

Transit Network Timetabling and Vehicle Assignment for Regulating Authorities

Valérie Guihaire^{*,a,b}, Jin-Kao Hao^b

^a*Perinfo SA, 1 rue de Metzeral 67100 Strasbourg, France*

^b*LERIA, Université d'Angers, 2 Boulevard Lavoisier, 49045 Angers Cedex 01, France*

Abstract

In the literature on transit planning, network timetabling and vehicle scheduling are usually treated as separate problems. In this paper, we focus on combining important features of these two steps and propose a simultaneous solution approach to redefine timetables with the objective of bringing improvements to both quality of service and vehicle costs incurred. This includes the objectives of quantity and quality of the transfers proposed, evenness of the line headways, fleet size and length of the deadheads. The model proposed for this simultaneous approach is adapted to the problem faced by regulating authorities, encouraging intermodality and taking into account a variety of practical features. We introduce an optimization procedure based on Iterated Local Search and present computational experiments carried out on data from a large existing transit network, showing substantial improvements in both quality of service and level of resources compared to the current practice.

Key words: Transit network, timetabling, transportation

1. Introduction

Transit network timetabling and vehicle assignment are two key steps in the process of transit network design (Ceder and Wilson, 1986; Guihaire and Hao, 2008). Any solution approach to the combined problem (TVA) must

*Corresponding author. Tel: +33 954304114; Fax:+33 388105671

Email addresses: vguihaire@perinfo.eu (Valérie Guihaire),
hao@info.univ-angers.fr (Jin-Kao Hao)

take into account a set of constraints and optimization objectives including in particular the quality of service (user side) and the resource utilization (supplier side).

Basically, the *timetabling problem* consists in creating the timetable of the line runs for the transit network. This problem has been the object of studies for decades, and keeps challenging researchers on problem definition, modeling and solution methods (Wong and Leung, 2004; Cevallos and Zhao, 2006). There exists many different approaches to this problem, targeting mostly quality of service (Ceder et al., 2001; Jansen et al., 2002). It has often been modeled as a Quadratic Semi-Assignment Problem (QSAP) (Bookbinder and Désilets, 1992) which aims at minimizing the global transfer waiting time of passengers in the network by setting the first departure time of each line.

Timetabling-like problems can also be found in other transportation areas like railway and aircraft, but the constraints and objectives considered are often very different from those encountered in transit network systems.

The *vehicle assignment problem*, on the other hand, consists in assigning vehicles to line runs and depots, thereby creating the so-called vehicle services. Several aspects have been studied, considering different levels of complexity, such as number of depots (Single Depot Vehicle Scheduling Problem (Freling et al., 2001) or Multiple Depot Vehicle Scheduling Problem (Pepin et al., 2009; Laurent and Hao, 2009; Hadjar et al., 2006; Laurent and Hao, 2009b)) or fleet homogeneity/heterogeneity.

Recent developments in transit planning include the multiplication of proposals for the integration of steps such as transit network design and frequencies setting (Zhao and Zeng, 2006) and timetabling (Zhao and Zeng, 2008) or vehicle and driver scheduling (Huisman et al., 2005; Laurent and Hao, 2007, 2008). Each step has direct influence on the next one and transitively on all the following steps. Although it has been shown that these integrations provide benefits, the simultaneous approach of timetabling and vehicle assignment has not been much explored so far. While some studies integrate constraints on the number of available vehicles in the transit network timetabling problem (Castelli et al., 2004), the first study that we are aware of that considers the number of vehicles as an objective is (Chakroborty et al., 2001), in which the case of a single transfer stop with multiple lines is studied. In 2004, Fleurent et al. (2004) also mention including this objective in their approach, though they hardly bring up their resolution method.

To better understand this part of the transit planning process, let us

quickly describe the roles of each actor of the regulated market. Regulating authorities define the transit network structure and timetables, with the objective of securing a certain level of quality of service. Then they delegate the actual transportation service by network portions (often consisting of groups of geographically close lines) to transit operators through calls for tender. The price to pay depends on one main factor, the number of vehicles needed, and on a secondary one, the number of kilometers. Since this problem is aimed at being addressed before the choice of one transit operator, details concerning the fleet of vehicles and depot locations are not yet available. The fleet is thus assumed to be homogeneous with unlimited capacity: one assumes overcrowding is taken care of by the frequencies.

In this paper, we introduce a rich model and a heuristic solution approach to the integrated problem of timetabling and vehicle assignment. Instead of considering timetabling and vehicle assignment in a sequential way, we propose to tackle them simultaneously in the hope of offering improved solutions. We integrate a set of realistic constraints and objectives within our model and devise a heuristic solution method based on the Iterated Local Search metaheuristics combined with an exact linear quasi-assignment algorithm. The final goal of the work is to create a flexible decision-aid tool that enables the human planner to favor or combine any set of criteria like quality and number of transfer opportunities, headway evenness, number of vehicles needed and length of deadheads.

The paper is organized as follows. In section 2, a description of the TVA problem is presented, including definitions, input, output, constraints and objectives. This description is associated to a formulation of the problem as a constrained optimization problem. In section 3, we expose our solution method. Results are presented and analyzed in sections 4, based on computational experiments carried out on a real transit network from the area of Orléans, France.

2. Problem description and formulation

The problem considered in this paper can be informally described as follows. Given a pre-defined lines network, including stops, line routes and timetables; groups of lines sharing resources; period-dependent running times; period-dependent expected headways with variation margins; and levels of importance of the transfers; the goal is to define a synchronized network timetable and create the associated vehicle services with respect to a set of

imperative constraints and objectives. A network timetable is composed of line timetables, which in turn correspond to the set of all arrival and departure times for the stops served by each line run. A vehicle service is the complete and ordered sequence of line runs it is assigned to.

