

# Planning Rapid Transit Networks

G. Laporte<sup>a,\*</sup>, J.A. Mesa<sup>b</sup>, F.A. Ortega<sup>c</sup>, F. Perea<sup>d</sup>

<sup>a</sup>Canada Research Chair in Distribution Management, HEC Montréal, 3000, chemin de la Côte-Sainte-Catherine, Montréal, Canada H3T 2A7

<sup>b</sup>Departamento de Matemática Aplicada II, Escuela Técnica Superior de Ingeniería, Universidad de Sevilla, Camino de los Descubrimientos s/n, 41092 Sevilla, Spain

<sup>c</sup>Departamento de Matemática Aplicada I, Escuela Técnica Superior de Arquitectura, Universidad de Sevilla, Avenida Reina Mercedes s/n, 41012 Sevilla, Spain

<sup>d</sup>Departamento de Estadística e Investigación Operativa Aplicadas y Calidad, Universidad Politécnica de Valencia, Camino de Vera s/n, 46021 Valencia, Spain

---

## Abstract

Rapid transit construction projects are major endeavours that require long-term planning by several players, including politicians, urban planners, engineers, management consultants, and citizen groups. Traditionally, operations research methods have not played a major role at the planning level but several tools developed in recent years can assist the decision process and help produce tentative network designs that can be submitted to the planners for further evaluation. This article reviews some indices for the quality of a rapid transit network, as well as mathematical models and heuristics that can be used to design networks.

*Keywords:* planning, rapid transit network design, mathematical programming, heuristics.

---

## 1. Introduction

In the area of passenger transportation, there has been a tendency in recent years to increase investments in public transit projects and to reduce them in road construction. This shift comes partly as a response to environmental concerns and as a realization that dependence on fossil fuels for transportation is not a sustainable option [1]. Many cities have constructed new metro systems or have expanded or upgraded old ones. According to Wikipedia [2] there are now approximately 140 metro systems in the world, as opposed to 90 fifteen years ago [3]. There is no precise definition of a metro system but these are always independent of other road or pedestrian traffic. They are therefore designed with physical separation [4]. Metros are often underground but in many cities, like in London, the suburban part of the network is overground and sometimes shares some infrastructure with the main railway network. In addition, fully overground light-rail systems are quite common, and there exist other systems such as light metros, pre-metros, commuter trains, light-profile rapid transit

systems, monorails, etc. For convenience we will refer to all these systems as *rapid transit networks*. Since bus routes interact with street traffic and main railway lines are not normally part of the urban or metropolitan setting, we will exclude these from our discussion.

Rapid transit construction projects are major endeavours that require long-term planning. These projects are very costly, fraught with uncertainties, and subject to schedule or budget overruns. The number of stakeholders involved in such ventures is important as politicians, urban planners, engineers, management consultants, and citizen groups take part in the decision process. These players often have conflicting objectives and constraints, which are hard to formally define in simple terms. This problem is exacerbated by the fact that rapid transit projects have strong externalities ranging from their impact on the urban fabric, to changes in traffic patterns and reductions in air pollution. In such a context, it is unrealistic to contemplate solving the problem through a direct application of standard mathematical models and optimization algorithms. In fact, Gendreau et al. [3], who examined some 40 rapid transit projects, came to the conclusion that planners used little or no operations research. Vuchic [6] also notes that most of the academically oriented literature on systems analysis and operations research fails to reach actual ap-

---

\*Corresponding author

Email addresses: gilbert.laporte@cirreil.ca (G. Laporte), jmesa@us.es (J.A. Mesa), riejos@us.es (F.A. Ortega), perea@us.es (F. Perea)

plications, partly because it is too theoretical or pursues incorrectly formulated objectives. However, this author sees a need and a clear potential for the use of operations research methods in transit planning. We also share this view in so far as large segments of the planning process are rather well defined and lend themselves to the use of analytical techniques. Given the multi-player and multi-objective nature of the problem, we believe that the role of systems optimization techniques is mostly to construct and assess potential solutions to be later submitted to the decision makers.

In what follows, we will describe some operations research methods, developed over the past years, which we believe can assist the planning process. These methods can be used to assess network configurations, to generate rough networks and to locate stations. For an earlier review covering some of the pre-2000 research, see [7].

The remainder of this paper is organized as follows. Section 2 is devoted to the assessment of rapid transit networks. Mathematical models for the design of a network are described in Section 3, and heuristics for the location of simple alignments, multiple alignments and stations are presented in Section 4. Conclusions follow in Section 5.

## 2. Assessing rapid transit networks

The main objective of a rapid transit system is to improve population's mobility [8, 9]. A good network should be designed so as to provide short travel times to many people, while respecting some technical and budgetary constraints. This translates into providing reasonably direct service for a large number of trips. The major destinations covered by most networks are the main work and shopping areas of a city, typically the downtown core, transportation hubs such as bus terminals, train stations and airports, tourist attractions, entertainment sites, universities, hospitals, etc. The origins tend to be highly populated areas. With this in mind, planners will create alignments covering the main travel corridors while providing a sufficient degree of connectivity between the lines of the network.

A number of indices can be used to assess rapid transit configurations. The total line length and the number of stations are obvious ways of measuring the geographical extent of a network. Thus, if one excludes the Los Teques metro in Venezuela which has currently two stations but is still under construction, the smallest metros have only six stations. These are located in Catania, Kazan, Lima, Maracaibo and Diepropetrovs'k [5]. At the other extreme, the largest metro systems have

hundreds of stations (Table 1). Population coverage is also an important measure. According to Vuchic [6], it can be considered that most potential users living within five minutes walk (400 m) of a station will use the system, and the ridership falls practically to zero when the user-station walking distance increases to 10 minutes. In Paris, which has one of the densest networks in the world, most people can access a station within 10 minutes. This is not the case in London which is more spread out than Paris and contains large areas that are poorly served by the Underground. Note that the catchment area of a station is not always exclusively defined by pedestrian traffic. For example, stations with park-and-ride facilities also capture passengers arriving by car or by bus [10].

