Simultaneous frequency and capacity setting for rapid transit systems with a competing mode and capacity constraints

Alicia De-Los-Santos[1], Gilbert Laporte[2], Juan A. Mesa[1], and Federico Perea[3]

1 **Departamento de Matemática Aplicada II, Universidad de Sevilla, Spain**
aliciasantos@us.es, jmesa@us.es

2 **CIRRELT and Canada Research Chair in Distribution Management HEC Montreal, Canada**
gilbert.laporte@cirrelt.ca

3 **Departamento de Estadística e Investigación Operativa Aplicadas y Calidad, Universitat Politecnica de Valencia, Spain**
perea@eio.upv.es

Abstract

The railway planning problem consists of several consecutive phases: network design, line planning, timetabling, personnel assignment and rolling stocks planning. In this paper we will focus on the line planning process. Traditionally, the line planning problem consists of determining a set of lines and their frequencies optimizing a certain objective. In this work we will focus on the line planning problem context taking into account aspects related to rolling stock and crew operating costs. We assume that the number of possible vehicles is limited, that is, the problem that we are considering is a capacitated problem and the line network can be a crowding network. The main novelty in this paper is the consideration of the size of vehicles and frequencies as variables as well as the inclusion of a congestion function measuring the level of in-vehicle crowding. Concretely, we present the problem and an algorithm to solve it, which are tested via a computational experience.

1998 ACM Subject Classification G.2.2 Network problems

Keywords and phrases Line planning, railway, capacity, frequency, congestion

Digital Object Identifier 10.4230/OASIcs.ATMOS.2014.107

1 Introduction

Rapid transit planning can be divided into several consecutive phases, namely: network design, line planning, timetabling and vehicle and crew scheduling (see [13]). The second of these phases is the focus of this paper: line planning. We therefore assume that the network infrastructure (tracks and stations) as well as its associated lines are already given.

A line is characterized by several aspects: two different terminal stations, a sequence of intermediate stops, its frequency, and the vehicle capacity. The traditional line planning problem consists of finding a set of lines (a line plan) from a line pool and their frequencies providing a good service according to a certain objective, which is usually oriented towards the passengers or the operator. As in the rapid transit network design problems, the models can be classified into several categories depending on the point of view that is considered.

14th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS'14).
Editors: Stefan Funke and Matúš Mihalák; pp. 107–121
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

A classification of these models is presented in [21]. The author distinguishes between passenger oriented models and cost oriented models.

The common objective in cost oriented models is to minimize the costs related to the train operations. In [4] a mathematical programming model which minimizes the operating costs was defined. The cost structure includes fixed costs per carriage and hour, variable costs per carriage and kilometer and variable costs per train and kilometer. The problem consists of determining the lines from a line pool and their frequencies as well as the type of train operating a line and the number of carriages for each train. The passengers are assigned a priori by a modal split procedure to different types of trains. By means of binary variables representing whether a line ℓ is served by trains of type t with c carriages, a nonlinear programming model is formulated. Some techniques to make the problem more tractable are applied. A branch-and-cut approach based on the models of [4] is presented in [11]. [12] extends this model to the multi-type case in which not all trains need to stop at all stations. The authors first present a model to solve the problem of deciding for each line its frequency and the number of carriages per train. They define a set of possible combinations of lines, frequencies and capacities. Each element of this set is formed by a triple of the form (line, frequency and capacity). The only decision variable is a binary variable representing whether a triple is selected or not. They extend the model by considering different types of trains (regional, intercity and interregional). This problem is modeled as a multi-commodity flow problem.

On the other hand, passenger oriented models ensure a minimum level of quality for the passengers. One of most common objectives in the literature of passenger oriented models is to maximize the number of direct trips, see [2] and [3], which can be argued because this objective does not take travel times into account, and therefore it may yield solutions with few transfers but with too long travel times. In other papers such as [22], the objective function considered is the total travel time of all passengers, which is computed using a penalty for each transfer representing the inconvenience for the passengers. They define a graph structure named *Change&Go* graph to model the line planning problem. The problem consists of finding a set of lines and a path for each origin-destination pair, respecting a budget on the operating costs.

Our paper focusses on the line planning problem context taking into account aspects related to rolling stock and personnel costs. The problem consists of maximizing the net profit of a line plan by selecting its frequency as well as the train size of each line, assuming that all passengers preferring to travel in the Rapid Transit System (RTS) can be transported. So, we simultaneously determine the frequency and the number of carriages of the RTS trains considering at the same time, in-vehicle crowding. We assume that passengers choose their routes and their transport mode according to traveling times, which are affected by the selected frequencies and train sizes.

Our model is different from those in [4] and [12] because: the latter papers do not consider an alternative mode competing with the rapid transit transport, and we do present a model integrating the traffic assignment procedure in the optimization process and they model the problem from the operator's perspective considering different train types. So, another relevant aspect in our model is the objective function considered. Thanks to the incorporation of a logit function, the level of demand will depend on the quality of the services offered. This fact, together with the assumption that all passengers willing to travel in the RTS have to be transported, make that the model is community oriented. The component according to the operator not only expresses operating costs but also personnel and investment costs. This problem has a limitation on the number of carriages and the

RTS can become a congested network. Thus, a congestion function measuring the level of in-vehicle crowding is introduced in the model.

The remainder of the paper is structured as follows. In Section 2 we describe the problem, the needed data and notations as well as the objective function. An algorithm for solving the proposed problem is described in Section 3. The results of our computational experiments are shown in Appendix A.

