Efficient decomposition method for the stochastic optimization of public transport schedules

Sofia Zaourar-Michel
Xerox Research Centre Europe
sofia.michel@xrce.xerox.com

Abstract

We propose a new method to optimize public transport schedules by minimizing the waiting time during transfers. Using ticket validation data to construct a realistic scenario-based model of the waiting times, our goal is to design shifts of the current schedules that reduce the overall expected waiting time. We propose a parallel local search heuristic that exploits the structure of the problem to efficiently explore a large number of possible schedules. Compared to previous approaches, our algorithm should allow to both treat more transfers (bigger cities) and more scenarios (ensuring a better generalization). We provide promising preliminary results on transit data collected from Nancy, France.

1 Introduction

Public transport planning is a complex process that has a significant impact on the overall quality of the service. A key step in this process is to design timetables that take into account the transportation resources, the law and regulations and the specific constraints of the city while satisfying the demand and ensuring a smooth experience for the passengers. A major weakness of public transit, as perceived by passengers is the time lost during transfers. This latter can be significant: for example it represents an average of 23% of travel time for multi-modal trips in the UK [4]. Studies have shown that this transfer waiting time inherent to the system is poorly perceived by the passengers [1, 11].

In the Operations Research community, numerous works have been done in the domain of schedule synchronization. The goal is to coordinate the timetables of the different lines in order to minimize the waiting time at the connections (see e.g. the surveys [6, 10]). Most of the Operations Research approaches focus on the theoretical timetables and give the same importance to all the connections at all times of the day. If we observe the real usage of the system however, it is clear that some connections are more important than others in terms of the frequency and volume of passengers. That is why transportation authorities often incorporate their expert knowledge of the system, as well as experience's rules of thumbs, to design the timetables. However, with the constant growth and sophistication of public transport, these approaches reach their limit.

A natural alternative is to use the data that is generated by the transportation system to precisely quantify and model transfer waiting times while accounting for the stochasticity of the system. Among the first data-driven approaches to schedule optimization, [9] used the queries to an on-line trip planner to approximate the real usage of the system in order to build a two-stage stochastic linear program (LP) (see e.g. [2]). The idea was to compute shifts of the schedules that minimize the expected waiting times across a number of scenarios. A recent extension [8] proposed to use the real transit data taking into account the fact that passengers optimize their transfers according to the schedules. This led to a two-stage stochastic linear program with mixed-integer variables (MILP).

While these approaches give finer and more realistic models, they inevitably lead to very large-scale problems. In [8], only a sub-network was solved using an open-source MILP solver. Tackling an

30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain.

entire city with different modes and a substantial number of scenarios would be out of reach of standard optimization solvers.

In this paper, we propose a new formulation of the stochastic waiting time minimization problem as a quadratic mixed-integer program (MIQP). Although theoretically less appealing than the MILP models, this formulation allows us to exhibit an interesting separability property that we exploit to design our optimization algorithm. The basic idea is that for a given modification of the schedules, the evaluation of its impact on the waiting times can be done in parallel for the different scenarios. Even within a scenario, we show that the evaluation of certain aggregated transfers is separable and can be done in parallel. Moreover, we show that the optimization subproblem that needs to be solved for the evaluation has a closed-form solution, making the evaluation much more efficient. We take advantage of these properties to design an efficient parallel local search algorithm that explores random modifications of the current schedules looking for ones that reduce the waiting time. Although there are no optimality guarantees, we observe in our experiments that the algorithm quickly leads to a significant reduction of the expected waiting time.

The remainder of the paper is organized as follows. Section 2 describes the stochastic program for the waiting time optimization. Section 3 describes the separability property and the special structure of the evaluation subproblem and presents the local search heuristic that we propose. Finally Section 4 reports preliminary results of applying our algorithm to the city of Nancy, France.

2 Minimization of the transfer waiting time

2.1 Definitions and assumptions

We assume working with schedule-based public transport system, whose main components are: a set of nodes $\mathcal{N}$, a node being a stop or an aggregation of stops that allow connections from one to another within walking distance; a set of routes $\mathcal{R}$, a route being a transit service serving a series of stops and is composed of a collection of trips $r = \{p_1, p_2, \ldots, p_n\}$; a set of trips $\mathcal{P} := \cup_{r \in \mathcal{R}} r$; a set of connections $\mathcal{C}$, a connection being a triplet $(r, s, i) \in \mathcal{R} \times \mathcal{R} \times \mathcal{N}$ such that the node i is visited by both routes r and s. We denote by Δ_{rsi} the average walking time to transfer from route r to route s at node i.

