located. The constraints (11) require that each node in the line does not have more than two incident edges. The constraints (12)–(14) ensure that a line is activated if, at least, a link is constructed within it. M_1 is a large enough number: $M_1 \geq |A|$.

The constraints (15) require that the lines do not make cycles. This last constraint is not explicitly assumed in the model, however when a cycle appears in the optimal solution, an explicit constraint is added and the model is run again. B is any subset of N .

- The mode demand splitting constraints (MDSC) produce an all or nothing mode assignment: if $u_w^{\text{pri},t}$ is inferior or equal to $u_w^{\text{pub}}(t)$, the demand is assigned to mode PRI ($p_w(t) = 0$) and if not, it is assigned to PUB ($p_w(t) = 1$):

$$\text{MDSC}(w, t) : u_w^{\text{pub}}(t) - \mu^t u_w^{\text{pri},t} - M_2(1 - p_w(t)) \leq 0 \quad \forall w \in W \quad \forall t \in T, \quad (16)$$

where M_2 is a large enough number:

$$M_2 \geq \frac{1}{\lambda} \sum_{(i,j) \in A} d_{ij} - \max_{w \in W, t \in T} \{\mu^t u_w^{\text{pri},t}\}.$$

- The location-allocations constraints (LAC) are as follow:

LAC(ij, w, t):

$$f_{ij}^w(t) + f_{ji}^w(t) \leq \sum_{l \in L} x_{ij}^l(t) \quad \forall (i, j) \in A, \quad i < j, \quad \forall w \in W \quad \forall t \in T, \quad (17)$$

$$f_{ij}^w(t) \leq \sum_{l \in L} y_j^l(t) \quad \forall i = o, \quad \text{if } (o, d) = w; \quad \forall w \in W, \quad \forall t \in T, \quad (18)$$

$$f_{ij}^w(t) \leq \sum_{l \in L} y_i^l(t), \quad \forall j = d, \quad \text{if } (o, d) = w; \quad \forall w \in W \quad \forall t \in T. \quad (19)$$

- RTNDCE considering the above constraints is separable by period time, but some additional constraints establish the relations between the separated RTNDCE(t), they are the logical constraints (LogC), so if a station or a link is located in a period, they are obligatory located in future periods.

$$\text{LogC}(l) : y_i^l(t) \geq y_i^l(t-1) \quad \forall i \in N \quad \forall l \in L \quad \forall t \in T, \quad t > 1, \quad (20)$$

$$x_{ij}^l(t) \geq x_{ij}^l(t-1) \quad \forall (i, j), \quad i < j \quad \forall l \in L \quad \forall t \in T, \quad t > 1, \quad (21)$$

$$h^l(t) \geq h^l(t-1) \quad \forall l \in L \quad \forall t \in T, \quad t > 1. \quad (22)$$

2.5. RTNDCE shortly formulation

The RTNDCE may be expressed in terms of the above constraints as follows:

$$\text{Min}_{x,y,h,f,p \in \{0,1\}} z = \sum_{t \in T} z(t)$$

subject to : CCC(t), RDC(i, w, t), LLC(l, t), MDSC(w, t), LAC(ij, w, t), LogC(l).

3. Heuristic methodology

The problem of RTNDCE is its large size. It is easy to calculate the number of RTND variables for a network like that of Seville, which will be introduced in the next section. For this network with data set: $|N| = 24$, $|E| = 264$, $|W| = 552$, $|L| = 5$, $|T| = 5$, the number of variables is 735325 and the number of constraints is 1153575, without taking into account the anti-cycle constraints. For these large size problems, branch and bound cannot obtain the optimal solution and a heuristic approach is justified.

RTNDCE is time separable if the constraints LogC(l) are not considered, so a heuristic has been defined in order to relax them. For each period t , we define a relaxed RTND, RRTND(t) model with the above constraints (1)–(19) considering the input of each model the set of the installed lines at previous periods. This set is updated in each period t with the optimal lines $\{y^{l*}, x^{l*}\}$ installed in the previous one.

$$L_0(t) = \left\{ l \in L : y_i^{l*}(t-1) = 1, x_{ij}^{l*}(t-1) = 1 \right\}. \quad (23)$$

In the first period, $L_0(t)$ is defined with the initial installed lines if some exists. At each period a new model RRTND(t) is defined using the updated $L_0(t)$ as a reference. We try to locate the node and link forming lines, but we consider that some lines have already been located. At each period some previous lines can be extended or new lines can be assigned. At each period t , $L(t)$ is updated as: $L(t) \leftarrow L_0(t) \cup L_1(t)$, where $L_0(t)$ is the above set and $L_1(t)$ is the set of the alternative lines to be located at period t .

