Transit Network Re-Timetabling and Vehicle Scheduling

Valérie Guihaire[1,2] and Jin-Kao Hao[2]

[1] PERINFO, 41 avenue Jean Jaurès, 67100 Strasbourg, France
`vguihaire@perinfo.com`
[2] LERIA, Université d'Angers, 2 boulevard Lavoisier, 49045 Angers, France
`hao@info.univ-angers.fr`

Abstract. In the transit planning literature, network timetabling and vehicle scheduling are usually treated in a sequential manner. In this paper, we focus on combining important features of these two steps, and underline how their simultaneous optimization is meaningful and can bring important improvements to both quality of service and level of resources required. We deal with the objectives of networkwide quality of service through number and quality of the transfers and evenness of the line headways, and with the resources side through number of vehicles needed. Our approach is adapted to the problem faced by regulating authorities, treating among others intermodality, multi-periods for headways and travel times, and complex timetable schemes. We introduce an optimization procedure based on Iterated Local Search and present computational experiments carried out on a real large transit network, showing substantial improvements in both quality of service and level of resources compared to the current practice.

Key words: Mass transit, Timetabling, Scheduling, Transfer Synchronization, Iterated Local Search

1 Introduction

Transit network timetabling is the step of transit planning during which the quality of service of the offer is determined and the level of resources needed is strongly influenced (lower-bounded). In general, the offer is defined first, to create transfer possibilities and respect headway bounds. Vehicle allocation only occurs afterwards, thoroughly constraining the problem of minimizing the number of buses needed. Despite the strong relation between these two problems, most studies focus on a single side, due to the intrinsic complexity of each of them.

Depending on the focus of the problem treated, denominations for transit network timetabling include Schedule Synchronization [6], Transfer Time Optimization [3] or Transfer Coordination [12]. Objectives assigned to these problems include minimizing total waiting time, minimizing transfer waiting time or maximizing the number of simultaneous arrivals [9]. This problem has often been modeled as a Quadratic Semi-Assignment Problem (QSAP) [6, 1] which aims at minimizing the global transfer waiting time of passengers in the network by setting the first departure time of each line. However, QSAP cannot capture some important operational constraints such as variable headways. The author of [11] proposed a constructive heuristic to set line runs departure times. The objective was to minimize the total transfer waiting time while allowing variable headways between runs. However, the evenness of headways was not considered as an objective, neglecting an important factor in service quality. In many studies, headway evenness is rather computed as a function of initial waiting time weighted by the number of users waiting [7]. This implies availability of the desired boarding times of the passengers or the assumption of arrival time distribution functions.

In order to include additional degrees of freedom, the authors of [4] proposed a Non-Linear Mixed Integer Problem model in which stopping times are allowed to vary. However, due to the level of complexity involved, this study considered a single transfer node. In [7], the authors used a Genetic Algorithm on a networkwide basis. A single common period is considered for the whole network, meaning fixed running times and unvarying headway demand for each route.

Vehicle scheduling, 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 or fleet homogeneity/heterogeneity [5]. In the case of regulating authorities, information regarding depots and fleet are unavailable.

So far, the simultaneous approach of optimal fleet distribution and timetabling of transit systems has only been superficially explored. It has often been narrowed to integrating constraints on the number of available vehicles in the timetabling problem and considering schedules without interlining. The first study that we are aware of that considers the number of vehicles as an objective of the transit network timetabling problem is reported in [4]. The authors proposed a genetic algorithm to tackle a combination of transfer coordination and vehicle scheduling problems. However, since the representation is cumbersome, the restrictive case of a single transfer stop with multiple lines is studied.

In this paper, we consider the combination of transit network timetabling with vehicle scheduling from the point of view of regulating authorities, meaning the consideration of depots is not needed yet. Given a pre-defined lines network; the current timetable; groups of lines sharing resources; running times; headway periods; and levels of importance of the transfers; the goal is to define a synchronized network timetable and the associated vehicle assignment with respect to a set of constraints and objectives. Our approach is based on three different levels of evaluation: headway evenness is calculated per line, level of resources is computed per group of lines, and transfer optimization works at the network level. We present a solution method combining both perspectives, leading to a very flexible decision-aid tool. An Iterated Local Search procedure is developed, which is based on the exploration of two types of neighborhoods aiming at alternatively intensifying and diversifying the search.