In this section, we present the timetabling and vehicle assignment problem, its input, output, constraints and objectives. We start by a few definitions:

- A *line* is an ordered and directional sequence of stops. A *line run* is a trip on the line characterized by a departure time from its initial terminal. An *external line* is any activity (e.g. a train line, a factory quitting) that can be connected with the transit network through intermodal transfers.
- The *headway* of a line is the time separating the service of its main stop by consecutive runs. It is the inverse of the frequency over a time period.
- The *turnaround time* is the time needed by a vehicle at the end of a line run to get ready for the next trip and possibly catch up with some lateness.
- A *deadhead* is a trip without passengers that links consecutive line runs on a given vehicle service.

In this section, we present our model for the combined transit network timetabling and vehicle assignment problem as a constrained optimization problem. This formulation allows us to formally state the problem and constitutes a basis for developing our heuristic solution approach.

2.1. Notations

Set of line groups \mathcal{G} , set of activities \mathcal{A} (including the set of internal lines \mathcal{L} and the set of external lines \mathcal{M}), set of line runs \mathcal{R} , set of bus stops \mathcal{S} , set of transfers \mathcal{T} , set of vehicles \mathcal{V} , planning horizon \mathcal{H} .

2.2. Input

- For each line $l \in \mathcal{L}$:
 - m_l : main stop.
 - \mathcal{P}_l : set of homogeneous headway periods. For each $p \in \mathcal{P}_l$:

- * $bt_{l,p}, et_{l,p}$: time of beginning and end of the period.
- * $h_{l,p}, v_{l,p}$: expected headway and allowed variation margin.
- $tt_{l,s}$: turnaround time per final stop $s \in \mathcal{S}$.
- g_l : line group comprising l . Interlining is allowed inside line groups.
- \mathcal{R}_l and r_l^α, r_l^ω : set of runs among which, first and last runs.
- For each line run $r \in \mathcal{R}$:
 - l_r : internal line to which r belongs.
 - $\mathcal{S}^r \subset \mathcal{S}$: set of served stops.
 - $\mathcal{S}^{*r} \subset \mathcal{S}^r$: restricted set of important served stops.
 - s_r^α, s_r^ω : first and last stop.
 - s_r^+ : stop following s .
 - $h_{r,s}^-, h_{r,s}^+$: arrival and departure time for $s \in \mathcal{S}^r$ in the existing timetable.
 - $st_{r,s}$: stopping time per stop $s \in \mathcal{S}^r$.
- For each external line $m \in \mathcal{M}$:
 - s : connecting point with the network.
 - \mathcal{C}_m : set of arrivals at s or departures from s .
 - $h_{m,c}^-, h_{m,c}^+$: arrival, departure time of $c \in \mathcal{C}_m$.
- For each transfer $t(a_1, s_1, a_2, s_2) \in \mathcal{T}$ (with $a_1, a_2 \in \mathcal{A}, s_1, s_2 \in \mathcal{S}$):
 - il_t : importance level.
 - $\lfloor wt_t \rfloor, [wt_t], \lceil wt_t \rceil$: minimal, ideal and maximal waiting time.
- $dt_{s,s'}$: deadheading time between stops s and s' .
- $rt_{s,s'}^h$: running time between s and s' per time point $h \in \mathcal{H}$.

2.3. Output - Decision variables and values

The output of the TVA problem is defined by: transit lines timetables, fleet size and vehicle services.

In most studies, (see e.g. Jansen et al. (2002)), the decision variables for the timetabling problem are the line departure times, from which the complete timetable can be deduced. Indeed, they consider fixed headways, as well as fixed stopping and running times.

In this paper, we consider a more flexible model with variable headways, fixed stopping times at stops and given period-dependent running times. With this model, all arrival and departure times can be deduced for the rest of the stops from the starting time of each line run.

- One *decision variable* is a line run on the timetable.
- The *value* of a decision variable is a (starting time, vehicle) pair (see Figure 1). The domain of these values is discrete (unit of time is the minute) and finite. The number of vehicles is at most equal to the number of line runs.
- One *configuration* σ is the complete assignment of values to the set of decision variables: $\sigma : \mathcal{R} \rightarrow (\mathcal{H} \times \mathcal{V})$.

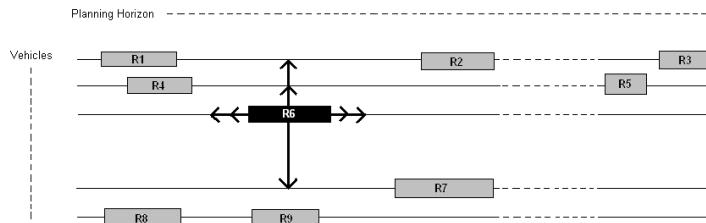


Figure 1: Timetable and Schedule coupling

Rectangles represent the decision variables, i.e. the line runs.

A configuration corresponds to the positioning of each line run on the two-dimension space created by the fleet of vehicles and the time frame.

Additionally, a set of state variables is maintained to ease the calculation of values related to transfers and headway evenness. They include the arrival

and departure times of line runs for a restricted set of stops. This set is composed of the main stop of the line, the first and last stop of each line run, and the stops at which transfers can take place. $\forall r \in \mathcal{R}, \forall s \in \mathcal{S}^r, \pi_{r,s}^{\leftarrow} \in \mathcal{H}, \pi_{r,s}^{\rightarrow} \in \mathcal{H}$.

The state variables related to vehicle resources include the size of the set \mathcal{V}_g of vehicles used in each line group g , and the sequences of line runs they are assigned to. $\forall v \in \mathcal{V}, \Phi(v) = \{\mathcal{R}\}^n, n \in \mathbb{N}$.

We use r_l^+ (resp. r_v^+) to designate the run following run r on line l (resp. service of vehicle v) and we use r_v^α and r_v^ω to represent the first and last runs vehicle v is assigned to.

2.4. Constraints

Our model considers eight types of constraints which are encountered in a real timetabling and vehicle assignment problem.