2 The problem

We now formally describe the Simultaneous Frequency and Capacity Problem (SFCP), which consists of maximizing the net profit of a line plan by selecting the frequency and the train size of each line, assuming that all passengers willing to travel in the RTS can be transported. A maximum allowed capacity and frequency for each line makes that the network can become congested if a sufficiently high number of passengers want to travel in the RTS. We introduce the crowding effect by means of a congestion function which depends on the load on each arc. This function assigns a time penalty on each congested arc, therefore modifying the problem instance. The crowding effect is assumed to be the in-vehicle crowding. Thus, we want to remark that solutions in which the platform crowding appears are not taken into account.

2.1 Data and notation

The SFCP takes the following input data:

- We assume the existence of a set of stations, $N = \{i_1, \ldots, i_n\}$ and a set of lines $\mathcal{L} = \{\ell_1, \ldots, \ell_{|\mathcal{L}|}\}$ in the RTS. For the sake of readability we will identify a station with its subindex whenever this creates no confusion.
- Let n_ℓ be the number of stations of line ℓ. Each line $\ell \in \mathcal{L}$ consists of a subset of pairs of stations of N whose associated (directed) arcs form two-paths. In other words, $\ell = \{(i_1, i_2), (i_2, i_3), \ldots, (i_{n_\ell - 1}, i_{n_\ell})\}$ in such a way that i_1, i_{n_ℓ} are the terminal stations of the line, and that $\{i_1, i_2, i_3, \ldots, i_{n_\ell}\}$ and $\{i_{n_\ell}, i_{n_\ell - 1}, \ldots, i_1\}$ are paths in the network. Each couple of arcs (i_{j_1}, i_{j_2}) and (i_{j_2}, i_{j_1}) can be replaced by an edge (undirected) $\{i_{j_1}, i_{j_2}\}$.
- We define the set of edges E of the network as the union of all edges of all lines. In order to compute traffic flows we need the set of (directed) arcs associated with E. We therefore define A as the set of (directed) arcs of the network. Note that $E = \{\{i, j\} : (i, j) \in A, i < j\}$. Let $((N, E), \mathcal{L})$ be a RTS line network describing the RTS system.
- Let $d_{ij} = d_{ji}$ be the length of edge $\{i, j\} \in E$. The parameter d_{ij} can also represent the time needed to traverse edge $\{i, j\}$ by considering a parameter λ, which represents the average distance traveled by a train in one hour (commercial speed). We consider the same value of λ for all trains.
- Let ν_ℓ be the cycle time of line ℓ, that is, the time necessary for a train of line ℓ to go from the initial station to the final station and returning back. Note that $\nu_\ell = 2 \cdot len_\ell / \lambda$, where len_ℓ is the length of line ℓ.
- An undirected graph $G_{E'} = G(N, E')$, which represents the competing (private car, bus, etc.) mode network is introduced. The nodes are assumed to be coincident with those of the rapid transit mode: they could represent origin or destination of the aggregated demands; however, edges are possibly different. For each edge $\{i, j\} \in E'$, let d'_{ij} be the traversing time of such link by the competing mode.
- Let $W = \{w_1, \ldots, w_{|W|}\} \subseteq N \times N$ be a set of ordered origin-destination (OD) pairs, $w = (w_s, w_t)$. For each OD pair $w \in W$, g_w is the expected number of passengers per

hour for an average day and u_w^{ALT} is the travel time using the alternative mode of OD pair w, respectively.

With respect to costs, we distinguish three types: related to the operation, the personnel and the investment.

- Concerning rolling stock, we define a cost for operating one locomotive per unit of length c_{loc} as well as a cost representing operating cost of one carriage c_{carr} per length unit. Both parameters include running costs such as fuel or energy consumption. These terms can be easily adapted to another type of rolling stock.
- Related to the personnel costs, a cost c_{crew} per train and year is given.
- For the rolling stock acquisition, we consider two costs: the purchase price of the necessary locomotives I_{loc} per train and the purchase price of one carriage I_{carr}.
- Concerning capacity, let Θ be the carriage capacity measured in number of passengers seating and standing. We consider a minimum number $\delta^{\min}$ of carriages and a maximum number δ^{max} of carriages that can be included in a train. The capacity associated to a train is the maximum number of passengers that it can transport at any given time. More precisely, the capacity of a train of a line ℓ is equal to the capacity of a carriage (Θ) times the number of carriages forming the train (δ_ℓ). The carriage capacity is defined as the nominal capacity or crush capacity ([18], [14]) which includes both seating and standing.
- We consider a fixed finite set of possible frequencies $\mathcal{F}$ for lines of the RTS. We assume that the frequency of each line takes values between a minimum and maximum frequency in order to guarantee a certain level of service in the network.
  To be more precise, not all feasible frequency values between this minimum and maximum can be considered. Note that in real systems the frequencies have to produce a regular timetable. To take this requirement into account, we describe the set of ordered possible frequencies as $\mathcal{F} = \{\phi^1, \phi^2, \ldots, \phi^{|\mathcal{F}|}\}$, where each $\phi^q \in \mathbb{N}$, $1 \leq q \leq |\mathcal{F}|$ and $|\mathcal{F}| \geq 2$.
- Let ρ be the total number of hours that a train is operating per year and let η be the fare per trip (including the passenger subsidy) which is the same for all trips regardless of their length/duration. A parameter needed to compute the transfer time is uc_i, which represents the time spent between platforms at station i.
- We define a parameter σ in order to allow solutions that exceed the capacity by a small number of passengers.

2.2 Variables and objective function

The following variables are needed to describe our model.