2 Problem Description

Our problem consists in assigning a departure time and a vehicle to each line run in the network, with respect to a set of objectives and constraints. In this part, we detail the characteristics, variables and domains, inputs, constraints and objectives of our approach. Let us state a few definitions first:

- A *route* is a sequence of stops, and a line is a route with a direction. In the rest of this paper, we will consider only lines.
- A *line run* is a trip on the line, characterized by a departure time.
- An *external line* is any activity connected in time and space with the transit network (e.g. a train line, a factory quitting). This information is used to create intermodal transfers.
- 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 accumulated with respect to the planned schedule.
- 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.
- 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 run.
- A *vehicle assignment* is the entire sequence of line runs assigned to a vehicle.

2.1 Properties of our Approach

Our model includes a set of interesting properties that render it particularly flexible, realistic and adapted to planners' needs. It aims at defining a compromise between flexibility and complexity for the design of high quality timetables. For this purpose, we consider period-dependent travel times and headways, while keeping fixed stopping times in stations.

We also base our model on realistically available data such as planner-defined importance level of transfers and period-dependent headways. Indeed, while origin-destination data is often taken as input for the models in the literature, the complexity induced by path-assignment is such that the option of re-routing passengers during the process of optimization is usually abandoned and the demand considered fixed and inelastic. Therefore, we directly assign a weight to each transfer based on the transit operator's knowledge and experience.

Also, we consider that users are not captive and will not use a transfer requiring more than a certain waiting time limit. This implies that we need to take into consideration both number and quality (waiting time with respect to provided minimum, ideal and maximum waiting times) of transfers.

Our model supports complex line timetable schemes and uses real-world timetables, in which itineraries can vary with the line runs. This includes skipping or adding stops to the "main" line itinerary in some runs, serving stops in variable order, and lines with branches (when several patterns of itinerary are used, serving different stops, usually at extremities of the line).

2.2 Input

The problem considered in this paper has the following inputs.

- The lines network structure, composed of ordered sets of stops with a planner-defined main stop for each line.
- The current timetable (it is assumed the lines are already in use), comprising for each line run the set of arrival and departure times of the served stops.
- For each line, a set of headway periods with associated variation margins.
- The constitution of groups of lines on which interlining is allowed. Each vehicle will be assigned to line runs exclusively inside the same lines group. This allows to cover the cases in which portions of various sizes of the network are serviced by the same operator, as well as the common urban case of vehicles running back and forth on a single line.
- All needed information concerning external lines are also available, namely connecting point and times with the network.
- The set of parameterized transfers. Parameters include relative level of importance, and minimum, ideal and maximum waiting time for users.
- The deadhead running times between all line run termini of each lines group, which are used in the vehicle assignment part of the problem.

2.3 Variables

The set of *decision variables* is composed of all the line runs on the timetable. The value to be assigned to each variable is a (starting-time, vehicle) pair. The domain of these variables is discrete and finite. The planning horizon is comprised in a time-frame usually of one day long, and discretised in minutes without loss of generality. The number of vehicles is at most equal to the number of line runs in the timetable.

Since we use fixed stopping times and the period-dependent running times are also given, we can then deduce all arrival and departure times for the rest of the stops from the starting time of each line run.

2.4 Constraints

Feasibility Constraints A feasible solution must meet three conditions:

- The stopping time at each stop is equal to the initial stopping time at the same stop of the same line run.
- The running times between stops match the period-dependent data.
- In any vehicle schedule, the time gap separating the arrival at the last stop of a line run and the departure from the first stop of the next line run assigned to it must be greater or equal to the turnaround time of the terminal plus the deadheading trip duration.

Timetable Structure Within a given line, the set of stops served and their order can vary with the runs along the day. The structure of each run is fixed, and the order in which the runs are served cannot be modified. Additionally, 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.

Complete Assignments A (starting-time, vehicle) couple value must be assigned to each decision variable. In order to be consistent, we require the assignment to be complete such that:

- A departure time must be assigned to each line run.
- One vehicle must be assigned to each line run.

Group Interlining All resources must be assigned to line runs belonging to the same line group.

2.5 Objectives

Fleet Size The number of vehicles is the main resource objective.