Table 1: The five metro systems with the largest number of stations (Wikipedia [2])

City	Name	Date opened	Number of stations	Length (km)
New York City	New York City Subway	1870	468	369
	PATH	1908	13	22
Paris	Métro de Paris	1900	368	214
London	London Underground	1863	270	408
	Docklands Light Railways	1987	40	34
Madrid	Metro de Madrid	1919	294	284
Seoul	Seoul Subway	1974	293	317

Some measures relate to the topology of a network  $G = (N, E)$ , where  $N$  is the node set and  $E$  is the edge set. The first quality indices, put forward by Musso and Vuchic [11], include the number of stations, the total length of the network, the number of lines and the number of multiple stations. Five more elaborate measures defined by the same authors are:

$C$ : the number of minimal cycles (not embedding any other cycles):

$$C = |E| - |N| + 1 ;$$

$\alpha$ : a cycle availability index, defined as the ratio between  $C$  and the largest value it could take for a network with  $|N|$  nodes and  $|E|$  edges:

$$\alpha = (|E| - |N| + 1) / (2|N| - 5) ;$$

$\beta$ : a measure of the network complexity:

$$\beta = |E| / |N| ;$$

$\gamma$ : a connectivity indicator equal to the ratio of  $|E|$  to the maximal number of edges that could exist in a planar network with  $|N|$  nodes:

$$\gamma = |E| / 3(|N| - 2) ;$$

$\delta$ : a measure of directness of service equal to the number of origin/destination ( $O/D$ ) paths that can be traveled without transfers.

In addition to these measures, Laporte et al. [12] have also defined the passenger/network effectiveness index of a network and the passenger/plane effectiveness index. To define the passenger effectiveness index  $\lambda$ , first compute the total length of the network (in terms of travel times):

$$T = \sum_{(i,j) \in E} t_{ij},$$

where  $t_{ij}$  is the travel time on edge  $(i, j)$ . Then define the total passenger cost as the sum, over all  $O/D$  pairs  $(i, j)$ , of passenger travel times  $\theta_{ij}$  using shortest paths  $P$ . The value of  $\theta_{ij}$  includes the travel time on the edges of  $P$  and the transfer time. The travel time  $\theta_{ij}(P)$  on path  $P$  is equal to

$$\theta_{ij}(P) = \sum_{(i,j) \in P} t_{ij} + r(P)t_f + (s(P) - r(P) - 1)t_s,$$

where  $r(P)$  and  $s(P)$  are, respectively, the number of transfers and edges on  $P$ , and  $t_f$  and  $t_s$  are the corresponding transfer and stopping times. The value of  $\theta_{ij}$  is computed as

$$\theta_{ij} = \min_P \theta_{ij}(P),$$

and

$$\lambda = \sum_{\substack{i,j \in N \\ i < j}} \theta_{ij} / T.$$

The passenger/plane effectiveness index  $\mu_p$  compares passenger travel time on the network to what it would be if travel was made on the street network, assuming that an  $\ell_p$  norm is used ( $p \in [1, 2]$ ). To compute  $\mu_p$ , first define the matrix  $\Theta = (\theta_{ij})$  and the matrix  $M^p = (m_{ij}^p)$ , where  $m_{ij}^p$  is the travel time on the street network according to metric  $\ell_p$ . Then

$$\mu_p = \|\Theta - M^p\| / |N|,$$

where  $\|\cdot\|$  is the Frobenius norm, *i.e.*, given a matrix  $A = (a_{ij})$ ,  $\|A\| = (\sum_i \sum_j a_{ij}^2)^{1/2}$ . Note that lower values of  $\lambda$  and  $\mu_p$  are preferred.

To illustrate the use of these norms, consider the three simple networks depicted in Figure 1 and the corresponding statistics reported in Table 2 with  $t_{ij} = 1$  for all  $(i, j) \in E$ ,  $t_f = 0.3$  and  $t_s = 0.1$ . The results presented in Table 2 indicate that the two effectiveness measures tend to be inversely related to the number of cycles in the network, complexity, connectivity and directness.

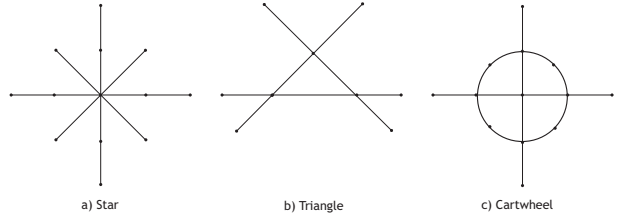


Figure 1: Three simple networks

The worst network topology is the star, whereas the triangle and the cartwheel are rather effective configurations. These measurements are purely topological and make no assumption about passenger volumes and modal competition. A follow up study by the same authors [13] has shown that the conclusions one can derive from Table 2 still hold when traffic volumes and modal competition are considered. It should also be noted that the quality of a network should not be assessed on purely topological measurements, but also on service frequency and timetabling synchronization between the lines at transfer nodes. Derrible and Kennedy [15] consider such nodes as a relevant characteristic of complex networks.

Table 2: Indices for three simple networks

	Star	Triangle	Cartwheel
$ N $ : number of stations	13	9	13
Total length	12	9	14.2832
Number of lines	4	3	3
Number of multiple stations	1	3	5
$C$ : number of cycles	0	1	4
$\alpha$ : cycle availability	0	0.0769	0.1902
$\beta$ : complexity	0.9230	1	1.2307
$\gamma$ : connectivity	0.3635	0.4285	0.4848
$\delta$ : directness	0.3333	0.5	0.5897
$\lambda$ : passenger/network effectiveness	34.1667	18.4667	25.9055
$\mu_{1.5}$ : passenger/plane effectiveness	0.8770	0.5948	0.5341

The three simple configurations depicted in Figure 1 and several others are encountered in many cities. An interesting reading on this topic is the book by Oveden [14] which contains maps of the world's major metro systems. To illustrate, the Minsk metro is now a simple star consisting of two lines (Figure 2a) but the plan is to add a third line to form a triangle (Figure 2b). The Prague Metro (Figure 3) is a perfect example of a triangle configuration. The London Underground (Figure 4) is a highly complex system with several intersecting branches, particularly within the Circle Line, and creating triangles. This network combines the advan-

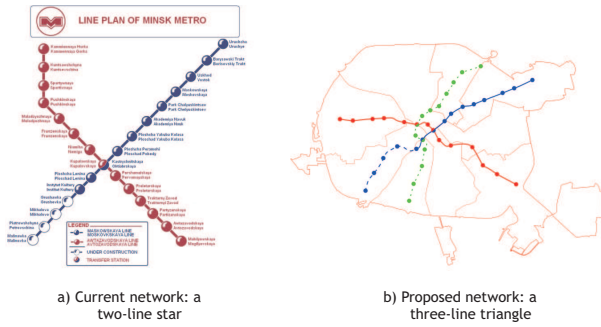


Figure 2: The Minsk Metro (Wikipedia [16])

tages of the triangle and cartwheel configurations. Similar topologies are to be found in Moscow, Shanghai and Paris, for example.