- $\psi_\ell \in \mathcal{F}$ is the frequency of line ℓ (number of services per hour). A service is defined as the trains with the same route and stop stations.
- $\delta_\ell \in \{\delta^{min}, \ldots, \delta^{max}\}$ represents the number of carriages used by trains of line ℓ.
- $u_w^{RTS} > 0$ is the travel time of pair w using the RTS network.
- $f_w^{RTS} \in [0,1]$ is the proportion of OD pair w using the RTS network.
- $\tilde{f}_{ij}^{w\ell} = 1$ if the OD pair w traverses arc $(i,j) \in A$ using line ℓ, 0 otherwise.
- $\tilde{f}_i^{w\ell\ell'} = 1$ if demand of pair w transfers in station i from line ℓ to line ℓ', 0 otherwise.
- $\kappa_{ij}^\ell = \sum_{w\in W} g_w f_w^{RTS} \tilde{f}_{ij}^{w\ell} \geq 0$ is the number of passengers traversing arc (i,j) of ℓ per hour.
- $Nb_\ell = \Theta\delta_\ell\psi_\ell$ is the maximum number of passengers who can travel on line ℓ per hour.

As mentioned before, we consider the existence of public economic support for the operation of the RTS during a certain planning horizon. This assumption is very common in the rapid transit networks around the world. Usually, governments provide subsidies on the basis of the number of passengers or passenger-kilometer in order to guarantee certain positive margin to companies exploiting the transportation system.

The objective function considered is the net profit of the rapid transit network ([15], [8]). This profit is expressed as the difference between revenue and total cost in terms of monetary units over a planning horizon. The total revenue for the $\hat{\rho}$ years is computed as the number of passengers who use the RTS during the planning horizon, times η (defined as the passenger fare plus the passenger subsidy), which is the same for all passengers independently of the length of their trips. So, the revenue is mathematically expressed as

$$z_{REV} = \eta\rho\hat{\rho} \sum_{w \in W} g_w f_w^{RTS}. \tag{1}$$

The operation cost of a network is expressed by means of a fixed cost z_{FOC} and a variable cost z_{VOC}. The fixed operating cost includes maintenance costs and overheads. The fixed operating cost depends on the infrastructure. This term does not affect the objective function and is not considered, see [8]. The variable operating cost z_{VOC} over the planning horizon is defined as the sum of the crew operating cost z_{CrOC} and the rolling stock cost z_{RSOC}.

The crew operating cost z_{CrOC} includes the crew cost induced by the operation of all trains in the time horizon $\hat{\rho}$. This cost is affected by the required fleet size B_ℓ. The required fleet for each line ℓ can be defined by means of the product of its frequency and its cycle time ν_ℓ as follows:

$$B_\ell = \lceil \psi_\ell \nu_\ell \rceil = \lceil 2\psi_\ell \cdot len_\ell / \lambda \rceil,$$

where $\lceil \cdot \rceil$ is the ceiling of a number. Thus, the crew operating cost in the planning horizon is

$$z_{CrOC} = \hat{\rho} \cdot c_{crew} \sum_{\ell \in \mathcal{L}} B_\ell. \tag{2}$$

The rolling stock operation cost of a train in one hour is defined as the distance λ traveled by the train, times the cost of moving the train with δ_ℓ carriages and which is approximated by $c_{loc} + c_{carr}\delta_\ell$ ([10]). Therefore, the rolling stock operation cost in the whole planning horizon z_{RSOC} is

$$z_{RSOC} = \hat{\rho}\rho \sum_{\ell \in \mathcal{L}} B_\ell \lambda (c_{loc} + c_{carr}\delta_\ell), \tag{3}$$

and the variable operating cost in the planning horizon is $z_{VOC} = z_{RSOC} + z_{CrOC}$.

The fleet investment cost for each train is the cost of purchasing the locomotives and the carriages. Therefore, the fleet acquisition cost of all trains z_{FAC} is computed as

$$z_{FAC} = \sum_{\ell \in \mathcal{L}} B_\ell (I_{loc} + I_{carr} \cdot \delta_\ell). \tag{4}$$

So, the net profit associated to the rapid transit network is

$$z_{NET} = z_{REV} - (z_{VOC} + z_{FAC}). \tag{5}$$

2.3 Crowding

An interesting aspect to take into consideration in this problem is the crowding levels as a consequence of assuming a limited capacity. In overcrowding situations, many passengers choose an alternative path or a different transportation mode. So, congestion not only causes an increase in the traveler's disutility, but also a revenue loss to operators. The load factor ϱ^{ℓ}_{ij} is defined as the ratio $\kappa^{\ell}_{ij}/Nb_{\ell}$. Observe that if $\varrho^{\ell}_{ij} \leq 1$, the arc $(i,j) \in \ell$ is not affected by the congestion. Therefore, if the train capacity of a line ℓ is not enough to transport all passengers traveling inside ℓ, the rapid transit network can become a congested network. In recent research, the load factor is introduced to estimate the crowding levels. There exists four crowding types: in-vehicle crowding, platform crowding, excessive waiting time and increased dwell time. We will concentrate on the analysis of in-vehicle crowding effects, which can be defined by means of crowding penalties. This term can be expressed in three possible ways: time multiplier, the monetary value per time unit, and the monetary value per trip. We will use the time multiplier in our problem. Since each transport mode is different, it is not possible to define a general crowding function valid for all transport modes. [6] proposed an exponential function for the crowding penalty in the context of railway system using a load factor. This crowding function is expressed as

$$CF(\varrho^{\ell}_{ij}) = 1 + \frac{\varsigma_1}{1+\exp(\varsigma_2(1-\varrho^{\ell}_{ij}))} + \varsigma_3 \exp(\varsigma_4(\varrho^{\ell}_{ij} - \varsigma_5)), \tag{6}$$