At RRTND(t) initial lines are data: $\bar{y}_i^{l'} = 1$ and $\bar{x}_{ij}^{l'} = 1$ for all l' belong to $L_0(t)$, and the following variables are decision variables, defined in the set of the alternative lines:

- $y_i^{l'}(t) = 1$, if line $l' \in L_1(t)$ is located using the node i , 0 otherwise.
- $x_{ij}^{l'}(t) = 1$, if line $l' \in L_1(t)$ is located using the edge ij , 0 otherwise.

RRTND(t) is defined from the above RTNDCE considering $L(t)$ instead of the previous L , and considering the modified LAC as follows:

$$\text{LAC: } f_{ij}^w(t) + f_{ji}^w(t) \leq \sum_{l \in L_1(t)} x_{ij}^l + \sum_{l' \in L_0(t)} \bar{x}_{ij}^{l'} \quad \forall (i, j) \in A, \quad i < j \quad \forall w \in W \quad \forall t \in T, \quad (24)$$

$$f_{ij}^w \leq \sum_{l \in L_1(t)} y_j^l + \sum_{l' \in L_0(t)} \bar{y}_j^{l'} \quad \forall i = o, \text{ if } (o, d) = w; \quad \forall w \in W \quad \forall t \in T, \quad (25)$$

$$f_{ij}^w \leq \sum_{l \in L_1(t)} y_i^l + \sum_{l' \in L_0(t)} \bar{y}_i^{l'} \quad \forall j = d, \text{ if } (o, d) = w; \quad \forall w \in W \quad \forall t \in T. \quad (26)$$

RRTND(t) parameters must be adequately modified according to its evolution over time.

3.1. Heuristic algorithm

With the previous considering the heuristic algorithm is defined as follows:

1. **Initialization:** Define the data to be used during the capacity expansion period t : $c_{ij}^t, c_i^t, c_{\max}^t, u_w^{\text{pri},t}, g_w^t, \mu^t$. Define an initial $L_0(0)$ by the lines initially located.
2. **Do for each t until $|T|$:**
 Redefine $L_0(t)$ taking as data the last y_i^{l*}, x_{ij}^{l*} .
 Run RRTND(t) using branch and bound. Let y_i^{l*}, x_{ij}^{l*} be the optimal solution.
End Do.

4. Computational experiments

In the computational experiments the model RTNDCE is tested to verify its utility and the RTNDCE and heuristic solutions are compared.

Branch and bound is used to solve RTNDCE and each RRTND(t) in the Heuristic algorithm. Branch and Bound have been implemented with the help of Gams 21.7 which calls CPLEX 9.0, using a laptop with a processor AMD Athlon 64 \times 2 Dual, 4200+, 2.22 GHz, 2.00 GB RAM memory.

In the RTNDCE the annual parameter changes have been defined by the following exponential growth factors:

- Fca: alignment construction cost uniform incremental factor.
- Fcn: station construction cost uniform incremental factor.
- Fb: budget uniform incremental factor.
- Fg: demand uniform incremental factor.

In each period t , the alignment construction costs are defined by: $c_{ij}^t = \text{Fca}^{t-1} c_{ij}^1$. The station construction costs are: $c_i^t = \text{Fcn}^{t-1} c_i^1$. The construction budget is: $c_{\max}^t = \text{Fb}^{t-1} c_{\max}^1$. The demand is defined by: $g_w^t = \text{Fg}^{t-1} g_w^1$. The private user cost is affected by the congestion incremental factor by $u_w^{\text{pri},t} = \text{Fb}^{t-1} u_w^{\text{pri},1}$. In all cases the value

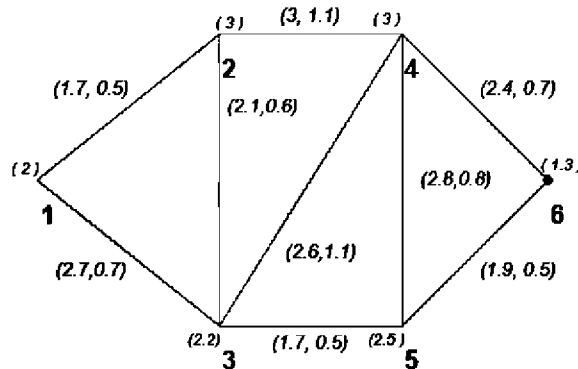


Fig. 1. Test network R1.

in the first period is taken as reference. In the computational experiments the values are as follows: $F_{ca} = 1.03$, $F_{cn} = 1.03$, $F_b = 1/0.95$, $F_g = 1.025$. The congestion factor is $\mu^t = 1.05$. In these experiments, the relative weight of the first term of the objective function in relation to the others has been taken as $\eta = 0.96$.

A first experiment uses the network previously used by Laporte et al. (2004). This network is denoted by R1 and has six nodes and nine edges. This network is shown in Fig. 1. Each node has an associated initial construction cost c_i and each edge a pair (c_{ij}, d_{ij}) of weights: the construction cost and the distance.