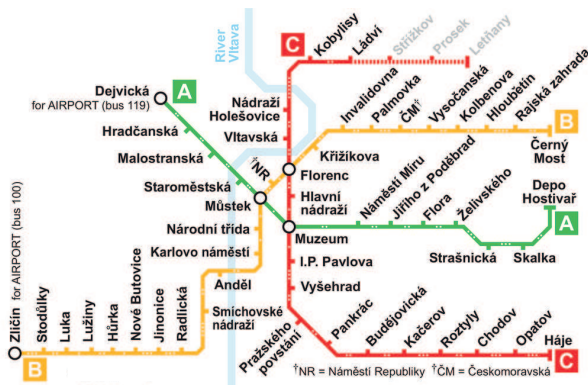


Figure 3: The Prague Metro (Mappery [17])

The above indices are computed under the assumption that the lines forming the transit network operate independently and have no common segments. It is, however, possible to improve the directness and efficiency of a system by allowing line sharing. For example, the star network of Figure 5 a can be operated as two independent (diagonal) lines (a) or as six lines (b), among other possibilities. Of course, the latter mode of operation is more difficult to manage since it involves multiple routing and scheduling decisions. Parts of the London Underground (e.g., the Northern Line, Figure 6) operate in this fashion. For a good exposition of line planning, the reader is referred to Chapter 4 of [6]. A very good reference on scheduling is the paper by Liebchen [19].

Another line of research in the area of transit network evaluation relates to the concept of robustness. Robustness in network design can be defined in several ways. Most references are devoted to the construction of low cost communication networks that can survive in the presence of edge failures [21]. In recent years, several

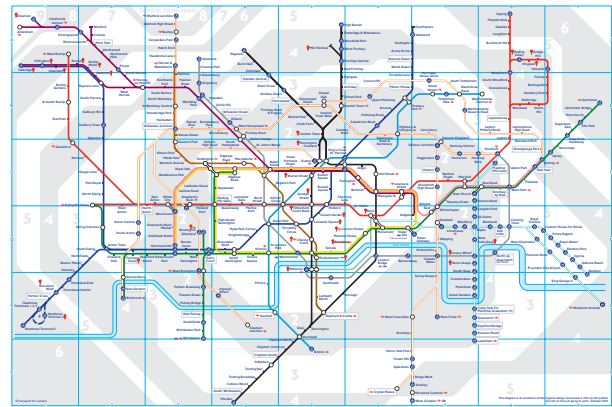


Figure 4: The London Underground and Docklands Light Railway (Transport of London [18])

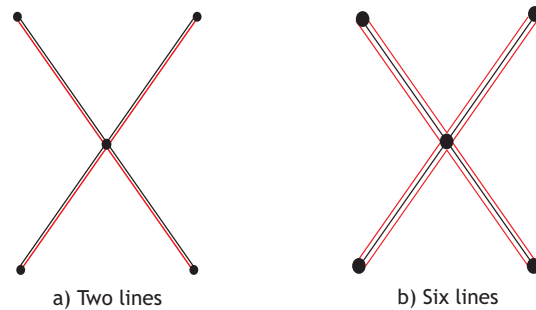


Figure 5: Star configuration operated as two or six lines

authors have assessed the topological configurations of transportation infrastructure by means of graph-related indices, like connectivity [11, 12], efficiency [13], performance [22], vulnerability [23] and survivability [24]. In this paper, a network is said to be robust when it reacts well to disruptions of links or stations. Intuitively, a network, such as a cartwheel, which offers alternative itineraries, should be more robust than a star which provides no protection in case of failure. Albert and Barabási [25] and Klau and Weiskircher [26] have developed robustness indices for topological networks, i.e., networks in which all edges have the same length. Nagurney and Qiang [27, 28] have measured the robustness of transportation networks with respect to congestion and user behaviour. Recently, De-Los-Santos et al. [29] have proposed robustness indices for rapid transit systems for two scenarios. Under the first scenario, passengers affected by an interruption have to wait for the failure to be repaired in order to pursue their journey. Under the second scenario, a bus service is provided to carry the affected passenger to the end of the failing link. The authors show how a new link con-

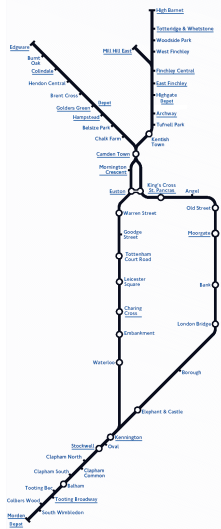


Figure 6: The Northern Line, London Underground (TrainWeb [20])

necting the stations Atocha and Nuevos Ministerios in Madrid has helped improve the robustness of the network under both scenarios.



Figure 7: The Boston Subway (W3 [30])

In recent years, some researchers [31, 32] have applied the concept of small-world networks to the assessment of metro systems. Small-world networks arise in the study of some social, biological and communication networks. They are characterized by low average inter-node distances and high average local connectivity. In general networks, these properties have been substituted by global and local efficiency, respectively. This means that in a small-world transportation network it is easy to travel both at the global and local levels. Using these concepts, the Boston Subway (Figure 7) was shown to

have a relatively good global efficiency but a low local efficiency, thus not representing an example of a small-world network. Since network robustness can be interpreted as a loss of efficiency, this low local efficiency means that the Boston Subway is not tolerant against disruptions. However, an opposite conclusion can be reached by incorporating the complementary bus network in the analysis [31], or by providing temporary bus-bridging operations.

### 3. Mathematical models for the design of rapid transit networks

The design of a rapid transit network is a highly complex process. Operations research techniques borrowed from the fields of network design and metaheuristics can support the planning process by providing candidate solutions meeting some coverage and cost criteria.