- *Stopping time* for each stop and line run is equal to the one provided in the initial timetable.

$$\forall r \in \mathcal{R}, \forall s \in \mathcal{S}^{*r}, \quad \pi_{r,s}^{\leftarrow} + st_{r,s} = \pi_{r,s}^{\rightarrow} \quad (1)$$

- *Running time* for each pair of consecutive stops in a given line run matches the value provided per time period.

$$\forall r \in \mathcal{R}, \forall s, s_r^+ \in \mathcal{S}^{*r}, \quad \pi_{r,s}^{\rightarrow} + rt_{s,s_r^+}^{\pi_{r,s}^{\rightarrow}} = \pi_{r,s_r^+}^{\leftarrow} \quad (2)$$

- The *structure of each run* in terms of sequence of served stops is identical to the one provided in the initial timetable.

$$\forall r \in \mathcal{R}, \forall s \in \mathcal{S}^{*r}, \quad h_{r,s}^{\rightarrow} \in \mathcal{H} \Leftrightarrow \pi_{r,s}^{\rightarrow} \in \mathcal{H} \quad (3)$$

- The *sequences of runs* for each line is identical to the one provided in the initial timetable.

$$\forall l \in \mathcal{L}, \forall r_1, r_2 \in \mathcal{R}_l, \quad (h_{r_1, m_l}^{\rightarrow} > h_{r_2, m_l}^{\rightarrow}) \Leftrightarrow (\pi_{r_1, m_l}^{\rightarrow} > \pi_{r_2, m_l}^{\rightarrow}) \quad (4)$$

- The *line daily bounds* are respected: the first (resp. last) run of a line cannot serve the main stop before (resp. after) the start (resp. end) time of the first (resp. last) headway period of the line.

$$\forall l \in \mathcal{L}, \quad (\pi_{r_l^\alpha, m_l}^- \geq bt_{l,0}) \wedge (\pi_{r_l^\omega, m_l}^- \leq et_{l,|\mathcal{P}_l|}) \quad (5)$$

- In any *vehicle service*, the time gap separating the arrival at the final stop of a line run and the departure from the initial terminal of the next line run is greater or equal to the turnaround time plus the deadheading trip duration, leaving a positive or null remaining waiting time.

$$\forall g \in \mathcal{G}, \forall v \in \mathcal{V}_g, \forall r, r_v^+ \in \Phi(v), \quad \pi_{r_v^+, s_{r_v^+}^\alpha}^- - \pi_{r, s_r^\omega}^- \geq tt_{l_r, s_r^\omega} + dt_{s_r^\omega, s_{r_v^+}^\alpha} \quad (6)$$

- *Interlining* occurs only inside line groups: each vehicle is assigned to line runs belonging to the same line group.

$$\forall v \in \mathcal{V}, \forall r_1, r_2 \in \Phi(v), \quad gl_{r_1} = gl_{r_2} \quad (7)$$

- The *assignment* is complete: a (starting time, vehicle) couple value must be assigned to each line run.

$$\forall r \in \mathcal{R}, \quad r = (\pi \in \mathcal{H}, v \in \mathcal{V}) \quad (8)$$

2.5. Objectives

Our model integrates four realistic objectives that concern both the quality of service and the resource utilization. The associated costs are combined in an aggregated weighted function (see Section 3.2.3).

2.5.1. Number and quality of transfer possibilities

In the literature, the objective of transfer quality is often the straightforward minimization of the total weighted transfer waiting times in the network. This model is simple but fails to render an accurate vision of reality.

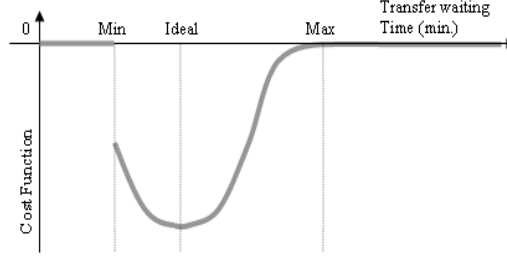


Figure 2: Cost Function for the Transfer Objective
The ideal time gap secures smooth transfers for most passengers.

Here, each transfer is characterized by a minimum, an ideal and a maximum time gap between arrival and departure (see Fleurent et al. (2004)). The minimum gap includes the time needed by the users to reach the departure point and a buffer time to compensate small hypothetical delays. The maximum waiting time defines the limit after which users will stop using the connection.

For this objective, we define a nonlinear cost function of the waiting time that favors the most heavily close-to-ideal waiting times (see Figure 2). The cost is also weighted by the relative importance of the transfer, a value assigned a priori by the expert human planner. We compute the transfer cost for each couple of intersecting line runs. The cost function relative to transfers is always negative, symbolizing the additional value brought to the timetable by each additional transfer opportunity. The function is given by the following formula (Formula 9) where σ is a candidate solution. Recall that il_t are weights associated to the transfers.

$$cost_1(\sigma) = \sum_{\substack{t \in \mathcal{T} \\ t=(l_1, s_1, l_2, s_2)}} \left[il_t * \sum_{\substack{r_1 \in \mathcal{R}_{l_1} \\ r_2 \in \mathcal{R}_{l_2}}} fTr(t, \pi_{r_1, s_1}^{\leftarrow}, \pi_{r_2, s_2}^{\rightarrow}) \right] \quad (9)$$

where

$$fTr(t, h_1, h_2) = \begin{cases} \text{decreasing over } [-0.5; -1] & \text{if } \lfloor wt_t \rfloor < h_2 - h_1 < \lceil wt_t \rceil \\ \text{increasing over } [-1; 0[& \text{if } \lfloor wt_t \rfloor < h_2 - h_1 < \lceil wt_t \rceil \\ 0 & \text{o.w.} \end{cases}$$

Only costs related to intramodal transfers are exposed in this formula for

the sake of simplicity, but the actual $cost_1(\sigma)$ also includes, under the same conditions, the costs concerning intermodal transfers.