where all parameters are positive values and ς_1 and ς_3 should be calibrated. The last parameter $\varsigma_5 > 1$ is the threshold from which the passengers start to perceive overcrowding (the crowding penalty can grow exponentially). Note that this function reflects the inconvenience associated with in-vehicle crowding. Observe that if the load factor $\varrho^{\ell}_{ij} \leq 1$, $CF(\varrho^{\ell}_{ij})$ is approximately one; the second term in Equation (6) is approximately zero for a proper value of parameter ς_2 and the third term $\varsigma_3 \exp(\varsigma_4(\varrho^{\ell}_{ij} - \varsigma_5))$ is close to zero (recall $\varrho^{\ell}_{ij} < \varsigma_5$). Similarly, when the load factor $1 \leq \varrho^{\ell}_{ij} \leq \varsigma_5$, in-vehicle crowding starts influencing the time of arc $(i,j) \in \ell$. The penalty impact will depend on the ς_2 parameter.

Due to the fact that we are only including in-vehicle crowding effects, solutions whose load factor is greater than the parameter σ are not allowed. Observe that if $\varrho^{\ell}_{ij} > \sigma$, penalties according to the excess waiting time, platform crowding and increased dwell time would have to be included in the model.

We consider $\bar{d}^{\ell}_{ij} = CF(\varrho^{\ell}_{ij}) \cdot d_{ij}$ as the perceived time to traverse arc (i,j) of ℓ using the rapid transit system. As commented before, if the arc $(i,j) \in \ell$ is not congested, $\bar{d}^{\ell}_{ij} \simeq d_{ij}$. The average travel time associated to the OD pair w using the rapid transit network under crowding can be explicitly defined as follows:

$$\begin{aligned} u^{RTS}_w \quad &= \sum_{\ell\in\mathcal{L}} \sum_{j:\{w_s,j\}\in\ell} \frac{60\tilde{f}^{w\ell}_{w_s j}}{2\psi_\ell} + (60/\lambda) \sum_{\ell\in\mathcal{L}} (\sum_{\{i,j\}\in\ell} \tilde{f}^{w\ell}_{ij} \bar{d}^{\ell}_{ij}) \\ &+ \sum_{\ell\in\mathcal{L}} \sum_{\ell':\ell'\neq\ell} \sum_{i\in\ell\cap\ell'} \tilde{f}^{w\ell\ell'}_i (\frac{60}{2\psi_{\ell'}} + uc_i),\ w = (w_s, w_t) \in W. \end{aligned} \tag{7}$$

The first term in (7) is the waiting time at the origin station, which is also assumed to be half of time between services of this line. The second term represents the in-vehicle time which can be affected by congestion. Finally, the third term is the transfers time.

Another variable that can be explicitly defined is the assignment f^{RTS}_w of demand to the RTS system. As mentioned, we assume the number of passengers who use a transport system varies depending on the provided service. More specifically, the proportion of an OD pair using each mode may be different depending on the characteristics of the RTS to be designed and on the competing transport mode. Therefore, the demand is split between the RTS and the alternative mode according to the generalized cost of each mode. The modal split is modeled by using logit type functions, see [19], as opposed to binary variables which are used in very complex problems. In [20] the route decisions are integrated in the line planning problem. To this end, the authors consider a *Change&Go* on which, a modified Dijkstra algorithm is applied and adapted to compute shortest paths of origin-destination pairs.

In order to define the logit function, we need two positive real parameters α and β for each transport mode. The parameter α simulates the market share for each mode and β weights the importance of each mode, see [17]. We consider, α^{RTS} for the RTS mode and α^{ALT} in the alternative mode. In order to express the same importance to both modes, the parameter β is independent of the mode as in [9]. Let us denote $\alpha = \alpha^{ALT} - \alpha^{RTS}$. Therefore, the proportion of the OD pair w using the RTS mode is

$$f_w^{RTS} = \frac{1}{1 + e^{(\alpha - \beta(u_w^{ALT} - u_w))}}, \, w \in W. \tag{8}$$

The logit model estimates the proportion of users assigned to each mode for each origin-destination pair in a continuous way. Note that this proportion depends on the travel time in each transport mode, which is modified if the congestion function is activated. Concretely, the congestion effect influences the travel time of each path, and, therefore, the number of passengers in the RTS. The passengers' behavior is different in congestion presence and, as a consequence, it is different for each instance. It can be observed that the penalization process stops when the network is not congested or a fixed point is found. In other words, passengers take a different path or mode and an equilibrium is searched (all passengers can be transported). The solution reflects not only the number of carriages and frequencies, but also a medium-term analysis of the passenger's behavior under congestion.

This problem can be extended to situations where the excess waiting time is taken into account, e.g. platform crowding. This term affects passengers waiting for next train if the first train was full and they were left behind, therefore increasing waiting time and discomfort to travel. [18] presented a formal definition of this type of crowding in the context of bus transport. They expressed the waiting time by means of headway and crowding level. However, the inclusion of excess waiting time effects in our model is not immediate. To this end, the travel time of all passengers waiting for next train is increased according to an additional time which depends on the frequency of the congested line. Rerouting passengers is very complicated because the passengers affected by the excessive waiting time have different travel time than the rest of passengers and, as a consequence, a different instance associated. So, the initial instance is divided into two different instances: one associated to in-vehicle crowding and the other one, related to excessive waiting time. Analogously, the origin-destination matrix is divided into two matrixes: one containing the passenger associated to the in-vehicle crowding and other one, according to the excessive waiting time. The crowding phenomenon is also treated as the congestion effect at train stations; the access/egress to/from the station, on platforms (see [7]) and on the increased dwell times as [16].