#### 3.1. Network location models

Formally, the metro network design problem is that of embedding a set of interconnected transit lines within a larger undirected network  $G = (N, E)$ , where  $N = \{1, \dots, n\}$  is a node set and  $E = \{(i, j) : i, j \in N, i < j\}$  is an edge set. The nodes correspond to population centroids in a city, while the edges correspond to potential connections to be built between vertex pairs. Let  $p_i$  be the population associated with node  $i$ . The vertex population is normally defined as the population living within a reasonable walking distance of the vertex, e.g., 500 m. Let  $c_{ij}$  be the cost of constructing a link between  $i$  and  $j$  and let  $c_i$  be the cost of constructing a station at node  $i$ . Three main criteria are used in mathematical models for metro network design: 1) the total construction cost; 2) the total population covered by the network; 3) the total  $O/D$  traffic captured by the network. The problem can naturally be cast within the framework of multi-criteria optimization. Alternatively, one can minimize cost subject to a population or traffic coverage constraint or maximize coverage subject to a budget constraint.

With few exceptions, such as [41] and [47], most available mathematical programming models can be used to locate a network but do not decompose it into distinct transit lines. Given a transit network, a set covering problem can later be solved to locate the lines but there will likely be several overlaps between them. Many networks have no overlapping lines (e.g. Minsk and Prague) while others, like the London Underground, contain several.

Network location models with a minimum cost or maximal coverage objective belong to the class of

*Steiner tree problems with profits (STPP)* [33]. Some classical examples are the following:

1. In the *Prize-collecting STPP*, the objective is a linear combination of the construction cost and of the population covered by the network. Let  $x_{ij}$  be a binary variable equal to 1 if and only if edge  $(i, j)$  belongs to the network, let  $y_i$  be a binary variable equal to 1 if and only if vertex  $i$  belongs to the network, and let  $\alpha$  be a user-controlled positive parameter. Let  $T \subseteq N$  be a set of nodes that must necessarily belong to the network ( $T$  may be empty). The model is then

$$\text{(PCSTPP)} \quad \text{minimize} \quad \sum_{(i,j) \in E} c_{ij}x_{ij} + \sum_{i \in N} c_i y_i - \alpha \sum_{i \in N} p_i y_i \quad (1)$$

subject to

$$\sum_{(i,j) \in E} x_{ij} = \sum_{i \in N} y_i - 1 \quad (2)$$

$$\sum_{i,j \in S} x_{ij} \leq \sum_{i \in S \setminus \{k\}} y_i \quad k \in S \subseteq N, |S| \geq 2 \quad (3)$$

$$y_i = 1 \quad i \in T \quad (4)$$

$$x_{ij} = 0 \text{ or } 1 \quad (i, j) \in E \quad (5)$$

$$y_i = 0 \text{ or } 1 \quad i \in N. \quad (6)$$

The objective function minimizes the network construction cost, minus a multiple of the population covered by the network. Constraints (2) and (3) force the network to be a tree. This model with  $T = \emptyset$  was used by Lucena and Resende [35]. Margot et al. [36] have proved that the  $x_{ij}$  variables need not be declared as integer.

2. In the *Quota STPP* [33], the objective is to minimize cost while ensuring a minimal population coverage  $\beta$ :

$$\text{(QSTPP)} \quad \text{minimize} \quad \sum_{(i,j) \in E} c_{ij}x_{ij} + \sum_{i \in N} c_i y_i \quad (7)$$

subject to

$$(2) - (6)$$

$$\text{and} \quad \sum_{i \in N} p_i y_i \geq \beta. \quad (8)$$

3. In the *Fractional STPP* [34], the objective is maximize to a population-to-cost-ratio. The model is:

$$\text{(FSTPP)} \quad \text{maximize} \quad \alpha \sum_{i \in N} p_i y_i / \left( \sum_{(i,j) \in E} c_{ij}x_{ij} + \sum_{i \in N} c_i y_i \right) \quad (9)$$

subject to

$$(2) - (6).$$

Similar models have been studied in relation with the *Traveling Salesman Problem with Profits (TSPP)* by a number of authors (see [37] for a survey). In this case, a specific vertex, called the depot, is always part of the solution (*i.e.*,  $|T| = 1$ ) and the solution is always a cycle spanning some of the nodes of  $G$ .

Another class of models can be derived by explicitly considering the population not covered by the network. Such models make use of variables  $z_{ij}$  equal to 1 if and only if node  $i$  not on the network is assigned to node  $j$  on the network, *i.e.*,  $j$  is the node on the network closest to  $i$ . Let  $d_{ij}$  be the cost of reaching  $j$  from  $i$  (*e.g.*,  $d_{ij}$  is the walking distance). Then any of the above models can be modified by adding a term  $\gamma \sum_{(i,j) \in E} d_{ij} z_{ij}$  to the objective function, where  $\gamma$  is a positive user-controlled parameter, or by imposing an upper bound on  $\sum_{(i,j) \in E} d_{ij} z_{ij}$ . In such models, technical constraints must be imposed to ensure that a node cannot be at the same time on the network and outside it. Examples are provided in [38, 39, 40].

### 3.2. Line location models

We now introduce a new model to locate a set  $L$  of lines covering a part of  $G$ . As in the *Prize-collecting STPP*, the function to be minimized is a linear combination of the construction cost and of the population covered by the transit network. This is again achieved by subtracting from the construction cost the population covered by the network multiplied by a parameter  $\alpha$ . The model combines the PCSTPP formulation described in Section 3.1 and a simplified version of the model presented in [41]. We first define the following binary variables:  $x_{ij}^l = 1$  if and only if edge  $(i, j)$  belongs to line  $l \in L$ ,  $x_{ij} = 1$  if and only if edge  $(i, j) \in E$  belongs to a line, and  $y_i^l = 1$  if and only if a station is built at node  $i$  on line  $l$ . Let  $z_k^l$  be the number of edges incident to a node  $k \in N$  belonging to line  $l \in L$ , excluding the edges of line  $l$ . The construction costs  $c_{ij}$  have the same meaning as above, and  $c_i^l$  is the cost of constructing a station at node  $i$  on line  $l$ . The prize collecting line location model is then

$$\text{(PCLL)} \quad \text{minimize} \quad \sum_{(i,j) \in E} c_{ij}x_{ij} + \sum_{l \in L} \sum_{i \in N} c_i^l y_i^l - \alpha \sum_{i \in N} p_i y_i \quad (10)$$

subject to

$$x_{ij}^l \leq y_i^l \quad (i, j) \in E, l \in L \quad (11)$$

$$x_{ij}^l \leq y_j^l \quad (i, j) \in E, l \in L \quad (12)$$

$$x_{ij}^l \leq x_{ij} \quad (i, j) \in E, l \in L \quad (13)$$

$$\sum_{i < k} x_{ik}^l + \sum_{j > k} x_{kj}^l \leq 2 \quad k \in N, l \in L \quad (14)$$