2.5.2. Headway evenness

The cost relative to headway evenness can be computed as the sum of the costs generated by all pairs of consecutive line runs' arrival or departure times from the main stop. However, timetables can have complex schemes and all runs might not serve the line's main stop. To deal with this case, we estimate the arrival or departure time of the main stop if this stop were to be included into the itinerary. This permits to consider runs that for instance serve different branches of a line but are overlapping on other parts of the route.

We focus here on the most common case, i.e. when consecutive runs belong to the same headway period. The cost function penalizes the gap between the observed interval and the expected one. It is correlated to the size of this gap and to whether this gap respects the allowed variation margin or not.

$$cost_2(\sigma) = \sum_{l \in \mathcal{L}} \left[\sum_{r, r_l^+ \in \mathcal{R}_l} fHw(l, r, r_l^+) \right] \quad (10)$$

We only present here the most common case of two consecutive line runs serving the main stop during the same headway period $p \in \mathcal{P}_l$. Let gap_{l,r,r_l^+} be the current headway, corresponding to $\pi_{r_l^+, m_l}^{\rightarrow} - \pi_{r, m_l}^{\rightarrow}$.

$$fHw(l, r, r_l^+) = \left[\frac{gap_{l,r,r_l^+} - h_{l,p}}{v_{l,p}} \right]^2 \quad (11)$$

2.5.3. Fleet size

The number of required vehicles represents the main objective to be minimized among those related to the vehicle resources. It is the sum of the number of vehicles assigned to the line runs in each line group.

$$cost_3(\sigma) = \sum_{g \in \mathcal{G}} |\mathcal{V}_g| \quad (12)$$

2.5.4. Length of deadheading trips

Deadheads are the set of trips made by the vehicles between successive line runs. They are to be avoided or minimized since they represent unproductive time and gas consumption.

$$cost_4(\sigma) = \sum_{g \in \mathcal{G}} \sum_{v \in \mathcal{V}_g} \left[\sum_{r, r_v^+ \in \Phi(v)} dt_{s_r^\omega, s_{r_v^+}^\alpha} \right] \quad (13)$$

2.6. Modelization choices

The problem considered in this paper is intended to be realistic and adapted to the planners' needs. For this purpose, commonly available data is used, such as period-dependent travel times and headways, and importance level of transfers. Also, users are not captive and will not use a transfer requiring more than a certain waiting time limit. This implies that both number and quality (waiting time with regard to the ideal value) of transfer opportunities need to be considered. Complex real-world timetable schemes are handled, in which itineraries can vary among runs of the same line. This includes skipping or adding stops in some runs, serving stops in variable order, and lines with branches. Since we consider the perspective of regulating authorities, information regarding the depots is unavailable. The deadheads considered in this paper are thus limited to those linking line runs inside vehicle services.

3. Solution Approach for Timetabling and Vehicle Assignment

3.1. General solution procedure

We base our solution approach on Iterated Local Search (ILS) (Lourenco et al., 2002). Two neighborhoods are alternatively used during the ILS. At each iteration, the timetable is altered and the optimal trip assignment is recomputed.

3.2. Initial solution

Our ILS solution approach needs an initial solution. This initial solution is built in two phases using a linear quasi-assignment model.

- First, a departure time is assigned to each line run of the network based on the existing timetable.

- Second, a vehicle is assigned to each line run through an exact algorithm, i.e. $(v_1, v_2, \dots, v_{|\mathcal{R}|}) = \operatorname{argmin}_{(v'_1, v'_2, \dots, v'_{|\mathcal{R}|}) \in (\mathcal{V} \times \mathcal{R})} (w_3 * \operatorname{cost}_3(\sigma) + w_4 * \operatorname{cost}_4(\sigma))$.

This linkage of runs can be modeled as a network-flow-based linear quasi-assignment problem and solved optimally by an efficient Auction algorithm (Freling et al., 2001). This algorithm consists in assigning trips/source to trips/sink. Half the cost of a vehicle is held by the links between the source and trips and between trips and the sink. Our model differs slightly from Freling et al.'s one in that we do not include in this cost any value related to the deadheading time for pull-in and pull-out trips, since information regarding the depots is unavailable. The links joining trips are assigned a value depending on the feasibility of the connection, deadheading time involved and weight assigned to this objective. The algorithm (*AuctionAlgo*) gives optimal results in a very short time, especially on sparse networks. This fits well our context of limited groups of lines for the vehicle assignment.

Considering that the time part of the assignment is based on the existing timetable, constraints 1 to 5 are satisfied in essence. If we consider that additionally the fleet size is a priori unlimited, constraint 8 is also respected. This variable fleet size combined to our group-by-group exact assignment method insures that the remaining constraints 6 and 7 apply. Therefore, the initial solution is always feasible.

3.2.1. ILS Algorithm for TVA

The Iterated Local Search algorithm for TVA (Algorithm 1) relies on two types of neighborhood relations that alternatively intensify and diversify the search. The elements constituting the ILS algorithm are detailed in the following subsections.

Algorithm - *ILS_for_TVA*

```
 $\sigma \leftarrow \text{GenerateInitialSolution}(t \leftarrow ET, v \leftarrow \text{AuctionAlgo})$   
 $time \leftarrow \text{StartTimer}(), nbIter \leftarrow 0, nbStagn \leftarrow 0$   
 $\sigma \leftarrow \text{Local Search}(\sigma, \text{RunShift})$   
 $best\sigma \leftarrow \sigma$   
repeat  
   $\sigma_1 \leftarrow \text{Perturbation}(\sigma, \text{LineShift})$   
   $\sigma \leftarrow \text{Local Search}(\sigma_1, \text{RunShift})$   
  if ( $f_1(\sigma) \leq f_1(best\sigma)$ )  
     $best\sigma \leftarrow \sigma$   
   $\text{Update}(nbIter, nbStagn)$   
until ( $time > maxTime$  or  $nbIter > maxIter$  or  
 $nbStagn > maxStagn$ )  
return  $\sigma$ 
```

3.2.2. Configuration and search space

A configuration for the TVA problem is constituted by a (time, vehicle) value for each line run (variable). The unconstrained associated search space can be defined by $\mathcal{S} = (\pi \in \mathcal{H}, v \in \mathcal{V})^{|\mathcal{R}|}$, of size $(|\mathcal{V}| * |\mathcal{H}|)^{|\mathcal{R}|}$. However, our algorithm remains inside the feasibility domain, which is considerably smaller, given the set of constraints defined in section 2.4. The two neighborhoods used in the algorithm permit to explore different parts of this search space, *LineShift* exploring a subpart of the space explored by *RunShift* (see section 3.2.4).