Number and Quality of Transfer Possibilities As mentioned in 2.1, we base our transfer quality evaluation on bounds and ideal value provided by the planner. The cost function we use is a nonlinear function of the waiting time which favors the most heavily close-to-ideal waiting times. The cost incurred to the configuration is also pondered by the relative level of importance of the transfer. The time gaps between arriving and departing runs belonging to lines meeting in the network are computed. Each gap belonging to the allowed interval means a transfer opportunity and generates a negative cost to the configuration.

Headway Evenness In the context of a minimization problem, we model this objective as one of minimizing the sum of evenness defaults. Dealing with multiple headway periods and their transitions is a challenging problem [2]. Let us describe the basics of our method to deal with the three types of situations which can arise (see Fig.1):

- *Case 1: Consecutive runs belong to the same headway period.* The observed interval should be as close as possible to the expected headway.
- *Case 2: Consecutive runs belong to adjacent headway periods.* Only gaps shorter than both or longer than one of the expected headways are penalized.
- *Case 3: Observed interval with start and end of the day.* For each line, two of these intervals occur: between the allowed start of the day and the first actual run, and between the last run and the end of the day. It is assumed that these values can be "too long" but not "too short", and we compare them to the variation margin.

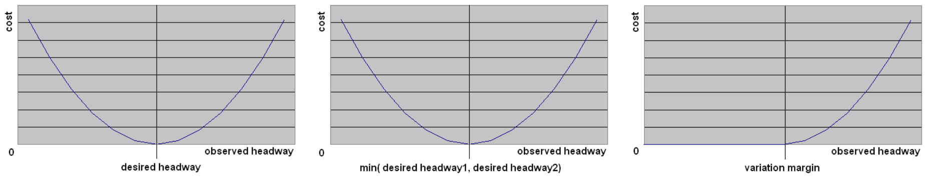

Fig. 1. Individual headway cost functions incurred respectively for case 1, 2 and 3

3 Solution Approach

Our transit network timetabling and vehicle scheduling problem enables several important features: complex timetable schemes, multiple headway periods for each line and variable running times along the day. These features induce at the same time additional difficulties for solving the problem. Given the intrinsic complexity of the model, we choose to employ a heuristic solution approach rather than exact methods. For this purpose, we use Iterated Local Search for its simplicity and efficiency.

In what follows, we recall the basic idea of ILS and the algorithm used for the vehicle assignment, then introduce the evaluation function and the neighborhoods employed, and from there we present the developed approach.

3.1 Iterated Local Search

Iterated Local Search (ILS) is a simple and robust metaheuristic [10]. It is based on the principles of Local Search combined with perturbation movements that are used when the search is caught in a local optimum. If this perturbation satisfies a given acceptance criterion, another round of local search is applied to the current solution, eventually leading to another local optimum. These alternate phases of intensification and diversification permit an exploration of the local optima of the search space that can provide effective results. In our case, we use "Markovian" walk dynamics rather than a history. This choice is based on the fact that the defined perturbation movements diversifies the search enough to limit the probability of leading back to the last local optimum.

3.2 Vehicle Assignment

The vehicle assignment part of our problem, i.e. the linkage of runs, can be modeled as a network-flow-based *quasi-linear assignment problem* and solved optimally by an efficient auction algorithm [8]. This algorithm consists in assigning the source and the trips to trips or to the sink (in order to create sequences). We assign half the cost of a vehicle to the links between the source and trips and between trips and the sink. The cost is somewhat different from Freling's model of [8] 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. Therefore all the links leaving the source and entering the sink have the same value. The other ”intermediate” links, joining two trips, are assigned a value depending on the feasibility of the connection. The algorithm 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.

3.3 Evaluation Function

An aggregated weighted sum is used to evaluate the global cost of a solution. While the number of vehicles involved and number of feasible transfers are countable, notions of transfer quality and headway evenness are more fuzzy. Non-linear cost functions are defined for these objectives (see 2.5). This evaluation function is used in the context of our minimization problem.

3.4 Moves and Neighborhoods

RunShift: Recall that the value of a variable is composed of a (starting-time, vehicle) couple. The *RunShift* move modifies only the time value of one single variable and the vehicle values of none to all of the variables.

Given a solution (i.e. a timetable and a schedule), a neighboring solution can be obtained by shifting the departure time of a single randomly selected line run by +/- n minutes. n is chosen randomly inside a restricted interval defined to both respect the timetable structure constraint and prevent large shifts that are likely to incur high costs on the headway evenness objective.