The origin–destination demand g_w in the first period is given by the following matrix G :

$$G = \begin{pmatrix} - & 9 & 26 & 19 & 13 & 12 \\ 11 & - & 14 & 26 & 7 & 18 \\ 30 & 19 & - & 30 & 24 & 8 \\ 21 & 9 & 11 & - & 22 & 16 \\ 14 & 14 & 8 & 9 & - & 20 \\ 26 & 1 & 22 & 24 & 13 & - \end{pmatrix}.$$

The private user cost matrix u_w^{pri} in the first period is defined by the following matrix U^{pri} :

$$U^{\text{pri}} = \begin{pmatrix} - & 1.6 & 0.8 & 2 & 2.6 & 2.5 \\ 2 & - & 0.9 & 1.2 & 1.5 & 2.5 \\ 1.5 & 1.4 & - & 1.3 & 0.9 & 2 \\ 1.9 & 2 & 1.9 & - & 1.8 & 2 \\ 3 & 1.5 & 2 & 2 & - & 1.5 \\ 2.1 & 2.7 & 2.2 & 1 & 1.5 & - \end{pmatrix}.$$

A second larger network R2 has also been defined. The network R2 has nine nodes and 16 edges. The network R1 is obtained from R2 by deleting the nodes 7, 8 and 9 and their adjacent edges. The network R2 is shown in Fig. 2.

As in R1, in R2 each node i has an associated initial construction cost c_i and each edge ij a pair (c_{ij}, d_{ij}) of weights: the initial construction cost c_{ij} and the distance d_{ij} . The origin–destination demand g_w for each $w \in W$ in the first period is defined by the matrix G :

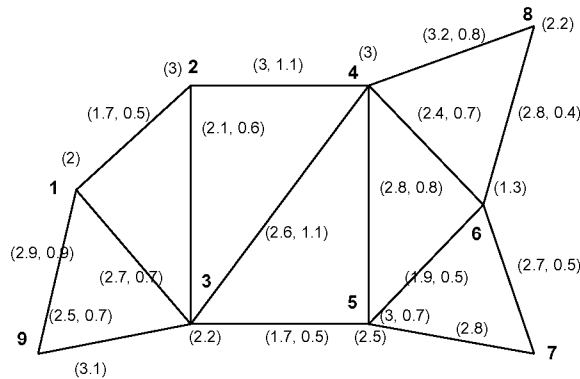


Fig. 2. Test network R2.

$$G = \begin{pmatrix} - & 9 & 26 & 19 & 13 & 12 & 13 & 8 & 11 \\ 11 & - & 14 & 26 & 7 & 18 & 3 & 6 & 12 \\ 30 & 19 & - & 30 & 24 & 8 & 15 & 12 & 5 \\ 21 & 9 & 11 & - & 22 & 16 & 25 & 21 & 23 \\ 14 & 14 & 8 & 9 & - & 20 & 16 & 22 & 21 \\ 26 & 1 & 22 & 24 & 13 & - & 16 & 14 & 12 \\ 8 & 6 & 9 & 23 & 6 & 13 & - & 11 & 11 \\ 9 & 2 & 14 & 20 & 18 & 16 & 11 & - & 4 \\ 8 & 7 & 11 & 22 & 27 & 17 & 8 & 12 & - \end{pmatrix}.$$

The private cost u_w^{pri} for each $w \in W$ in the first period is defined by the matrix U^{pri} :

$$U^{\text{pri}} = \begin{pmatrix} - & 1.6 & 0.8 & 2 & 2.6 & 2.5 & 3 & 2.5 & 0.8 \\ 2 & - & 0.9 & 1.2 & 1.5 & 2.5 & 2.7 & 2.4 & 1.8 \\ 1.5 & 1.4 & - & 1.3 & 0.9 & 2 & 1.6 & 2.3 & 0.9 \\ 1.9 & 2 & 1.9 & - & 1.8 & 2 & 1.9 & 1.2 & 2 \\ 3 & 1.5 & 2 & 2 & - & 1.5 & 1.1 & 1.8 & 1.7 \\ 2.1 & 2.7 & 2.2 & 1 & 1.5 & - & 0.9 & 0.9 & 2.9 \\ 2.8 & 2.3 & 1.5 & 1.8 & 0.9 & 0.8 & - & 1.3 & 2.1 \\ 2.8 & 2.2 & 2 & 1.1 & 1.5 & 0.8 & 1.9 & - & 0.3 \\ 1 & 1.5 & 1.1 & 2.7 & 1.9 & 1.8 & 2.4 & 3 & - \end{pmatrix}.$$

Finally a network representing the city of Seville (Spain) has been defined. The network of Seville is a medium-large size with 24 nodes, 264 edges, and up to 552 demands. A reduced Seville network is shown in Fig. 3.