3.2.3. Configuration evaluation

For the TVA problem, the evaluation function used by the ILS algorithm consists in the following aggregated weighted function (Formula 14).

$$f_1(\sigma) = w_1 * cost_1(\sigma) + w_2 * cost_2(\sigma) + w_3 * cost_3(\sigma) + w_4 * cost_4(\sigma) \quad (14)$$

Typically, the weights are given by the human planner to reflect the relative importance that he would associate to each objective (cost). Notice that changing the weights will impact the search behavior of the ILS algorithm.

3.2.4. Neighborhoods and moves

The two neighborhoods defined permit to explore the search space while always remaining inside the feasibility domain.

LineShift The *LineShift* move modifies the time value of a definite set of variables and the vehicle values of none to all of the variables. A neighboring solution can be obtained by shifting the departure time of all the runs of one line by $\pm n$ minutes. The vehicle assignment is then completely recomputed on the neighboring timetable using the Auction algorithm (see section 3.2).

LineShift is an original neighborhood in the sense that it impacts both the time and vehicle values of each run, an approach which is, as far as we know, particular to our work. However, it has equivalents on the single aspect of time assignments, and has been used in this partial form to address transfer synchronization problems (see Jansen et al. (2002)).

In most studies, the variables used are the line global departure times. Shifting the departure time of one line is equivalent to shifting the departure times of all its runs. Our choice of variables defines a larger search space by allowing the shifting of individual runs.

RunShift The *RunShift* move modifies the time value of one single variable and the vehicle values of none to all of the variables. A neighboring solution can be obtained by shifting the departure time of one line run by $\pm n$ minutes. The vehicle assignment is then completely recomputed on the neighboring timetable using the Auction algorithm (see section 3.2).

RunShift is an original neighborhood in the same way that *LineShift* is, due to its impact on both types of assignment. Additionally, it has sometimes been suggested in its partial (time only) form as a possibility for future research but not actually implemented in studies we are aware of. In our algorithm, it is used during the intensification phase.

Note that in both neighborhoods, the time shift $n \in \mathbb{N}^*$ is chosen randomly inside a restricted interval defined to respect constraints 3, 4 and 5, and also prevent large shifts, likely to impact the headway evenness objective very negatively. Note also that *RunShift* defines a larger neighborhood than *LineShift* and a *LineShift* move can be considered as a combination of several *RunShift* moves.

The neighborhood associated to *LineShift* is thus smaller than the one associated to *RunShift* in terms of number of neighboring configurations. On the opposite, the distance between neighboring configurations in terms of number of decision variables being assigned a different value, is greater for *LineShift* than for *RunShift*.

3.2.5. Intensification and diversification mechanisms

The ILS algorithm alternates an intensification phase and an diversification phase. The intensification phase is based on the larger (and faster) *RunShift* neighborhood relation. At each iteration of the ILS algorithm, the current solution is repeatedly replaced by a neighboring solution of better quality according to the function defined by Formula 14. This intensification phase stops when no improving neighbor is found after a prefixed number *maxStagn* of consecutive tested moves.

At this point, the diversification is launched, which consists in perturbing the best solution found during the intensification phase. More precisely, we change the neighborhood from *RunShift* to the smaller *LineShift*. Since moves using this neighborhood are likely to substantially modify a solution, only one *LineShift* move is applied in order to avoid excessive deterioration of the solution. This move is selected through a sampling process returning the lowest-cost move among m generated ones. This diversification phase is then followed by a new round of intensification phase that takes the perturbed solution as its input.

3.2.6. Acceptation and stopping criteria

In the ILS for TVA, only moves of non-negative impact are accepted during the intensification phase, while no condition applies during the perturbation phase. The stopping criteria of the whole ILS algorithm rely on computational time *maxTime*, number of iterations *maxIter*, and number of iterations without improvement *maxStagn*.

4. Experimentations and numerical results

In this section, we will assess the performance of our simultaneous approach. For this purpose, we perform experiments on a dataset from a real transit network. For the purpose of comparison, we rely on two references: the existing solution (i.e. the timetable and vehicle services used in the network) and the solution obtained with a conventional sequential approach.

4.1. Experimental settings

Our experimentations are based on a real extra-urban transit network of a large French area involving 3 medium-size cities and numerous villages. The main numerical characteristics of this network on a typical day of operations are summed up in Table 1.

Each line is assigned to one of three different operators, represented here by groups of 8, 16 and 26 lines. External activities include train and school flows. Transfer types correspond to pairs of activities meeting at a stop and thus between which transfers can hypothetically be generated. In our extra-urban context, most of the lines run in a non-continuous pattern, with peaks of activity between 7am and 9am, 12am and 2pm, and 5pm and 7pm. Besides, major lines offer a more complete coverage of the day.

	Lines	Stops	Groups	Runs	Ext. activities	Transfers Types
Quantity	50	673	3	318	30	282

Table 1: Main numerical features of the data set representing a real transit network of a large French area.

Our algorithm was coded in C++, compiled with VC++ 9.0, on a laptop equipped with a 2.8 Ghz Intel(R) Pentium(R) 4 and 1.5Go RAM running Windows XP. For these tests, we allow our ILS algorithm to run 300 minutes of CPU time. This time limit is used in these tests as the unique stopping criterion.