The vehicle assignment is then recomputed on the neighboring timetable, resulting in the possible modification of the vehicle value of up to the entire set of variables. These values are determined in order for the schedule to remain inside the realisability domain by respecting the sequence feasibility and group interlining constraints, and also to be cost-optimal on the vehicle assignment problem.

LineShift: The *LineShift* move modifies the time value of a definite set of variables at a time 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 single randomly selected line by +/- n minutes. As in *RunShift*, n is chosen randomly inside a restricted interval defined to both respect the timetable structure constraint and prevent excessively large shifts that could be detrimental to the headway evenness objective.

The vehicle assignment is then completely recomputed on the neighbor timetable.

3.5 Initial Solution and General Procedure

Our ILS algorithm is based on a heuristic initial construction and uses two neighborhoods aiming at alternatively intensifying and diversifying the search.

The initial solution is built according to the following three steps.

- 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 the linear assignment algorithm (see section 3.2).
- Third, a descent method combined with the *LineShift* neighborhood is applied to explore quickly parts of a restricted but diverse search space.

It should be clear that this initial solution corresponds to a local optimum with respect to the small-sized *LineShift* neighborhood. At this point, we use a larger neighborhood (*RunShift*) to intensify the search in the vicinity of this local optimum. Additional areas (that were not reachable with *LineShift*) of the search space (which is now complete) can be explored through *RunShift* moves. This is the heart of the ILS: at each iteration, a descent method is applied with *RunShift* up to a local optimum, at which point a perturbation is applied.

This perturbation consists of a short sequence of *LineShift* moves applied in the context of a descent method. Moves from *LineShift* are likely to modify substantially the current solution. Therefore, in order not to loose too many of the good properties acquired so far, only moves with negative or null impact on the evaluation function are accepted. The descent is stopped when a number of moves have been accepted or evaluated, or a time limit has been reached.

The acceptation criterion is met when the quality of the current solution is equal to or better than the quality of the best local optimum recorded so far. The stopping criterion of the whole ILS algorithm relies on computational time, number of iterations, and number of iterations without improvement.

4 Experimentations and Numerical Results

4.1 Data Set

Our experimentations are based on a real extraurban transit network of a large French area involving 3 medium-size cities and numerous villages. The network is composed of 50 oriented lines serving 673 stops. Each line is assigned to one of three different operators, represented in our model by line groups (of 8, 16 and 26 lines). On the typical day of operation considered in this study, 318 line runs are scheduled. Additionally, 30 external activities (train and school flows) interact with the bus network. Considering only the spatial structure of the transit network, 282 different kinds of intramodal and intermodal transfers can hypothetically be generated.

4.2 Computational Results

For this particular test, we carried out 10 runs on the transit network instance, allowing for each run 10 minutes of CPU time (corresponding to a reasonable time cutoff on real situations). Our algorithm was coded in C++, compiled with VC++ 9.0, on a laptop equipped with a 1.73 Ghz Intel(R) Pentium(R) M and 1Gb RAM running Windows XP. Averaged results are plotted in Fig. 2 and 3.

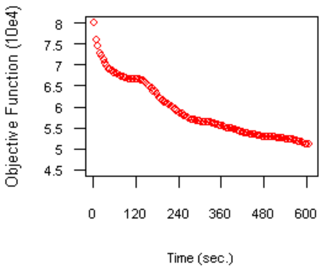

Fig. 2. Evaluation Function Evolution

The first part of the algorithm, based on $LineShift$, stops on average at 130s. We can observe on Fig. 2 that in this short initial period, the descent method based on this neighborhood provides drastic improvement to the solution very fast, and stabilizes at a local optimum. In the second phase consisting of the heart of the ILS, the additional level of detail provided by $RunShift$ permits to explore new areas of the search space and find (important) improvements therein.