In addition to these static characteristics of the system, transfer waiting times depend on the actual transfers that passengers make. These vary from day to day and are unknown at the moment of designing the schedules. Instead, we assume that we have access to historical transit data, from which we can extract observations of the unknown parameters. We treat this latter as a random variable ξ. We call the set of observations *scenarios* denoted by $\mathcal{S}$. Then for each scenario $\xi \in \mathcal{S}$, we denote by $T_{pqi}(\xi)$ the number of passengers transferring from trip p to trip q at node i and $T_{psi}(\xi)$ the number of passengers transferring from trip p to route s at node i ($T_{psi} := \sum_{q \in s} T_{pqi}(\xi)$). And we denote by $t_{pi}(\xi)$ the observed arrival (and departure) time of trip p at node i in the scenario ξ. Because in the data that we use for our experiments there is only one (approximate) value for the arrival and departure time of a vehicle at a stop, we simplify our presentation by using the notation $t_{pi}(\xi)$ for both. It would be straightforward to adapt our model were the two values available.

Figure 1 summarizes the main notations: it represents a schematic view of a transfer between two trips of two routes and the passengers and waiting time associated with the transfer (in a given scenario).

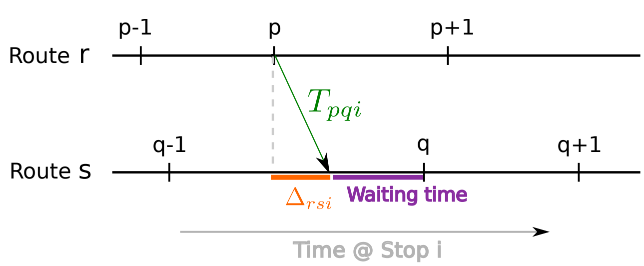

Figure 1: Illustration of a transfer from trip p of route r to trip q of route s at stop i in a given scenario

Similarly to previous works (e.g. [9, 8]) we assume that the underlying probability distributions of the demand in transfers and arrival times remain valid when the schedules are modified. This assumption seems reasonable when the schedules are shifted by only a few minutes. Another standard assumption is that the capacities of the vehicles are always sufficient to embark all entering passengers, so we do not consider capacity constraints.

2.2 Problem formulation

Our goal is to find a shift (or offset) of each transit trip such that the expected transfer waiting time is minimized.

We describe the ingredients of the optimization problem; some of them are common with [8] but for sake of completeness we quickly recall them as well.

Decision variables.

- *Offsets.* For every trip $p \in \mathcal{P}$, we introduce the continuous variable that represents its offset:

$$x_p \in \mathbb{R}: \text{ offset of the schedule of trip } p.$$

  Note that these variables define a unique offset per trip, i.e. an offset will be applied to the departure time form the first stop, and the travel times will be preserved.
- *Transfer choice.* For every pair of trips $(p, q) \in r \times s$ of a connection $(r, s, i) \in \mathcal{C}$, we define the binary variable that expresses whether passengers seeking to go from route r to s at node i and optimizing their journey would use this transfer:

$$y_{pqi}(\xi) = \begin{cases} 1 & \text{if passengers would transfer from } p \text{ to } q \text{ at node } i \text{ in scenario } \xi, \\ 0 & \text{otherwise.} \end{cases}$$

  Of course this passenger choice will depend on the schedule offsets.

Note that the x variable does not depend on the scenarios because we want a unique offset that will work for all the scenarios.

Constraints. We present here the minimal set of constraints; additional operational constraints that depend on the city are not described here.

- *Offset bounds.* Transportation operators may impose bounds on the time shifts:

$$l_p \leq x_p \leq u_p. \tag{1}$$

- *Transfer necessity.* For a trip $p \in r$ of a connection (r, s, i), such that there are passengers transferring from p to s in a scenario ξ (i.e. $T_{psi}(\xi) > 0$), these passengers would choose at least one transfer from p to a trip of s. This can be formulated as:

$$T_{psi}(\xi)(\sum_{q \in s} y_{pqi}(\xi) - 1) \geq 0.$$

- *Transfer feasibility.* A transfer from $p \in r$ to $q \in s$ of a connection (r, s, i) is feasible if there is enough time between the arrival time of p and the departure time of q at node i for passengers to walk from the stop of p to the stop of q. Let us denote by w_{pqi} this difference for the shifted schedules:

$$w_{pqi}(x, \xi) := (t_{qi}(\xi) + x_q) - (t_{pi}(\xi) + x_p) - \Delta_{rsi}. \tag{2}$$

  Passengers can only choose feasible transfers hence the constraint: $w_{pqi}(x, \xi)\, y_{pqi}(\xi) \geq 0$. Note that this constraint and the previous one link the variable y to the offset x.