$$\sum_{i,j \in S} x_{ij}^l \leq |S| - 1 \quad S \subseteq N, |S| \geq 2, l \in L \quad (15)$$

$$\sum_{\substack{i,j \in N \\ i < j}} x_{ij}^l \geq \sum_{i \in N} y_i^l - 1 \quad l \in L \quad (16)$$

$$z_k^l \geq \sum_{h \in L \setminus \{l\}} \left( \sum_{i < k} x_{ik}^h + \sum_{j > k} x_{kj}^h \right) - M(1 - y_k^l) \quad k \in N, l \in L \quad (17)$$

$$z_k^l \leq \sum_{h \in L \setminus \{l\}} \left( \sum_{i < k} x_{ik}^h + \sum_{j > k} x_{kj}^h \right) + M(1 - y_k^l) \quad k \in N, l \in L \quad (18)$$

$$z_k^l \leq M y_k^l \quad k \in N, l \in L \quad (19)$$

$$\sum_{k \in N} z_k^l \geq 1 \quad l \in L \quad (20)$$

$$x_{ij}^l, x_{ij}, y_i^l = 0 \text{ or } 1 \quad (i, j) \in E, i \in N, l \in L \quad (21)$$

$$z_k^l \geq 0. \quad k \in N, l \in L \quad (22)$$

In the objective function,  $\alpha$  controls the fraction of the population covered by the network. If  $\alpha = 0$ , then no network is constructed. If  $\alpha$  is very large, then it is optimal to cover the entire population. The cost of constructing a station at node  $i$  is counted for each line because the cost of transfer stations depends on the number of lines that connect through them. Constraints (11) and (12) state that edge  $(i, j)$  cannot be part of the network if no station is built at  $i$  and  $j$ . Constraints (13) imply that line  $l$  can use edge  $(i, j)$  only if  $x_{ij} = 1$  and hence the corresponding construction cost is incurred. Constraints (14) force the degree of each node to be 0, 1 or 2. Constraints (14) and (15) taken together mean that all lines will be a collection of paths. Escudero and Muñoz [42] relax these constraints and thus allow the creation of circular lines but they do not control their location nor their number of stations. Constraints (14–16) ensure that each line is made up of a single path. Constraints (17–20) prevent the formation of lines disconnected from the others. In constraints (17) and (18),  $M$  is a very large number. Constraints (17–19) define  $z_k^l$ . Constraints (20) force each line to intersect with at least another one. Finally, constraints (21) and (22) impose conditions on the variables. Marín [43] and Marín and García-Ródenas [44] present similar models without constraints (16–19), *i.e.*, the network may have disjoint lines. Our model prevents the construction of circular lines. It does, however, allow several lines to share the same edge. Also, it does not force all lines to be built. If a line  $l$  is empty, then the associated constraints are redundant.

As in Section 3.1, modifications to this basic model can easily be envisaged to deal with different objectives. Note that the model does not optimize the directness of traffic between  $O/D$  pairs, but only ensures that the network is connected. Setting up a concise model that optimizes directness of traffic appears to be rather challenging. Schöbel and Sholl [45] and Borndörfer et al. [46]

present several models for the problem of selecting a set of lines from a pool to connect several  $O/D$  pairs under a budget constraint. The objective is to minimize the overall travel time made up of the ride time and the transfer time. The size of the model grows very quickly with the number of  $O/D$  pairs. The model described in [44] includes so-called service constraints forcing the network to serve the demand defined by  $O/D$  pairs.

Recently, Laporte et al. [47] have proposed a model for the design of a robust rapid transit network, *i.e.*, a network in which the effect of disruption on total trip coverage is minimized. The model yields a network that provides several alternative routes for some  $O/D$  pairs in case of an interruption, thereby increasing the network robustness. Laporte et al. [41] have also proposed a game theoretic framework for the problem of designing an uncapacitated railway transit network in the presence of link failures and a competing mode. The problem was posed as a non-cooperative two-player zero-sum game with perfect information. The saddle points of the associated mixed enlarged game yield robust network designs.

#### 4. Heuristics for the design of rapid transit networks

Mathematical programming models such as PCLL and the two formulations provided in [41, 47] are generally intractable for all but small size instances. Further, they can only capture some aspects of the problem because not every problem specification lends itself to a characterization through linear constraints. There is therefore a need to develop heuristics.

##### 4.1. Locating a single alignment

A heuristic was proposed in [48] for the problem of constructing an alignment (*i.e.* a single line) of maximal population coverage, subject to a fixed number of stations and constraints on the minimal and maximal inter-station spacing (typically between 500 and 2000 m).

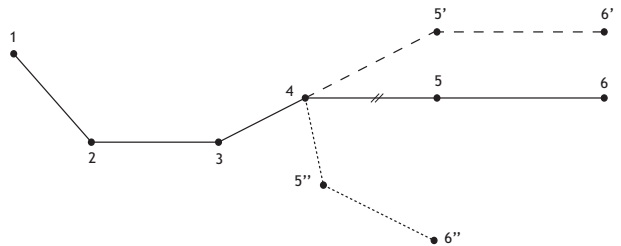


Figure 8: Locating a single alignment

Starting with a single edge, the heuristic iteratively extends an alignment in a greedy fashion (*i.e.* by maximizing the population coverage), while respecting the inter-station spacing constraints. In a second phase, the neighbourhood of the current solution is explored by tabu search. The neighbour of a solution is obtained by cutting an edge of the alignment and reconstructing several partial alignments from the break point. In Figure 8, edge (4, 5) is cut and the network is reconstructed in two different ways: (4, 5', 6') and (4, 5'', 6''). This heuristic was successfully applied to the city of Milan. Whereas the optimal solution is not known for this instance, it can be seen, by visual inspection, that the alignment covers the main population centres of the city (Figure 9).