The following section is devoted to introducing an algorithm to solve our problem.

3 An algorithm

In this section we introduce two algorithms that solve our problem: one with the nominal

Algorithm 1: The algorithm for the rapid transit network frequency and capacity setting problem under congestion with nominal capacity.

Data: A line network $(S, \mathcal{L})$, a set of possible frequencies and a minimum and maximum capacity

for *each possible combination of frequencies and carriages* **do**
    Loop III: Check capacity constraint
    for *each line* ℓ **do**
        Find the arc $(i, j) \in \ell$ with maximum load ϱ_{ij}^{ℓ};
        if $1 < \varrho_{ij}^{\ell} \leq \sigma$ **then**
            penalize the traverse time of each arc by means of CF-function;
            go Loop IV;
        end
    end
    Compute the profit;
end

Result: The solution with the maximum profit.

capacity and other one with number of seats. Each of them consists of analyzing each possible frequency (number of services per hour for each line) and each number of vehicles (number of carriages per train of each line). The idea is to iteratively check all possible combinations of frequencies and carriages. Once the frequencies and carriages have been set, the shortest path that takes into account transfer and waiting times on the rapid transit network for each OD pair can easily be calculated by a modified Dijkstra algorithm. From these shortest paths we compute the number of passengers traveling on each line and arc. At this point, the capacity constraint is checked on the arc with maximum load. If there exists a congested arc, the penalization process is activated. The travel time to traverse each arc is increased by means of its corresponding penalty. Once the penalization process is finished, the rerouting process is activated. To this end, the shortest path taking into account transfer and waiting times on the RTS for each OD pair is recalculated and the capacity constraint is rechecked and so on. Due to the travel time increase, the number of passengers on congested arcs is smaller than the previous iteration. Some passengers will take an alternative path or an alternative transport mode. This procedure breaks when the congestion ends or when a fixed point is found. Algorithm 1 shows the pseudocode to solve the SFCP with nominal capacity and Algorithm 3 is the pseudocode to solve the SFCP with the number of seats on each carriage. For the congestion with the seat capacity, the in-vehicle crowding is activated when the load factor reaches 140% or standing density is over four passengers per square meter (see [8]).

4 Conclusions

We have introduced a problem in the line planning context, in which the number of carriages is also a decision variable. Concretely, the problem consists of selecting, for each line, the number of services per hour and the number of train carriages in presence of a competing transportation mode. We have assumed that all passengers that want to use the RTS have a service and a certain net benefit is maximized. To this end, we have incorporated a long term public economic support for the operating and acquisition rolling stock. This problem can lead to congested networks since the maximum number of possible carriages is bounded.

Algorithm 2: Testing the fixed point.

Data: A line network $(S, \mathcal{L})$
Loop IV: Check fixed point
if *the number of iterations is equal to one* **then**
    $(S^{pre}, \mathcal{L}^{pre})=(S, \mathcal{L})$;
else
    if *the network* $(S, \mathcal{L})$ *is the same than* $(S^{pre}, \mathcal{L}^{pre})$ **then**
        break;
    else
        $(S^{pre}, \mathcal{L}^{pre})=(S, \mathcal{L})$;
        go Loop III;
    end
end
Result: A network.

The input data in the computational experiments has been based on real data in order to calibrate all parameters that appear in our problem. Moreover, we have randomly generated instances for different types of networks. The algorithm defined in Section 3 has been tested on small networks showing the effect of the congestion on the solutions. The congestion impact has been analyzed by means of a congestion function which measures the level of in-vehicle crowding. A total of 200 experiments were carried out in our analysis. From the results obtained, we observe that the profit is economically more interesting when the network is not congested (according to the randomly generated instances we have solved). In other words, the demand is sensitive to congestion and it is more profitable to add carriages than to lose passengers.

This problem can easily be extended to the case of a set of possible lines (a line pool) analyzing iteratively all combinations of lines. For each possible set of lines, the problem is reduced to our problem.

The proposed algorithm has to solve an underlying unconstrained non-linear optimization problem, and therefore the obtained solution is not guaranteed to be optimal. A potential line of research will be to check wether or not this algorithm is exact, that is, whether or not the solution returned is optimal.

Due to the complexity of the problem, and the fact that the proposed algorithm is only suitable for small-medium sized instances, future research will focus on heuristic approaches.

Another way of completing this research will be about existence and uniqueness of passenger flow equilibria and, if so, on the question whether the algorithm proposed converges to them.

Acknowledgements This research work was partially supported Ministerio de Economía y Competitividad (Spain)/FEDER under grant MTM2012-37048, by Junta de Andalucía (Spain)/FEDER under excellence projects P09-TEP-5022 and P10-FQM-5849 and by the Canadian Natural Sciences and Engineering Research Council under grant 39682-10.

References

1 A. Alfieri., R. Groot, L. Kroon and A. Schrijver. *Efficient Circulation of Railway Rolling Stock*. *Transportation Science*, 40:(3)378–391, 2006.

Algorithm 3: The algorithm for the rapid transit network frequency and capacity setting problem under congestion with seat capacity.

Data: A line network $(S, \mathcal{L})$, a set of possible frequencies and a minimum and maximum capacity

for *each possible combination of frequencies and carriages* **do**
  Loop III: Check the capacity constraint
  for *each line* ℓ **do**
    Find the arc $(i, j) \in \ell$ with maximum load ϱ^{ℓ}_{ij};
    Let $\hat{\varrho}^{\ell}_{ij}$ be the load with the nominal capacity;
    if $\varrho^{\ell}_{ij} > 1.4$ *and* $\hat{\varrho}^{\ell}_{ij} \leq \sigma$ **then**
      penalize the traverse time of each arc by means of CF-function;
      go Loop IV;
    end
  end
  Compute the profit;
end

Result: The solution with the maximum profit.