For the Seville network, as in previous networks, the initial data are defined.

4.1. Experiment comparing integral and heuristic solutions for the network R1

The experiments try to prove the high performance of the model in order to reproduce the multi-period capacity expansion problem, and to compare the integral and the heuristic solutions to study their computational efficiency. The RTNDCE solutions in the computational tests are referred to as “Integral”.

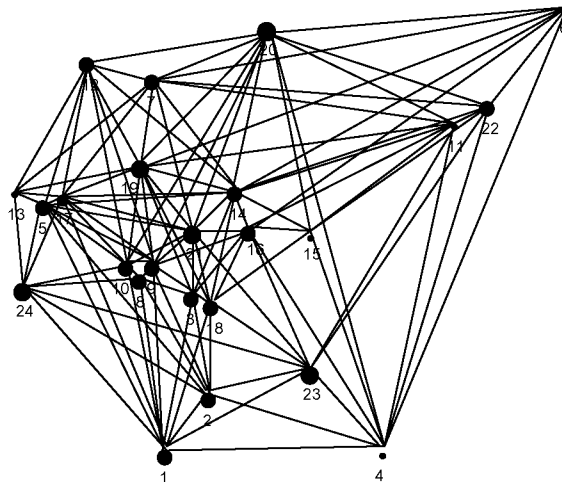


Fig. 3. Seville's test network.

Table 1
Comparing Integral and heuristic solutions, network R1 with 30 demands and two periods

Budget	Phase	PUBdem/dem		LocCost/budget		AdimRouting cost		Objective function		Relative gap		Computational time	
		Integral	Heuristic	Integral	Heuristic	Integral	Heuristic	Integral	Heuristic	Integral	Heuristic	Integral	Heuristic
8	1	0.11	0.11	0.86	0.86	0.98	0.98						
	2	0.44	0.28	0.88	0.93	0.78	0.95	-0.32	-0.17	0	0	335.8	3
10	1	0.19	0.19	0.96	0.96	0.89	0.89						
	2	0.6	0.44	0.96	0.66	0.83	0.78	-0.57	-0.42	0	0	1340.47	5.56
12	1	0.23	0.23	0.92	0.92	0.89	0.89						
	2	0.69	0.69	0.91	0.91	0.73	0.73	-0.69	-0.69	0	0	3275.28	5.36
14	1	0.28	0.28	0.89	0.89	0.96	0.96						
	2	0.9	0.9	1	1	0.7	0.7	-0.93	-0.93	0	0	3573.05	10.1
16	1	0.44	0.44	0.89	0.89	0.79	0.79						
	2	0.95	0.95	0.99	0.99	0.61	0.61	-1.13	-1.13	0	0	663.94	4.72
20	1	0.6	0.6	0.98	0.98	0.85	0.85						
	2	1	1	0.82	0.82	0.57	0.57	-1.35	-1.35	0	0	147.64	25

Table 2
Comparing integral and heuristic solutions, network R1 with 30 demands and three periods

Budget	Phase	PUBdem/dem		LocCost/budget		AdimRouting cost		Objective function		Relative gap		Computational time	
		Integral	Heuristic	Integral	Heuristic	Integral	Heuristic	Integral	Heuristic	Integral	Heuristic	Integral	Heuristic
8	1	0.11	0.11	0.86	0.86	0.98	0.98						
	2	0.44	0.28	0.88	0.93	0.78	0.95						
	3	0.69	0.46	0.88	0.72	0.75	0.83	−0.9	−0.51	0.31	0	10800.11	3.86
10	1	0.19	0.19	0.96	0.96	0.89	0.89						
	2	0.6	0.44	0.96	0.66	0.83	0.78						
	3	0.79	0.69	0.66	0.7	0.78	0.75	−1.24	−0.98	0.17	0	10800.13	6.75
12	1	0.23	0.23	0.92	0.92	0.89	0.89						
	2	0.69	0.69	0.87	0.91	0.77	0.73						
	3	0.95	0.9	0.66	0.55	0.59	0.63	−1.51	−1.47	0.13	0	10800.11	5.75
14	1	0.28	0.28	0.89	0.89	0.96	0.96						
	2	0.9	0.9	1	1	0.7	0.7						
	3	1	1	0.63	0.82	0.54	0.54	−1.8	−1.8	0	0	8987.23	10.33
16	1	0.44	0.44	0.89	0.89	0.79	0.79						
	2	0.95	0.95	0.99	0.99	0.61	0.61						
	3	1	1	0.51	0.51	0.54	0.54	−2.01	−2.01	0	0	3691.58	4.83
20	1	0.6	0.6	0.99	0.99	0.85	0.85						
	2	1	1	0.82	0.82	0.57	0.57						
	3	1	1	0	0	0.54	0.54	−2.24	−2.24	0	0	426.53	25.09

In the experiment of this Subsection, the solutions of the integral and heuristic solutions are compared with two and three periods, using the network R1, with 30 demands and a maximum of four lines. In the problems with two phases the integral and heuristic results have been run up to obtain the optimal solution (relative gap = 0). In the first example for two periods, the results for six different maximum location costs: 8, 10, 12, 14, 16, and 20 are compared in Table 1.