Regarding weights assigned to the objectives, we used the set of values shown in Table 2. These weight values allow us to simulate scenarios where the number of vehicles is considered to be a very important criterion from an economic point of view, and headways and transfers are crucial from a quality perspective. Deadheads are somewhat less important compared to the other objectives. It should be clear other scenarios can be easily simulated with different weighting strategies. The flexibility rendered by this mechanism constitutes a simple, yet desirable feature of the decision-aid tool for the human planner.

	Fleet size ($cost_1$)	Transfers ($cost_2$)	Headway ($cost_3$)	Deadheads ($cost_4$)
Weights	$w_1=1000$	$w_2=50$	$w_3=66.6$	$w_4=1.25$

Table 2: Set of weights used on the tests for each objective. It favors the fleet size predominantly, followed by quality of service (headways and transfers) and leaves deadheads as a subordinate objective.

4.2. References and criteria for comparison

Two solutions are used as our comparison reference: the existing solution (i.e. the current timetable and vehicle services used in the network) and the solution obtained with a conventional sequential approach where the timetabling and vehicle assignment problems are solved successively and separately.

4.2.1. Initial solution used by the transit network

The time values of the existing solution correspond to the currently applied timetable. From this data, we deduce the cost related to the headway evenness objective. This data combined with the real timetables of external activities permit to compute the cost related to intermodal and intramodal transfers.

The vehicle values of the initial solution, on the other hand, do not come from the current schedules, since these data are highly confidential. The initial vehicle assignment is thus computed using the Auction algorithm and results in a fleet size of 91 (lower bound). It is possible that the transit operators use in practice more vehicles than 91, however, fewer is impossible without violating hard constraints.

The detail of these costs can be found for comparison purpose in Tables 3 (Column 3) and 4 (Row 3). In addition to being used as a reference for comparison, the existing solution is also used as initial solution for our ILS algorithm.

4.2.2. Sequential approach

The sequential approach consists in defining the network timetable first, and then only, defining the vehicle services based on this timetable. Concretely here, the model from section 2 is altered so that we solve two distinct assignment problems one after the other.

In a first step, only time values are assigned to line runs. This is achieved through the ILS method relying on *RunShift* and *LineShift* moves restricted to changes of departure times of line runs. Therefore the evaluation of these moves will be much more faster since it is no more necessary to call the Auction algorithm. Consequently, the sequential algorithm permits to generate a much greater number of moves during the same running time.

Only once this first step is achieved, does the second step take place, through the assignment of vehicle values to the now fixed transit lines timetable.

This is done almost instantly through one unique call to the Auction algorithm.

Note that this sequential method, traditionally used in the literature (Ceder and Wilson, 1986), considers, by essence, sets of objectives one after the other. This could be considered to be equivalent to a hierarchical approach where the quality of service objectives would come before the resources ones. This approach opposes our simultaneous approach in which a weighted sum of all the objectives is used as the evaluation function.

4.2.3. Comparison criteria

We have defined two sets of comparison criteria.

- The quality of the solution according to the aggregated function (Formula 14, Section 2).
- The quality of the solution based on each objective: 1) the number of vehicles, 2) the number of transfer opportunities, 3) the number of problems spotted in the headway repartition (i.e. when the observed headway does not fit into the allowed variation margin) and 4) the number of minutes of deadheads.

4.3. Computational Results

For the purpose of simplicity, we will use Initial, SQ and SM to designate respectively the existing solution, the sequential approach and our simultaneous approach.

4.3.1. Simultaneous approach vs. existing solution

Let us first observe the behavior of the simultaneous method on each objective and analyze its performance with respect to the existing solution. Figure 3 presents the evolution of the best weighted cost found for each of the objectives along the search, the starting point corresponding to the values of the existing solution.

It is observed from this figure that at the beginning of the process, sharp cost reductions occur for the headways evenness and number of vehicles objectives. The improvements in the cost related to transfers are less massive at first but continue to steadily impact the global quality of the solution during a longer period. This may be linked to the structure of the evaluation of this objective and the neighborhood mechanism employed, which gives only small power to each move with respect to transfers. Focus has been put on the

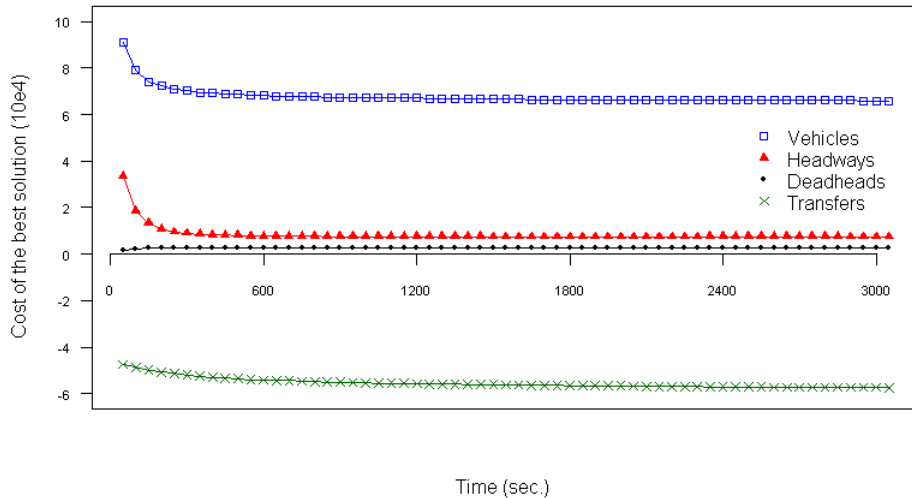


Figure 3: Evolution of the four weighted objective values (see Table 2) of the best solution found by the ILS algorithm. The main objectives ($cost_1$, $cost_2$, $cost_3$) benefit strongly from the algorithm. The time point $t=0s$ corresponds to the initial solution.

beginning of the process, since the improvements are much slower afterwards. Indeed this figure also shows that the algorithm can be stopped much earlier than the time limit used here ($t=300min$), and still provide great improvement. For instance, at $t=50min$, the algorithm has achieved, on average, 96.6% of the improvement provided by the longer ($t=300min$) runs. Notice that the evolution of each curve can be influenced through the tuning of the objectives' weights according to the human planner's needs.