2 M.R. Bussieck. *Optimal Line Plans in Public Rail Transport. Ph.D.* Technical University Braunschweig, 1998.

3 M.R. Bussieck, P. Kreuzer, and U.T. Zimmermann. Optimal lines for railway systems. *European Journal of Operational Research*, 96:54–63, 1997.

4 M. T. Claessens, N. M. van Dijk, and P. J. Zwaneveld. Cost optimal allocation of rail passenger lines. *European Journal of Operational Research*, 110(3):474–489, 1998.

5 J.-F. Cordeau, F. Soumis, and J. Desrosiers. A Benders decomposition approach for the locomotive and car assignment Problem. *Transportation Science*, 34(2):133–149, 2000.

6 A. De Palma, M. Kilani, and S. Proost. The localization and pricing of mass transit stations. 2010. In Proceedings of the 12th World Conference on Transport Research Society.

7 N. Douglas and G. Karpouzis. Estimating the cost to passengers of station crowding. 2005. In 28th Australasian Transport Research Forum (ATRF).

8 Q. Feifei. Investigating the in-vehicle crowding cost functions for public transit modes. *Mathematical Problems in Engineering*, 2014:13, 2014.

9 R. García, A. Garzón-Astolfi, A. Marín, J.A. Mesa, and F.A. Ortega. Analysis of the parameters of transfers in rapid transit network design. In L. G. Kroon and R. H. Möhring, editors, *5th Workshop on Algorithmic Methods and Models for Optimization of Railways*, Schloss Dagstuhl, Germany, 2006.

10 A. García, and M.P. Martín. Diseño de los vehículos ferroviarios para la mejora de su eficiencia energética. In *Monografías ElecRail*, volume 6, 2012.

11 J.-W. Goossens, S van Hoesel, and L. Kroon. A branch-and-cut approach for solving railway line-planning problems. *Transportation Science*, 38:379–393, 2004.

12 J.-W. Goossens, S. van Hoesel, and L. Kroon. On solving multi-type railway line planning problems. *European Journal of Operational Research*, 168(2):403–424, 2006.

13 V. Guihaire and J.K. Hao. Transit network design and scheduling: A global review. *Transportation Research Part A: Policy and Practice*, 42(10):1251–1273, 2008.

14 S. Jara-Díaz and A. Gschwender. Towards a general microeconomic model for the operation of public transport. *Transport Reviews*, 23(4):453–469, 2003.

15 Z.-C. Li, W. H. K. Lam, S. C. Wong, and A. Sumalee. Design of a rail transit line for profit maximization in a linear transportation corridor. *Transportation Research Part E: Logistics and Transportation Review*, 48(1):50–70, 2011.

16 T. Lin and N.H.M. Wilson. Dwell time relationships for light rail systems. *Transportation Research Record*, 1361:287–295, 1992.

17 A. Marín and R. García-Ródenas. Location of infrastructure in urban railway networks. *Computers & Operations Research*, 36(5):1461–1477, 2009.

18 R.H. Oldfield and P.H. Bly. An analytic investigation of optimal bus size. *Transportation Research Part B*, 22(5):319–337, 1988.

19 J.D. Ortúzar and L.G. Willumsem. *Modelling Transport*. Wiley, Chichester, 1990.

20 M. Schmidt and A. Schöbel. The complexity of integrating routing decisions in public transportation models. In T. Erlebach and M.E. Lübbecke, editors, *ATMOS10*, volume 14 of *OASICS*, pages 156–169. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, 2010.

21 A. Schöbel. Line planning in public transportation: models and methods. *OR Spectrum*, 34(3):491–510, 2011.

22 A. Schöbel and S. Scholl. Line planning with minimal traveling time. In L. Kroon and R.H. Möhring, editors, *5th Workshop on Algorithmic Methods and Models for Optimization of Railways, ATMOS 2005*, Dagstuhl, Germany, 2006. Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany.

A Computational experiments

All the calculations in this section were performed with a Java code in a computer with 8 GB of RAM memory and 2.8 CPU Ghz. In order to evaluate the performance of our algorithm, we have used several instances of networks (see A.1).

There are no previously reported solutions for the proposed problem as far as we know. We have performed tests to asses the impact of the congestion on the networks 6×2, 7×3 and 8×3. To this end, we have gradually increased the number of carriages and we have found a solution to the problem with our algorithm (see Algorithm 1). The results of these experiments are available from the authors. We noted that, when the maximum number of carriages is small, the optimal solution has high frequencies in order to transport all passengers. This is due to the problem definition: we have imposed that all passengers willing to travel in the RTS have to be transported.

A key factor to solve this problem was the introduction of the congested function defined in Section 2.3, which is based on in-vehicle crowding.

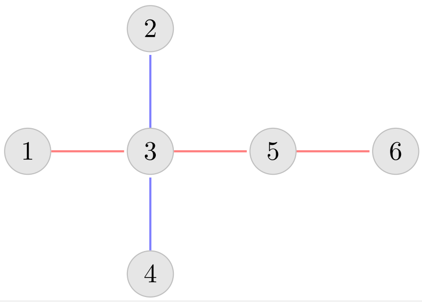
In the next section, we introduce all parameters needed to carry out the experiments as well as the considered networks. In order to keep the paper within the 15-page limit, detailed results of the experiments and other parameters of the problem such as number of possible trips in each instance have been omitted, but are available from the authors upon request.