The integral and the heuristic solutions have similar computational results. The heuristic objective function values are close to the integral values and for budget 12 or above the heuristic obtain the optimal solution. The advantage of the heuristic algorithm is that the solution is obtained in less computational time. The computational time is given in seconds for all the figures. Furthermore, the solutions have a correct behaviour: when the budget increases the location cost also increases, the public demand coverage increases, and the routing cost decreases.

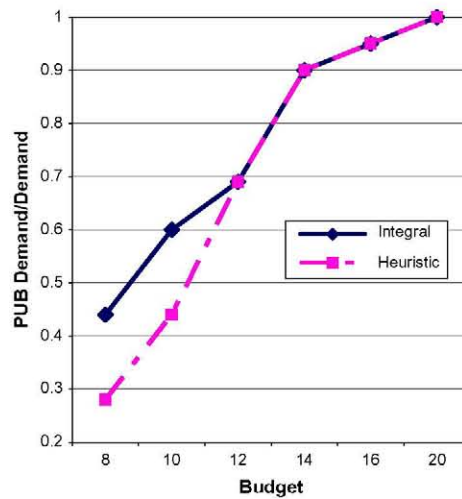


Fig. 4. PUB demand/demand vs budget network R1, 30 demands, two phases.

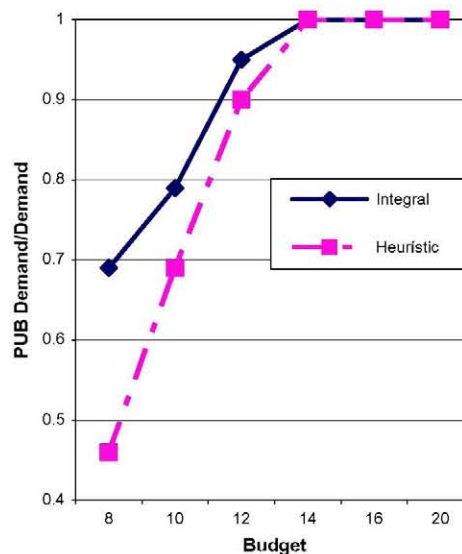


Fig. 5. PUB demand/demand vs budget network R1, 30 demands, three phases.

Table 3
Comparing integral and heuristic in network R2

Budget	Phase	PUBdem/dem		LocCost/budget		AdimRouting cost		Objective function		Relative gap		Computational time	
		Integral	Heuristic	Integral	Heuristic	Integral	Heuristic	Integral	Heuristic	Integral	Heuristic	Integral	Heuristic
8	1	0.06	0.06	0.84	0.84	0.97	0.97						
	2	0.31	0.18	0.99	0.81	0.83	0.93						
	3	0.52	0.32	0.63	0.71	0.73	0.87	−0.57	−0.24	0.78	0	10800	30.28
10	1	0.14	0.14	0.96	0.96	0.9	0.9						
	2	0.29	0.31	0.6	0.64	0.82	0.83						
	3	0.75	0.52	0.99	0.57	0.64	0.74	−0.85	−0.66	0.69	0	10800	52.7
12	1	0.16	0.17	0.93	0.96	0.91	0.93						
	2	0.52	0.48	0.75	0.85	0.74	0.82						
	3	0.99	0.75	0.89	0.9	0.53	0.62	−1.32	−1.05	0.24	0	10800	50.11
14	1	0.17	0.18	0.93	0.87	0.95	0.93						
	2	0.75	0.74	0.94	0.99	0.66	0.66						
	3	0.99	0.99	0.35	0.5	0.53	0.53	−1.55	−1.55	0.18	0	10800	106.88
16	1	0.29	0.31	0.99	0.99	0.83	0.86						
	2	0.99	0.72	0.98	0.84	0.56	0.8						
	3	0.99	0.99	0	0.63	0.53	0.51	−1.92	−1.66	0	0	581.56	65.38
20	1	0.32	0.32	0.87	0.85	0.86	0.88						
	2	0.99	0.91	0.71	0.89	0.56	0.66						
	3	0.99	0.99	0	0.3	0.53	0.53	−1.95	−1.87	0.02	0	10800	3423.72
22	1	0.5	0.5	0.98	0.98	0.83	0.83						
	2	0.99	0.99	0.7	0.76	0.53	0.53						
	3	0.99	0.99	0	0	0.51	0.51	−2.11	−2.11	0	0	2465.7	113.77
25	1	0.51	0.51	0.92	0.92	0.79	0.79						
	2	0.99	0.99	0.36	0.46	0.56	0.56						
	3	0.99	0.99	0	0	0.53	0.53	−2.14	−2.14	0.07	0	10800	447.2
28	1	0.7	0.7	0.96	0.96	0.7	0.7						
	2	0.99	0.99	0.19	0.23	0.56	0.56						
	3	0.99	0.99	0	0	0.53	0.53	−2.33	−2.33	0	0	4980.7	820.42
31	1	0.7	0.7	0.87	0.87	0.7	0.7						
	2	0.99	0.99	0.17	0.21	0.56	0.56						
	3	0.99	0.99	0	0	0.53	0.53	−2.34	−2.33	0.1	0	10800	6032.51
34	1	0.95	0.95	0.95	0.95	0.61	0.61						
	2	0.99	0.99	0	0	0.56	0.56						
	3	0.99	0.99	0	0	0.53	0.53	−2.57	−2.57	0	0	8241.4	239.11
40	1	0.99	0.99	0.93	0.93	0.56	0.56						
	2	0.99	0.99	0	0	0.53	0.53						
	3	0.99	0.99	0	0	0.51	0.51	−2.62	−2.62	0	0	58.78	3.36