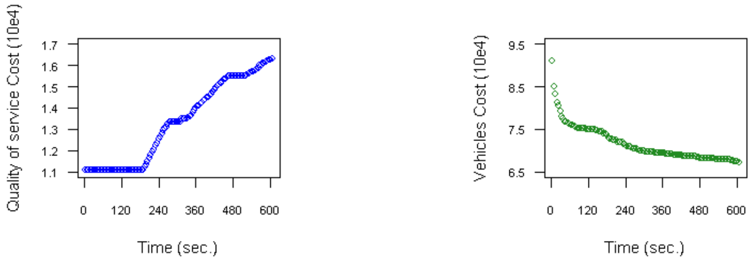

Fig. 3. Profile of the best solution according to the two sides of the problem

Fig. 3 discloses more details of the result. It is easily observed that the algorithm substantially improves both the quality of service (left) and the level of resources (right). The number of vehicles is evaluated at 91 with the current timetable, and reduced to a average of 67 by the algorithm. The number of feasible transfers on the other hand rises from 180 initially to 260 on average. Considering transfer quality and headway evenness, the global cost function related to quality of service is provided in the left graph of Fig. 3. We use a set of parameters that, while considering both aspects of the problem, slightly favors the resources, as is often desired by the regulating authorities. This is why the number of vehicles drops first, and the quality of service only rises afterwards.

5 Conclusion

The problem treated in this paper combines simultaneously two important steps usually treated sequentially in the transit planning process: timetabling and vehicle scheduling. Such a simultaneous approach is, to our knowledge, the first of this kind in the context of this study.

We introduced a natural and high level model in which a departure time and a vehicle must be assigned to each line run of the timetable. This model has the main advantage of being flexible, able to embrace various relevant features of real-world transit networks. In particular, we considered objectives and constraints concerning both quality of service and level of resources as well as other practical features.

A local search optimization procedure is proposed, combining two types of neighborhood in a descent-based Iterated Local Search method. A linear auction algorithm is used to assign vehicles to line runs. Tests carried out on a real and large network showed considerable improvements in both quality of service and level of resources required compared with the current practice. The algorithm has been integrated in a commercial software solution designed for transit operators, and is being used for re-timetabling and scheduling projects by regulating authorities. A path for future work is to integrate features to the solution method that belong to steps both forward and backward in the transit planning process.

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

References

1. Bookbinder, J.H., Désilets, A.: Transfer Optimization in a Transit Network. Transportation Science 26, 106–118 (1992)
2. Ceder, A.: Public Transport Timetabling and Vehicle Scheduling. In: Lam, W., Bell, M. (eds.) Advanced Modeling for Transit Operations and Service Planning. pp. 31-57. Elsevier Science Ltd, New York (2003)
3. Cevallos, F., Zhao, F.: Minimizing Transfer Times in a Public Transit Network with a Genetic Algorithm. Transportation Research Record 1971, 74–79 (2006)
4. Chakroborty, P., Deb, K., Sharma, R.K.: Optimal Fleet Size Distribution and Scheduling of Transit Systems using Genetic Algorithms. Transportation Planning and Technology 24(3), 209–226 (2001)
5. Daduna, J.R., Paixao, J.M.P.: Vehicle Scheduling for Public Mass Transit - an Overview. In: Daduna, J.R., Branco, I., Paixao, J.M.P. (eds.) Computer-Aided Transit Scheduling. pp. 76-90. Springer, Berlin (1995)
6. Daduna, J.R., Voss, S.: Practical Experiences in Schedule Synchronization. In: Daduna, J.R., Branco, I., Paixao, J.M.P. (eds.) Computer-Aided Scheduling of Public Transport. LNEMS, vol. 430, pp. 39-55. Springer, Berlin (1995)
7. Deb, K., Chakroborty, P.: Time Scheduling of Transit Systems with Transfer Considerations using Genetic Algorithms. Evolutionary Computation 6(1), 1–24 (1998)
8. Freling, R., Wagelmans, A., Paixao, J.: Models and Algorithms for Single-Depot Vehicle Scheduling. Transportation Science 35(2), 165–180 (2001)
9. Guihaire, V., Hao, J.K.: Transit Network Design and Scheduling: a Global Review. Transportation Research A (to appear)
10. Lourenco, H.R., Martin, O.C., Stutzle, T.: Iterated Local Search. In: Glover, F., Kochenberg, G. (eds.) Handbook of Metaheuristics. pp. 321-353. Kluwer Academic Publishers, Norwell (2002)
11. Quak, C.B.: Bus Line Planning. Master's thesis. Delft University of Techology, The Netherlands (2003)
12. Ting, C.J.: Transfer Coordination in Transit Networks. PhD Thesis, University of Maryland, College Park (1997)