A.1 Parameter setting

In Table 1 we report the values considered for the parameters of our algorithm. The data reported in this table are based on the specific train model *Civia*, usually used for regional railway passengers transportation in Spain by the National Spanish Railways Service Operator (RENFE). One important characteristic of *Civia* trains is that the number of carriages can be adapted to the demand. Each *Civia* train contains two electric automotives (one at

Table 1 Model parameters for SFCP.

Parameters		
Name	**Description**	**Value**
$\hat{\rho}$	years to recover the purchase	20
ρ	number of operative hours per year	6935
c_{loc}	costs for operating one locomotive per kilometer [€/km]	34
c_{carr}	operating cost of a carriage per kilometer [€/km]	2
c_{crew}	per crew and year for each train [€/ year]	$75 \cdot 10^3$
I_{loc}	purchase cost of one locomotive in €	$2.5 \cdot 10^6$
I_{carr}	purchase cost of one carriage in €	$0.9 \cdot 10^6$
Θ	capacity of each carriage (number of passengers)	$2 \cdot 10^2$
λ	average commercial speed in [km /h]	30
γ	maximum number of lines traversing an edge	4
ψ^{min}	minimum frequency of each line	3
ψ^{max}	maximum frequency of each line	20
ψ_ℓ	possible values	{3,4,5,6,10,12,15,20}



The lines are defined as:
red line $\ell_1 = \{1, 3, 5, 6\}$ and
blue line $\ell_2 = \{2, 3, 4\}$.

Figure 1 Representation of 6×2-configuration.

each end) and a variable number of passenger carriages. Each automotive or carriage has a maximum capacity of 200 passengers. In our experimentation, we will assume that the train is composed by only one electric locomotive (for traction purposes and null capacity) and several passengers carriages (which cannot move without a locomotive) as in [5] and [1]. The purchase price of rolling stock used in this experimentation is also based on the real data of *Civia* trains. The price of ticket and subvention considered in our experimentation, have been taken from the newspaper (http://www.20minutos.es/noticia/2028399/0/madrid/empresas-privadas/metro-ligero/).

In the experiments we have considered five network topologies. The first one is defined by six nodes, five edges and two lines as follows The second one is a star network with six nodes and three lines.

The following network is defined by eight nodes, nine edges and three lines. For each configuration, we have randomly generated 10 different instances for the OD-matrix and length data. To this end, the number of passengers of each OD pair w, was obtained according to the product of two parameters. The first one was randomly set in the interval [5,15] by using a uniform distribution, whereas the other one was set in a different interval for each configuration. Concretely, for the 6×2-network, the interval considered was set as [65,77], generating around 20.000 passengers at each instance of such configuration. For

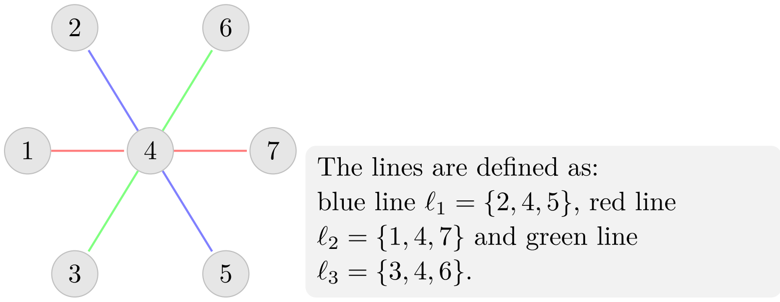


Figure 2 Representation of 7×3-configuration.

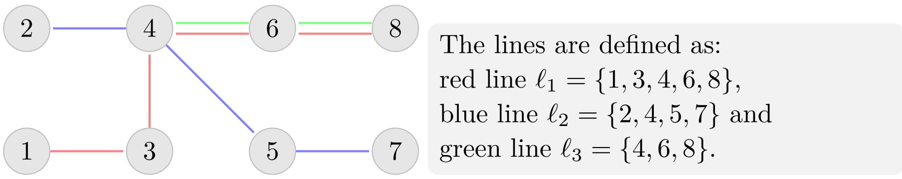


Figure 3 Representation of 8×3-configuration.

7×3 and 8×3-networks, the number of passengers was approximately 30.000 passengers at each case and the parameters were defined in the intervals [68, 80] and [51, 59], respectively. The parameter for the 20×6-configuration was set to 16 for all instances and [23, 25] for the 15×5-configuration.

To define each arc length, the coordinates of each station were set randomly by means of an uniform distribution. So, the arc length at each instance is different since each arc connects to different stations.

For the experiments, the travel times u_w^{alt} by the alternative mode, were obtained by means of the Euclidean distance and a speed of 20 km/h, whereas, the travel times in the RTS were obtained according to in-vehicle travel time, waiting and transfer times. The waiting time was supposed to be half of the corresponding time between services of lines at the origin station, whereas, the transfer time was assumed to be half time between two consecutive services at the line to transfer. We assume two possible values for the σ parameter: 1.1 and 1.2. So, for $\sigma = 1.1$, if the number of passengers traveling inside each line is 10% higher than its capacity, the solution is taken into account.