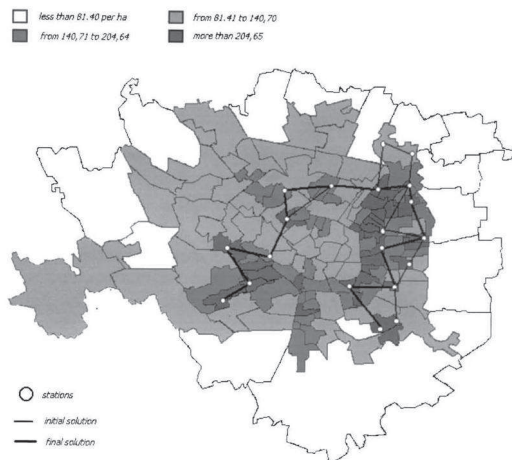


Figure 9: Single alignment covering the main population centres of Milan

#### 4.2. Locating several alignments

This heuristic was later extended in [49] to the multi-alignment case. To initialize the algorithm, the user specifies the shape of the network to be built (*e.g.*, a cartwheel), a corridor within which each line is to be located, the number of stations of each line and the inter-station spacing constraints. Figure 10 depicts three-line solution for the city of Milan in which each line is embedded within a predefined corridor.

The idea of specifying a priori corridors for the lines is interesting from a practical point of view because the planners' knowledge about a city's main traffic corridors can easily be incorporated within the formal solution process. In a mathematical model such as PCLL, it is straightforward to limit each line to a subset of nodes.

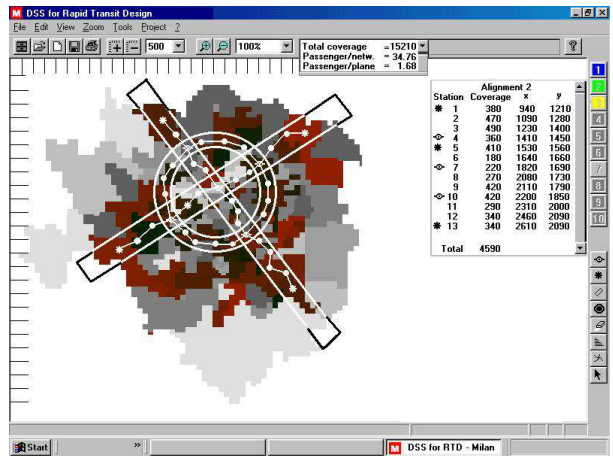


Figure 10: Cartwheel configuration for the city of Milan

In addition, the formulation can then easily be modified to allow the creation of some circular lines within predefined rings. This type of solution process therefore combines the benefits of human intervention and of mathematical programming.

A heuristic for the construction of a single alignment maximizing trip coverage (*O/D* pairs as opposed to population coverage) was developed in [50]. This heuristic uses the *O/D* matrix as an input, as well as the number of stations to be constructed and the interstation spacing constraints. Several constructive heuristics and an improvement procedure were developed and compared. The best results were obtained by the application of a simple greedy extension heuristic. This conclusion was confirmed on scenarios derived from data obtained in the city of Seville where the inter-station distance exceeds 1250 m. When smaller upper bounds are used (700 m and 1000 m), an insertion method followed by a post-optimization phase yields the best results.

The three heuristics developed in [48, 49, 50] are easy to implement and yield good solutions within very short computing times. Their usefulness lies in their ability to quickly generate promising scenarios that can be submitted to the decision makers for further analysis. These heuristics are highly flexible in the sense that they can easily accommodate a host of objectives and side constraints, and their computational behavior is not very much affected by non-linearities.

On an existing public transportation network, the integration of several planning steps, such as the determination of lines and their associated frequencies is addressed in Michaelis and Schöbel [51]. Finally Marín and Jaramillo have applied Benders decomposition algorithms to the rapid transit network design prob-



lem [52]. In another paper [53] they have formulated a multi-objective multi-period rapid transit network design in which existing lines can be extended and new lines can be added at each period. Because the problem is of large scale, it is solved as a sequence of single period problems taking as input the network defined at the previous period.

### 4.3. Locating stations on an alignment

The models introduced in Section 3 and the heuristics outlined in Sections 4.1 and 4.2 can be used to simultaneously locate alignments and stations. In the mathematical models it is assumed that short edges whose length is less than the acceptable inter-station spacing can be eliminated. In a post-processing step, additional stations can sometimes be located on long edges. The problem of locating stations along a line is typically formulated through the use of coverage models [54, 55] or as a bicriteria problem in which the two objectives are coverage and cost [56].

In a classical article on this topic, Vuchic and Newell [57] have considered a star system in which flows converge to the central point. Taking into account passenger distribution, access speed, dynamic characteristics of the trains, waiting time and intermodal transfer time, the authors used simultaneous difference equations to show that station spacing increases in the direction of passenger cumulation. Vuchic [6] compares the effect of locating stations at which all trains must stop to scenarios in which stations are served by alternate trains. Triangulation methods can also be applied.

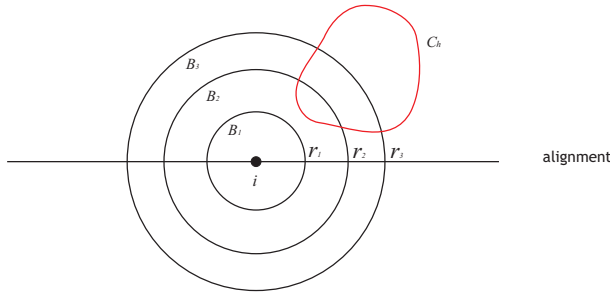


Figure 11: Catchment area of station  $i$  intersecting with census tract  $C_h$

Thus [58] combines the geometrical catchment area of a potential station (computed with respect to the existing street metric) with census data in order to compute an estimate of the number of passengers who will use a station. It is assumed that the catchment area of a potential station is made up of  $K$  concentric rings  $B_1, \dots, B_K$  with attraction decreasing with distance to the centre

(see Figure 11). Population counts are given by  $H$  census tracts  $C_1, \dots, C_H$ . Then the coverage provided by station  $i$  is

$$R(i) = \sum_{h=1}^H \sum_{k=1}^K \frac{a\rho_h}{\bar{r}_k^2} \text{Area}((B_k \setminus B_{k-1}) \cap C_h) \quad i = 1, \dots, n, \quad (23)$$

where  $n$  is the number of stations to be located,  $B_0 = \emptyset$ ,  $r_k$  is the radius of ring  $B_k$  centered at  $i$ ,  $\bar{r}_k$  is an intermediate value between  $r_{k-1}$  and  $r_k$  (e.g.,  $\bar{r}_k = (r_{k-1} + r_k)/2$ ),  $\rho_h$  is the population density of  $C_h$  and  $a$  is a parameter to be determined.