Table 3 presents in more details the results of our simultaneous approach (Column 5, SM) in comparison with the initial solution (Column 3, Initial) according to the aggregated cost function f_1 (Row 1) together with the individual weighted costs. In this Table, we show the average, minimal, maximal value and standard deviation on 20 independent runs. Similarly, Table 4 focuses on the second comparison criterion according to each of the four objectives and presents the results found in the Initial solution (Row 3, Initial) and those obtained by the simultaneous approach (Row 5, SM).

From these tables, we can observe that the simultaneous approach provides significant improvement compared with the existing solution on all but

the last (and least important) objective. Both quality of service and level of resources strongly benefit from the ILS for TVA. Most notably, the number of vehicles drops by 28.35%, the number of transfers increases by 112.33%, while the number of observed headways not fitting into the allowed variation margin plummets by 75.72%. Only the value of the deadheads objective deteriorates due to its status of least important one, reflected in the weights repartition.

		Initial	Sequential ILS (SQ)	Simultaneous ILS (SM)
Cost	Avg	78763.8	27420.0	16732.5
$f_1(\sigma)$	Min	-	25478.1	14869.9
	Max	-	29794.3	18441.6
	StD	-	1262.7	991.0
	Avg	91000.0	78706.0	65150.0
Vehicles $w_3 * cost_3(\sigma)$	Min	-	77000.0	62000.0
	Max	-	81000.0	67000.0
	StD	-	1213.0	1226.0
	Avg	-47373.1	-62648.7	-58716.6
Transfers $w_1 * cost_1(\sigma)$	Min	-	-64050.0	-61136.4
	Max	-	-60606.3	-55732.1
	StD	-	883.6	1428.6
	Avg	33526.9	8420.1	7469.7
Headway $w_2 * cost_2(\sigma)$	Min	-	7551.1	6345.3
	Max	-	9206.8	8307.9
	StD	-	434.7	414.0
	Avg	1610.0	2942.7	2829.4
Deadheads (internal) $w_4 * cost_4(\sigma)$	Min	-	2603.8	2462.5
	Max	-	3138.8	3218.8
	StD	-	154.4	203.1

Table 3: Comparative results of the existing solution (Initial), sequential approach (SQ) and the simultaneous approach (SM) according to the first type of comparison criteria: cost function f_1 and individual weighted costs (see Table 2). While both SM and SQ methods achieve important improvements compared with the initial situation, the simultaneous method clearly performs much better than the sequential one.

	Vehicles				Transfers				Headway				Deadheads			
	Avg	Min	Max	StD	Avg	Min	Max	StD	Avg	Min	Max	StD	Avg	Min	Max	StD
Initial	91.0	-	-	-	180.0	-	-	-	131.0	-	-	-	1288.0	-	-	-
SQ	78.7	77	81	1.2	398.4	352	428	21.9	34.9	28	39	2.9	2354.2	2083	2511	123.5
SM	65.2	62	67	1.2	382.2	361	425	17.7	31.8	23	39	3.4	2263.5	1970	2575	162.5

Table 4: Comparative results of the existing solution (Initial), sequential approach (SQ) and the simultaneous approach (SM) according to the second type of comparison criteria. The simultaneous approach performs much better than the sequential one, notably reducing the fleet size by 13.5 vehicles on average.

4.3.2. Simultaneous approach vs. sequential approach

We now turn our attention to the comparison between the results of SM and SQ by considering the results of Tables 3 and 4. From these tables, we can make one main observation: the simultaneous approach based on the *RunShift* and *LineShift* neighborhoods outperforms the sequential approach and permits to drastically reduce the size of the fleet size while sensibly improving the quality of service. Let us analyze this difference with more details.

Table 3 shows that on average, the best aggregated cost generated by SM is $f_1 = 16732.5$, while SQ attains 27420.0. This demonstrates the superiority of the simultaneous approach towards the sequential method. Using the second comparison criteria (see Table 4), the results obtained by SM are 14.84% better than those obtained by SQ on the number of vehicles, 2.37% better on the number of headway irregularities, 7.04% better on the length of deadheads, and 9% worse on the number of transfer opportunities. This is remarkable given that both SQ and SM start from the same initial solution.

As predictable, we can see that in SQ, quality of service is strongly favored compared to the level of resources. Indeed, during the first assignment, no resource considerations are taken into account, and thus fewer constraints prevent the algorithm from creating and improving transfer opportunities. This explains the better results on the transfer objective for SQ than for SM. Once the transit timetable is fixed however, the fleet size is also implicitly fixed, denying it its own relative weight in the objective function. The simultaneous approach permits to obtain a significantly smaller number of needed vehicles: 65.2 instead of 78.7 for SQ (i.e. a reduction of 17.15%).

To conclude this section, we observe that the global quality (with respect to the cost function) of the solution provided is far better with the simultaneous than with the sequential approach, due to the fact that the planner can balance at will the set of objectives, while in the sequential case, the fleet size is reduced to being a consequence of the decisions made separately and previously on the quality of service objectives.

5. Conclusion

The problem treated in this paper constitutes a combination of steps often treated sequentially in the transit planning process: timetabling and vehicle assignment. These problems, although closely inter-related, have been rarely considered simultaneously in the literature.

We described a flexible model that includes important features of a real transit network by considering a number of constraints and objectives relative to both quality of service (number and quality of transfers, headway evenness) and resource utilization (number of vehicles deployed, deadheading trips). Based on this rich model, we devised an Iterated Local Search algorithm which relies on two original neighborhoods and is combined with an Auction algorithm.

We presented simulation results on a real transit network of a large French area and compared the results with both the solution currently applied in the network and the solution obtained with a conventional sequential method. These experiments showed evidence that the simultaneous approach is a valuable alternative to the sequential method and permits to obtain improved timetables and a reduced number of required vehicles.