In Table 2 the computational results for R1 and 30 demands are run with three periods. In these experiments the time limit to obtain integral solutions was 10800 seconds; meanwhile for heuristic and each of the phases the optimal solution is obtained (gap = 0).

The integral and the heuristic solutions also have similar computational results. The heuristic objective function values are close to the integral values, for budget 14 or above the heuristic obtain the optimal solution. The advantage of the heuristic algorithm is that the solution is obtained in much less computational time, so the heuristic solution represents a good approach to the integral solution.

In the same way as in the example with two periods, the results for three periods have a correct behaviour: when the budget increases the location cost also increases, the public demand covertures increase, and the routing costs decrease. The example with three periods uses more computational time than using two periods, but the use of three periods is closer to the necessity of algorithms to help the decision takers. In the following experiments only the results for three periods will be shown.

Figs. 4 and 5 show the curves PUB demand/demand versus Budget to the Integral and heuristic solutions, for R1, 30 demands and two and three phases, corresponding to the previous Tables 1 and 2.

The sections of the curves with more pendants give an idea of the most beneficial budget to plan the network design, so we obtain practical information about the percentage of public demand increases by budget unit. A criterion may be that this pendant must be superior to 0.1.

In Fig. 4 the sections of the “integral” curves with more pendants are from 12 to 14 of the budget, this section has a pendant of 0.105. If we make the same calculation for the heuristic the sections of more pendants are from 10 to 14 (both with pendant superior to 0.1), but the first section from 10 to 12 is included because the heuristic does not give the optimal solution, so we must take careful consideration of the results obtained using only the heuristic. Normally this optimum interval is bigger in the heuristic than in the integral.

Fig. 5 corresponds to three phases; the section of the “integral” curve with maximum pendant is from 10 to 12 with a pendant of 0.08. In the heuristic curve two sections, from 8 to 12, have a pendant superior to 0.1, but for the same reason (the error of the heuristic solution) we must take into careful consideration the maximum interval obtained by the heuristic.

4.2. Experiment comparing integral and heuristic solutions for the network R2, with 42 demands and four lines

In these experiments the solutions of the integral and the heuristic methods are compared with three periods using the network R2, 42 demands and a maximum of four lines. In this problem the integral and heuristic results have been run up to obtain the optimal solution or with a time limit of 10800 seconds if the relative

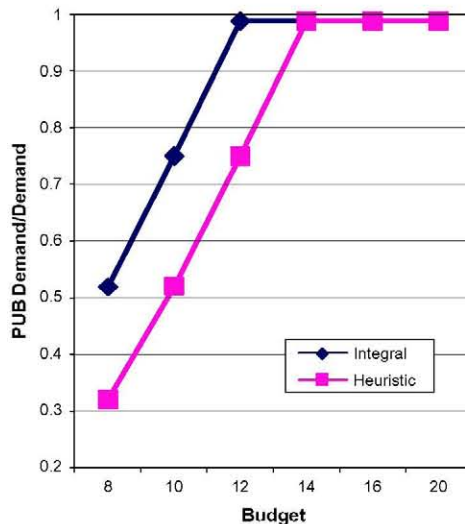


Fig. 6. PUB demand/demand vs budget network R2, 42 demands, three phases.