Laporte et al. [58] have combined (23) with an estimation of  $O/D$  patterns with population counts to yield a coverage measure. More precisely, each station pair  $(i, j)$  has an associated  $K \times K$  matrix called  $OD_{ij}(k, k')$ . Each element  $OD_{ij}(k, k')$  is the weighted sum of portions of  $t_{lm}$  values (predicted number of trips of zone  $Z_l$  attracted to zone  $Z_m$ ), where  $l, m = 1, \dots, L$ , and  $L$  is the number of zones over which the  $O/D$  pairs are computed:

$$OD_{ij}(k, k') = \sum_{\substack{l, m=1 \\ l \neq m}}^L \frac{a^2 t_{lm}}{\bar{r}_k^2 \bar{r}_{k'}^2} \frac{\text{Area}((B_{ik} \setminus B_{i,k-1}) \cap Z_l)}{\text{Area } Z_l} \frac{\text{Area}((B_{jk'} \setminus B_{j,k'-1}) \cap Z_m)}{\text{Area } Z_m}. \quad (24)$$

The trip coverage  $f_{ij}$  provided by a pair  $(i, j)$  of stations is then

$$f_{ij} = \sum_{k, k'} OD_{ij}(k, k').$$

These measurements were embedded within the heuristics put forward in [50] for the location of an alignment maximizing trip coverage.

## 5. Conclusions

In recent years, several cities throughout the world have constructed new rapid transit systems or built extensions of existing ones. Several new projects have also been planned. For a number of reasons, operations research techniques have not played a major role in such endeavors, although several tools developed over the past 15 years or so can assist the planning process by providing indices to measure the quality of a network, or mathematical models capable of suggesting new designs. Such models typically contain a very large number of integer variables and constraints. A number of heuristics based on classical construction and

improvement procedures or on advanced local search paradigms can be used instead of exact algorithms to yield high quality solutions. As a rule, they are rather flexible and can incorporate non-linear constraints that cannot be handled by exact algorithms.

## Acknowledgments

This research was partly funded by the Canadian Natural Sciences and Engineering Research Council under grant 39682-10, the Spanish Ministry of Science and Innovation under grant MTM 2009-14243 and the Junta de Andalucía, Spain, under grant P09-TEP-5022. This support is gratefully acknowledged. Figure 10 was kindly provided by Giuseppe Bruno. Thanks are due to a referee who provided several valuable comments on an earlier version of this paper.

## References

- [1] Rodrigue JP, Comtois C. Transport and sustainability. In: Rodrigue JP, Comtois C, Slack B (Eds), *The Geography of Transport Systems*, New York: Routledge 2009;274–279.
- [2] List of metro systems, [http://en.wikipedia.org/wiki/List\\_of\\_metro\\_systems](http://en.wikipedia.org/wiki/List_of_metro_systems) (accessed February 7, 2011).
- [3] Gendreau M, Laporte G, Mesa JA. Locating rapid transit lines. *Journal of Advanced Transportation* 1995;29:145–162.
- [4] International Association of Public Transport, <http://www.uitp.org/Public-Transport/metro/index.cfm> (accessed February 7, 2011).
- [5] World Metro Database, <http://mic-ro.com/metro/table.html> (accessed February 7, 2011).
- [6] Vuchic VR. *Urban Transit Operations, Planning and Economics*. Hoboken: Wiley 2005.
- [7] Laporte G, Mesa JA, Ortega FA. Optimization methods for the planning of rapid transit systems. *European Journal of Operational Research* 2000;122:1–10.
- [8] Blackledge DA, Humphreys EMH. The West Midland rapid transit study. *Proceedings of the Planning and Transport Research and Computation Ltd., Sussex* 1984;71–84.
- [9] Fukuyama M. Optimal station location for a two hierarchy transit system. *Proceedings of the Eight International Symposium on Transport and Traffic* 1981:264–291.
- [10] Mesa JA, Ortega FA. Park-and-ride station catchment areas in metropolitan rapid transit systems. In: Pursula, M., Niittmäki, J. (Eds), *Mathematical Methods on Optimization in Transportation Systems*. Dordrecht: Kluwer 2001:81–93.
- [11] Musso, A., Vuchic VR. Characteristics of metro networks and methodology for their evaluation. *Transportation Research Record* 1988;1162:22–33.
- [12] Laporte G, Mesa JA, Ortega FA. Assessing the efficiency of rapid transit networks. *Studies in Locational Analysis* 1994;7:105–121.
- [13] Laporte G, Mesa JA, Ortega FA. Assessing the efficiency of rapid transit configurations. *TOP* 1997;5:95–104.
- [14] Ovenden M. *Transit Maps of the World*. London: Penguin Group 2007.
- [15] Derrible S, Kennedy C. Characterizing metro networks: state, form and structure. *Transportation* 2010;37:275–297.
- [16] Minsk Metro, [http://en.wikipedia.org/wiki/Minsk\\_Metro](http://en.wikipedia.org/wiki/Minsk_Metro) (accessed February 7, 2011).
- [17] Prague Metro, <http://mappery.com/map-of/Prague-Metro-2008-Map> (accessed February 7, 2011).
- [18] London Underground, <http://www.tfl.gov.uk/assets/downloads/standard-tube-map.pdf> (accessed February 7, 2011).
- [19] Liebchen C. The first optimized railway timetable in practice. *Transportation Science* 2008;42:420–435.
- [20] Northern Line, London Underground, <http://www.trainweb.org/tubeprune/northern\%20line\%20diagrams.htm> (accessed February 7, 2011).
- [21] Grötschel M, Monma CL, Stoer M. Polyhedral and computational investigations for designing communication networks with high survivability requirements. *Operations Research* 1995;43:1012–1024.
- [22] Gattuso D, Miriello E. Compared analysis of metro networks supported by graph theory. *Networks and Spatial Economics* 2005;5:395–414.
- [23] Grubescic TH, Matisziw TC, Murray AT, Snediker D. Comparative approaches for assessing network vulnerability. *International Regional Science Review* 2008;31, 88–112.
- [24] Myung YS, Kim HJ. A cutting plane algorithm for computing  $k$ -edge survivability of a network. *European Journal of Operational Research* 2004;156, 579–589.
- [25] Albert R, Barabási A-L. Statistical mechanics of complex networks. *Review of Modern Physics* 2002;74:47–97.
- [26] Klau GW, Weiskircher R. Robustness and resilience. In: Brandes, U., Erlebach, T. (Eds), *Networks Analysis, Lectures Notes in Computer Science*, Berlin/Heidelberg: Springer 2004;3418:417–437.
- [27] Nagurney A, Qiang Q. A network efficiency measure for congested networks. *Europhysics Letters* 2007;3:1–5.
- [28] Nagurney A, Qiang Q. Robustness of transportation networks subject to degradable links. *Europhysics Letters* 2007;6:1–14.
- [29] De-Los-Santos, A., Laporte, G., Mesa, JA, Perea, F. Evaluating passenger robustness in a rail transit network. *Transportation Research Part C*. doi:10.1016/j.trc.2010.09.002.
- [30] Boston Subway, <http://www.w3.org/Conferences/WWW4/images/MBTA.gif> (accessed February 7, 2011).
- [31] Latora V, Marchiori M. Is the Boston subway a small-world network? *Physica A* 2002;314:109–113.
- [32] Criado R, Hernández-Bermejo B, Romance M. Efficiency, vulnerability and cost: An overview with applications to subway networks worldwide. *International Journal of Bifurcation and Chaos* 2007;17:2289–2301.
- [33] Costa AM, Cordeau J-F, Laporte G. Steiner tree problems with profits. *INFOR* 2006;44:99–115.
- [34] Klau GW, Ljubić I, Mutzel P, Pferschy U, Weiskircher R. The fractional prize-collecting Steiner tree problem on trees. *Lecture Notes in Computer Science*, Berlin/Heidelberg: Springer 2003;2832:1304–1315.
- [35] Lucena A, Resende MGC. Strong lower bounds for the prize collecting Steiner problem in graphs. *Discrete Applied Mathematics* 2004;141:277–294.
- [36] Margot F, Prodon A, Liebling TM. Tree polyhedron on 2-tree. *Mathematical Programming* 1994;63:183–192.
- [37] Feillet, D., Dejax, P., Gendreau, M. Traveling salesman problems with profits. *Transportation Science*, 2005; 39: 188–205.
- [38] Mesa JA, Boffey TB. A review extensive facility location in networks. *European Journal of Operational Research* 1996;95:592–603.
- [39] Labbé M, Laporte G, Rodríguez-Martín I, Salazar-González JJ. The ring star problem: Polyhedral analysis and exact algorithm.