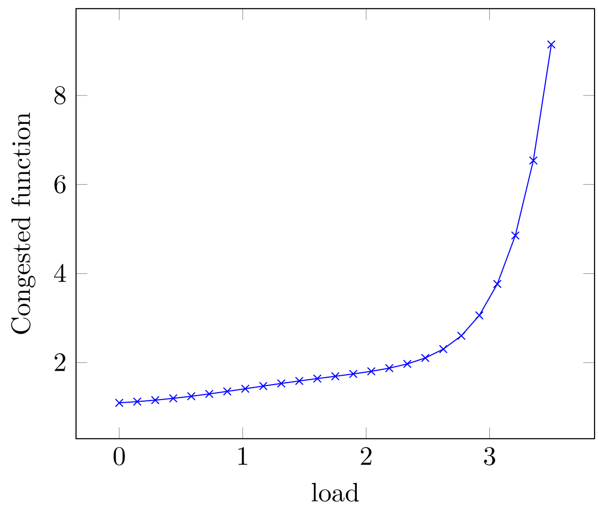
A.2 Computational experiments for our problem

To evaluate the performance of our algorithm, we have adapted the crowding function defined in Section 2.3 to our problem. Concretely, the crowding penalty was mathematically defined for the nominal capacity as

$$CF(x) = 1 + \frac{0.8}{1 + \exp(2 * (1 - x))} + 0.01 \exp(3 * (x - 1.3)). \tag{9}$$

The following figures show a representation of the crowding functions above defined.

In order to evaluate the impact of the in-vehicle crowding on the solution of our problem, we have gradually increased the maximum number of carriages in our experimentation. Moreover, we have analyzed the solutions obtained at the uncapacitated case (an unlimited number of carriages) when the in-vehicle crowding is introduced.



load	penalty
1.08	1.44
1.05	1.42
1.01	1.41

Figure 4 In-vehicle crowding function.

The parameter σ considered here was fixed to 1.1, which implies that if the number of passengers of each line is 10% higher than its capacity, the solution is taken into account. A total of 200 experiments of the 6×2, 7×3 and 8×3-configuration were tested. For 6×2 configuration our algorithm was able to obtain a solution in a very small CPU time (3 to 6 seconds). However, on larger instances, the algorithm took too long (in some instances over an hour). It is difficult, if not impossible, to conduct experiments over real instances, which indicates the need of applying heuristic strategies to solve the problem.

A.2.1 6×2-configuration

We have analyzed the solutions when the parameter δ^{max} is less than or equal to 8. For $\delta^{max} \leq 4$, the problem is always infeasible. For $\delta^{max} > 4$, most cases are feasible and the optimal solutions are not affected by congestion (see the seven column). This fact indicates the maximum number of carriages is a sufficient number in order to transport all passengers willing to use the RTS. The average CPU time is 6.15 seconds when $\delta^{max} = 8$.

A.2.2 7×3-configuration

For $\delta^{max} \leq 2$, the optimal solutions have high frequencies in order to transport all passengers. From the results we observed that the number of trains decreases when the maximum number of carriages increases. The profit starts to be economically interesting when the number of carriages is greater than two for some instances and it is greater than three for some others. The most cases, the optimal solution corresponds to a non-congested network. It is interesting to note that for $\delta^{max} = 6$ only two instances yield the same solution. This fact indicates the in-vehicle crowding directly affects the solutions. Indeed, the optimal solutions for the uncapacitated case are affected by the in-vehicle crowding when the congestion is introduced in our problem. For instance, we could observe that the optimal solution for one instance of the capacitated case has one more carriage than the solution to the uncapacitated case. In other words, when the congestion is taken into account, the passenger's

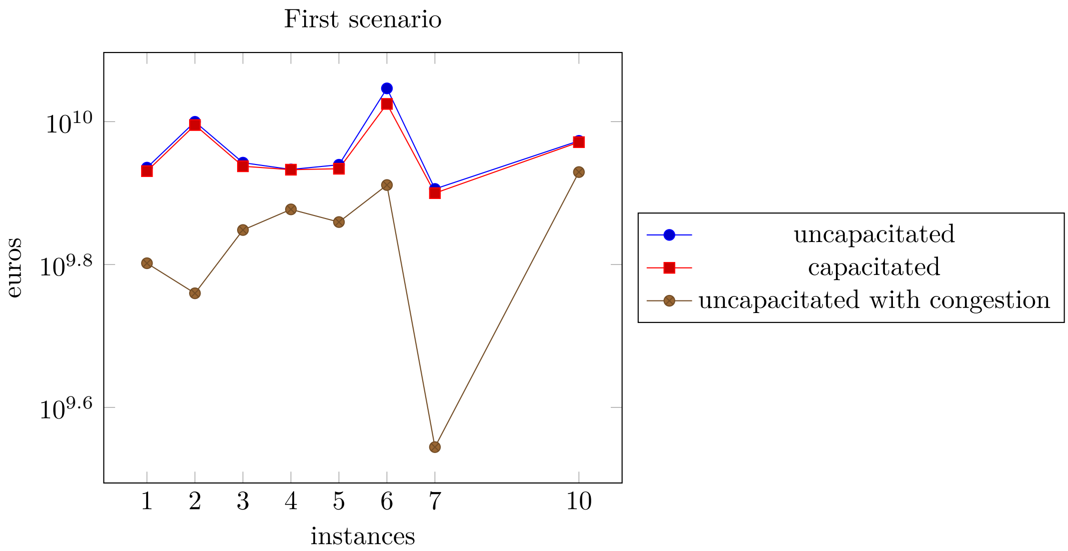


Figure 5 Profit for 7×3-configuration. Optimal solution for uncapacitated problem, capacitated problem and the uncongestion optimal solution with the congestion effect.

behavior changes, and it is economically more interesting to add a carriage than to lose passengers.

A.2.3 8×3-configuration

The results obtained reveal that, for most cases, the system becomes productive from three carriages on. In 7×3-configuration, the frequencies are high when the capacities are small, in order to transport all passengers willing to travel on the RTS. The average CPU time for $\delta^{max} = 1$ is 1.36 seconds whereas for $\delta^{max} = 6$ is 823. The optimal solutions for uncapacitated case are affected by the in-vehicle crowding at the capacitated case, as shown in in our experiments.