Table 4
Comparing Integral and heuristic solutions in the Seville network

Budget	Phase	PUB demand/demand		Location cost/budget		AdimRouting cost		Objective function		Relative gap		Computational time	
		Integral	Heuristic	Integral	Heuristic	Integral	Heuristic	Integral	Heuristic	Integral	Heuristic	Integral	Heuristic
2700	1	0.29	0.39	0.86	1.00	1.35	1.62						
	2	0.92	0.85	0.91	0.77	1.05	1.14						
	3	1.00	0.89	0.00	0.00	0.92	1.07	−2.06	−1.98	0.3000	0.5100	108000	69897
3000	1	0.41	0.43	0.99	1.00	1.62	1.61						
	2	0.92	0.89	0.96	0.70	1.01	1.12						
	3	0.94	0.94	0.00	0.00	0.95	1.00	−2.12	−2.12	—	0.3086	108000	47814
3400	1	0.47	0.48	0.96	0.96	1.57	1.56						
	2	0.97	0.90	0.64	0.60	0.98	1.07						
	3	1.00	0.94	0.00	0.00	0.92	1.00	−2.29	−2.17	—	0.0018	108000	106100
3800	1	0.48	0.49	1.00	0.86	1.58	1.55						
	2	0.97	0.90	0.37	0.53	0.98	1.07						
	3	1.00	0.94	0.02	0.00	0.92	1.00	−2.27	−2.19	—	0.0011	108000	20943
4000	1	0.39	0.49	0.96	0.82	1.72	1.55						
	2	0.97	0.90	0.62	0.51	0.99	1.07						
	3	1.00	0.94	0.13	0.00	0.92	1.00	−2.21	−2.19	—	0.0011	108000	21948
6000	1	0.49	0.47	0.99	0.54	1.59	1.58						
	2	0.97	0.90	0.03	0.34	0.98	1.07						
	3	1.00	0.94	0.00	0.00	0.92	1.00	−2.33	−2.17	0.0100	0.0114	22351	2127
10000	1	0.49	0.49	0.61	0.33	1.60	1.55						
	2	0.97	0.90	0.00	0.20	0.98	1.07						
	3	1.00	0.94	0.07	0.00	0.92	1.00	−2.32	−2.20	0.0096	0.0009	8163	2012
12000	1	0.49	0.51	0.46	0.34	1.59	1.56						
	2	0.97	0.94	0.00	0.12	0.98	1.01						
	3	1.00	0.97	0.01	0.00	0.92	0.95	−2.33	−2.28	0.0584	0.0007	280007	2590
15000	1	0.49	0.49	0.35	0.25	1.59	1.63						
	2	0.97	0.94	0.02	0.10	0.98	1.01						
	3	1.00	0.97	0.00	0.00	0.92	0.95	−2.33	−2.26	0.0160	0.0006	280007	1640

gap is not zero. This experiment compares the results for 12 different budget costs between eight and 40. The solutions are represented in Table 3.

As in the previous examples, the computational results have a correct behaviour. The same as in the above experiments, in this experiment for high budgets, the integral and the heuristic algorithms have similar optimal solutions. The advantage of the heuristic algorithm is that the solution is obtained in less computational time. In this example an effect is that when the budget increases the location costs tend to be concentrated in the first phase, so when the budget is above 34 only the first phase is used. For inferior budgets the three phases are used, which is desirable. From this point of view, it is recommended to use budgets 8–14, where the investments are well distributed through the periods.

In Fig. 6, the curves PUB demand/demand versus Budget corresponding to Table 3 show the Integral and Heuristic solutions. The sections of the integral curve with maximum pendant are from 8 to 12 with a pendant above 0.1. In the heuristic curve three sections, from 8 to 14, have a pendant above 0.1, but because of the error in the heuristic solution we must take into careful consideration the maximum interval obtained. From both points of view, we can conclude that the period from 8 to 14 is the most advantageous with regard to the efficient of the resource investment.

4.3. Experiment comparing integral and heuristic solutions for the Seville network, with 72 demands and four lines

Table 4 and Fig. 7 show the results of the heuristic and integral models for the Seville network, considering only 72 of the 552 possible demands, three periods, budget values between 2700 and 15000 monetary units, and using the same parameters defined in R1 and R2. Budgets with values inferior to 2700 give unfeasible solutions.

The heuristic finds solutions within the above budget range, but the relative gaps are lower than 0.1 only for budget values above 3400. Computational times are high (around 30 hour) for low budget values but, for budget values greater than 3800, the heuristic finds a solution in a few hours.

After 30 hour of running time, the integral model only finds feasible solutions (in the limit of 30 computational hours) for budget values above 2700, and these results have high relative gaps. For the budget range and under these conditions, the results of the heuristic are close to the integral model, in particular for budget values higher than 12000.

As Fig. 7 shows, the slope of the heuristic model curve is steepest for budget ranges from 500 to 1500. We cannot determine the steepest part of the integral model's curve because the model does not find valid results for the whole budget range.

From the results above, we can infer that the results of the heuristic model are more practical than those of the integral model: the heuristic model finds solutions for all values in the budget range; in the low-budget range, some of these solutions are sub-optimal, with high relative gaps. The integral model, by contrast, is not able to find solutions for the lowest budget values and on average, its solutions have high relative gaps.