- Networks 2004;43:177–189.
- [40] Labbé M, Laporte G, Rodríguez-Martín I, Salazar-González JJ. Locating median cycles in networks. *European Journal of Operational Research* 2005;160:457–470.
  - [41] Laporte G, Mesa JA, Perea F. A game theoretic framework for the robust railway transit network design problem. *Transportation Research Part B* 2010;44:447–459.
  - [42] Escudero LF, Muñoz S. An approach for solving a modification of the extended rapid transit network design problem. *TOP* 2009;17:320–324.
  - [43] Marín Á. An extension to rapid transit network design problem. *TOP* 2007;15:321–241.
  - [44] Marín Á, García-Ródenas R. Location of infrastructure in urban railway networks. *Computers & Operations Research* 2009;36:1461–1477.
  - [45] Schöbel A, Scholl S. Line planning with minimal traveling time. *ATMOS 2005: 5th Workshop on Algorithmic Methods and Models for Optimization of Railways* 2006:1–16.
  - [46] Borndörfer R, Grötschel M, Pfetsch ME. Models for line planning in public transport, in: *Computer-aided scheduling of public transport (CASPT)*. Lecture Notes in Economics and Mathematical systems 2008;600:363–378.
  - [47] Laporte G, Marín Á, Mesa JA, Perea F. Designing robust rapid transit networks with alternative routes. *Journal of Advanced Transportation* 2011;45:54–65.
  - [48] Bruno G, Gendreau M, Laporte G. A heuristic for the location of a rapid transit line. *Computers & Operations Research* 2002;29:1–12.
  - [49] Bruno G, Laporte G. An interactive decision support system for the design of rapid public transit networks. *INFOR* 2002;40:111–118.
  - [50] Laporte G, Mesa JA, Ortega FA, Sevillano I. Maximizing trip coverage in the location of a single rapid transit alignment. *Annals of Operations Research* 2005;136:49–63.
  - [51] Michaelis M, Schöbel A. Integrating line planning, timetabling, and vehicle scheduling: a customer-oriented heuristic. *Public Transportation* 2009;1:211–232.
  - [52] Marín Á, Jaramillo P. Urban transit network design: Accelerated Benders decomposition. *Annals of Operations Research* 2009;169:35–53.
  - [53] Marín Á, Jaramillo P. Urban transit network capacity expansion. *European Journal of Operational Research* 2008;191:45–60.
  - [54] Gleason JM. A set covering approach to bus stop location. *Omega* 1975;3:605–608.
  - [55] Poetranto Groß DR, Hamacher HW, Horn S, Schöbel A. Stop location design in public transportation networks: covering and accessibility objectives. *TOP* 2009;17:335–346.
  - [56] Schöbel A. Locating stops along bus or railway lines—A bicriteria problem. *Annals of Operations Research* 2008;136:211–227.
  - [57] Vuchic VR, Newell GF. Rapid transit interstation spacing for minimum travel time. *Transportation Science* 1968;2:303–339.
  - [58] Laporte G, Mesa JA, Ortega FA. Locating stations on rapid transit lines. *Computers & Operations Research* 2002;29:741–759.

**Gilbert Laporte** is Professor of Operations Research at HEC Montréal. His main research interests lie in the fields of distribution management, including vehicle routing and location. He has authored or coauthored 15 books and more than 375 scientific articles. In 2009 he received the Robert L. Herman Lifetime Award in Transportation Science from the Transportation Science and Logistics Society of INFORMS.

**Juan A. Mesa** is Professor at the Applied Mathematics II Department of the University of Seville. His research interests include location analysis, discrete mathematics, transportation planning and combinatorial optimization. He has published in several scientific journals. He has been the supervisor of six PhD theses and is the responsible of

several research projects funded by regional, national and European research agencies.

**Francisco A. Ortega** is Lecturer of Applied Mathematics at the Higher School of Architecture of the University of Seville. His research focuses on location models and planning transit networks. He has published several scientific papers on these topics. He is a member of the EURO Working Groups on Locational Analysis and on Transportation.

**Federico Perea** is Associate Professor at the Department of Applied Statistics and Operations Research and Quality at the Polytechnic University of Valencia. He holds a PhD degree in Operations Research. His main research interest are transportation and game theory.