The developed algorithm has been integrated into a commercial decision-support system that is currently used by both urban and extra-urban transit network planners. The commercial product includes additional constraints such as fixed departure time for some line runs and mandatory transfer opportunities. It also permits to create completely new timetables without initial existing timetable for some or all of the lines. It is used to prepare simulations and proposals using different sets of weights for the objectives and different sets of constraints. The product also allows the planner to make any modification needed to the generated output. Such a flexibility is crucial in preparation of calls for tender and discussions with transit operators. Instead of manually designing only one or two scenarios as it is often still the case, planners can use the decision-support system to quickly and efficiently generate, evaluate and compare varieties of new timetables and vehicle schedules.

Despite the great complexity of the underlying problem, the study shows that a simultaneous approach fits the planners' needs and permits to obtain solutions of better quality on the classical criteria. A path for future work could be the integration to the solution method of additional features that belong to steps both upstream (e.g. frequencies setting) and downstream (e.g. depot assignment, crew scheduling) in the transit planning process.

Acknowledgements

This work was partially supported by the French Ministry for Research and Education through a CIFRE contract (number 400/2005). The reviewers of

the paper are greatly acknowledged for their helpful comments.

References

- Bookbinder, J.H. & A. Désilets. (1992). Transfer optimization in a transit network. *Transportation Science*, 26, 106–118.
- Castelli, L., R. Pesenti & W. Ukovich. (2004) Scheduling multimodal transportation systems. *European Journal of Operational Research*, 155, 603–615.
- Ceder, A. & N.M.H. Wilson. (1986). Bus network design. *Transportation Research B*, 20, 331–344.
- Ceder, A., B. Golany & O. Tal. (2001). Creating bus timetables with maximal synchronization. *Transportation Research A*, 35, 913–928.
- Cevallos, F. & F. Zhao. (2006). Minimizing Transfer Times in a Public Transit Network with a Genetic Algorithm. *Transportation Research Record*, 1971, 74–79.
- Chakroborty, P., K. Deb & R.K. Sharma. (2001). Optimal fleet size distribution and scheduling of transit systems using genetic algorithms. *Transportation Planning and Technology*, 24(3), 209–226.
- Daduna, J.R. & S. Voss. (1995). Practical experiences in schedule synchronization. In: J.R. Daduna, I. Branco & J.P.M. Paixao (Eds), *Computer-Aided Scheduling of Public Transport, Lecture Notes in Economics and Mathematical Systems LNEMS*, vol. 430 (pp. 39–55). Berlin: Springer.
- Fleurent, C., R. Lessard & L. Seguin. (2004). Transit timetable synchronization: Evaluation and optimization. *Proceedings of the 9th International Conference on Computer-Aided Scheduling of Public Transport (CASPT)*. San Diego, California.
- Freling, R., A. Wagelmans & J. Paixao. (2001). Models and algorithms for single-depot vehicle scheduling. *Transportation Science*, 35(2), 165–180.
- Guihaire, V. & J.K. Hao. (2008). Transit network design and scheduling: a global review. *Transportation Research A*, 42, 1251–1273.

- Hadjar, A., O. Marcotte & F. Soumis. (2006). A branch-and-cut algorithm for the multiple depot vehicle scheduling problem. *Operations Research*, 54(1), 130–149.
- Huisman, D., R. Freling & A.P.M. Wagelmans. (2005). Multiple-Depot Integrated Vehicle and Crew Scheduling. *Transportation Science*, 39, 491–502.
- Jansen, L.N., M.B. Pedersen & O.A. Nielsen. (2002). Minimizing passenger transfer times in public transport timetables. In: *Proceedings of the 7th Conference of the Hong Kong society for Transportation Studies: Transportation in the Information Age* (pp. 229–239). Hong Kong.
- Laurent, B. & J.K. Hao. (2007). Simultaneous vehicle and driver scheduling: a case study in a limousine rental company. *Computers & Industrial Engineering*, 53(3), 542–558.
- Laurent, B. & J.K. Hao. (2008). Simultaneous vehicle and crew scheduling for extra urban transports. In: Nguyen, N.T., L. Borzemski, A. Grzech & M. Ali (Eds), *New Frontiers in Applied Artificial Intelligence, Lecture Notes in Artificial Intelligence 5027* (pp. 466–475). Berlin: Springer Verlag.
- Laurent, B. & J.K. Hao. (2009). Iterated Local Search for the Multiple Depot Vehicle Scheduling Problem. *Computers & Industrial Engineering*, 57(1), 277–286.
- Laurent, B. & J.K. Hao. (2009). List Graph Coloring for Multi-depot Vehicle Scheduling. *International Journal of Mathematics in Operational Research*, 1(1-2), 228–245.
- Lourenco, H.R., O.C. Martin & T. Stutzle. (2002). Iterated Local Search. In: Glover F. & G. Kochenberger (Eds), *Handbook of Metaheuristics, ISORMS 57* (pp. 321–353). Norwell: Kluwer Academic Publishers.
- Martins, C.L. & M. Vaz Pato. (1998). Search strategies for the feeder bus network design problem. *European Journal of Operational Research*, 106, 425–440.
- Pepin, A.S., G. Desaulniers, A. Hertz & D. Huisman. (2009). A comparison of five heuristics for the multiple depot vehicle scheduling problem. *Journal of Scheduling*, 12(1), 17–30.

- Wong, R.C.W. & J.M.Y. Leung. (2004). Timetable Synchronization for Mass Transit Railway. In: Proceedings of the 9th International Conference on Computer-aided Scheduling of Public Transport (CASPT), San Diego, California.
- Zhao, F. & X. Zeng. (2006). Simulated annealing-genetic algorithm for transit network optimization. *Journal of Computing in Civil Engineering*, 20(1), 57–68.
- Zhao, F. & X. Zeng. (2008). Optimization of transit route network, vehicle headways and timetables for large-scale transit networks. *European Journal of Operational Research*, 186, 841–855.