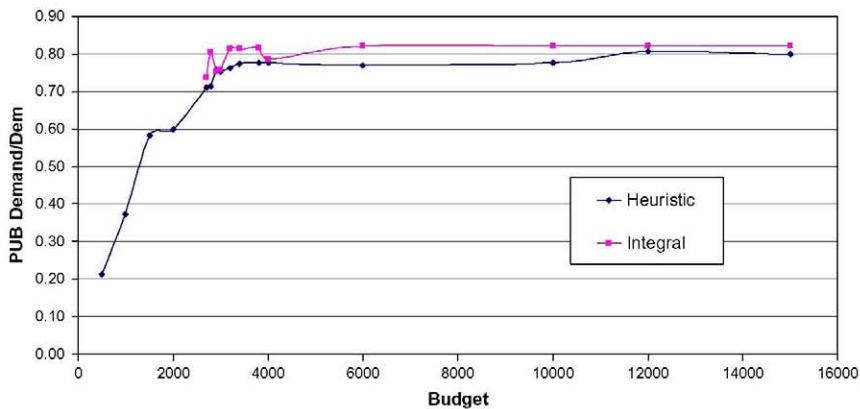


Fig. 7. PUB demand/demand versus budget. Seville network.

When the budget is high, the heuristic model finds good solutions, with relative gaps equal to zero, whereas the integral model finds acceptable solutions, with relative gaps between 0.05 and 0.01.

With this experiment the difficulty in solving medium/large real networks using the integral model is clear. For these real networks only the heuristic may be used. This result is no surprise given that the static RTND (several times inferior in size) cannot obtain solutions with more than nine nodes using branch and bound.

5. Conclusions

A first version of the capacity expansion problem of the rapid transit network design has been defined. Some constraint alternatives and capacity extensions have been discussed. The model finds a long-term expansion plan for the urban rapid network implemented by phases. Planning is based on the time evolution of demand, prices, congestion and available resources. The problem must be analyzed using different phases because the large network construction investments must be optimally distributed in order to consider the changes through time of budget availability, price of the money, demand and traffic. These factors are very important and the expansion of the infrastructure should be taken into account. They must be programmed efficiently and coherently according to the evolution of these variables through time.

The capacity expansion of the rapid transit network design model is a mixed integer programming model. Basically it is a large scale problem due to the high number of variables and constraints, which is difficult to solve. Some first computational tests have been run using branch and bound but the algorithm requires high computational time and only offers solutions with high relative gaps. This requires us to define a heuristic; the heuristic algorithm obtains a feasible coherent local optimum, which represents a good and practical alternative for the solution to the problem as it can reach solutions near to the exact solution in much less time.

Further research may be explored to efficiently solve the model: Firstly, the model can be strengthened in a number of ways. They include the addition of valid inequalities, reduction of coefficients and the incorporation of redundant cut constraints. The goal will be to find a set of constraints which are effective in improving the lower bound value with a reasonable computational effort. Secondly, the model may be solved using meta-heuristic algorithms and/or decomposition methods with the capacity to explore the internal mathematical structures of the model.

References

- Bruno, G., Ghiani, G., Improta, G., 1999. A multi-modal approach to the location of a rapid transit line. *European Journal of Operational Research* 104, 321–332.
- Bruno, G., Gendreau, M., Laporte, G., 2002. A heuristic for the location of a rapid transit line. *Computers and Operations Research* 29, 1–12.
- García, R., Marín, A., 2001. Urban multimodal interchange design methodology. In: Pursula, M., Niittymäki, J. (Eds.), *Mathematical Methods on Optimization in Transportation Systems*. Kluwer Academic Publishers, pp. 49–79.
- García, R., Marín, A., 2002. Parking capacity and pricing in Park'n ride trips: A continuous equilibrium network design problem. *Annals of Operations Research* 116, 153–178.
- Hamacher, H., Liebers, A., Schöbel, A., Wagner, D., Wagner, F., 2001. Locating new stops in a railway network. *Electronic Notes in Theoretical Computer Science*, 50.
- Laporte, G., Marín, A., Mesa, J.A., Ortega, F.A., in press. An integrated methodology for the rapid transit network design problem. *Lectures Notes in Computer Science. Algorithmic Methods for Railway Optimization*, 4359.
- Laporte, G., Mesa, J.A., Ortega, F.A., 2002. Locating stations on rapid transit lines. *Computers & Operations Research* 29, 741–759.
- Laporte, G., Mesa, J.A., Ortega, F.A., 2005. Maximizing trip coverage in the location of a single rapid transit alignment. *Annals of Operations Research* 136 (1), 49–63.
- Marín, A., in press. An extension to Rapid Transit Network Design Problem. *TOP*, Springer Verlag.