Objective function. The objective to minimize is the total expected waiting time. In our context, since we only have access to a discrete set of observations of the random variable ξ, we approximate the expectation by:

$$\min_x \ \mathbb{E}_\xi[W(x, \xi)] \approx \sum_{\xi \in \mathcal{S}} p_\xi W(x, \xi) \quad \text{with} \quad W(x, \xi) = \sum_{(r,s,i) \in \mathcal{C}} \sum_{(p,q) \in r \times s} T_{psi}(\xi)\, w_{pqi}(x, \xi)\, y_{pqi}(\xi).$$

where p_ξ is the probability of scenario ξ.

Problem formulation. The problem of minimizing transfer waiting times can then be written as:

$$W(x,\xi) := \begin{cases} \min\limits_{x,y} & \sum\limits_{\xi\in\mathcal{S}} p_\xi \sum\limits_{(r,s,i)\in\mathcal{C}} \sum\limits_{(p,q)\in r\times s} T_{psi}(\xi)\, w_{pqi}(x,\xi)\, y_{pqi}(\xi) \\ & l_p \le x_p \le u_p \qquad \forall p \in \mathcal{P} \\ \text{s.t.} & T_{psi}(\xi)(\sum_{q\in s} y_{pqi}(\xi) - 1) \ge 0 \quad \forall p \in r, (r,s,i)\in\mathcal{C}, \xi\in\mathcal{S} \\ & w_{pqi}(x,\xi)\, y_{pqi}(\xi) \ge 0 \qquad \forall (p,q)\in r\times s, (r,s,i)\in\mathcal{C}, \xi\in\mathcal{S} \\ & y_{pqi}(\xi) \in \{0,1\} \qquad \forall (p,q)\in r\times s, (r,s,i)\in\mathcal{C}, \xi\in\mathcal{S} \end{cases} \tag{3}$$

With $w(x,\xi)$ given by (2). Contrarily to previous works [9, 8], this problem is a mixed-integer quadratic program. Even by relaxing the integrality constraints, the problem is not convex, because of the form of the objective and the third family of constraints. However the problem has an interesting property that will help us to design efficient heuristics.

3 Decomposition strategy

3.1 Separability property

In problem (3), if we fix the offset variable x, the optimization subproblem for y becomes separable with respect to the scenarios and the aggregated trip-to-route transfers. Indeed, when the schedules are fixed, we only need to determine which are the optimal choices of transfers for the passengers and deduce the associated waiting time. This can be done independently for each scenario and aggregated transfer. To exploit this property, we consider a two-stage optimization process, where the problem can be rewritten (equivalently) as:

$$\min_{l\le x\le u} W(x) \quad \text{with} \quad W(x) = \sum_{\xi\in\mathcal{S}} p_\xi \sum_{(r,s,i)\in\mathcal{C}} \sum_{p\in r} W^*_{psi}(x,\xi), \tag{4}$$

where $W^*_{psi}(x,\xi)$ is the optimal value of the subproblem:

$$W^*_{psi}(x,\xi) = \begin{cases} \min\limits_{y} & \sum_{q\in s} T_{psi}(\xi)\, w_{pqi}(x,\xi)\, y_{pqi}(\xi) \\ & T_{psi}(\xi)(\sum_{q\in s} y_{pqi}(\xi) - 1) \ge 0 \\ \text{s.t.} & w_{pqi}(x,\xi)\, y_{pqi}(\xi) \ge 0 \qquad \forall q\in s \\ & y_{pqi}(\xi) \in \{0,1\} \qquad \forall q \in s. \end{cases} \tag{5}$$

3.2 Subproblem structure

We can compute a closed-form optimal solution of the subproblem (5):

$$y^\star_{pqi} = \begin{cases} 1 & \text{if } q = \text{argmin}_{q\in s}\{w_{pqi}(x) : w_{pqi}(x) \ge 0\}, \\ 0 & \text{otherwise.} \end{cases}$$

Note that if the argmin is not a singleton, one can simply take any index of the set. Therefore, for a fixed offset, we can compute in parallel the solutions of the subproblems for the different scenarios and potentially for batches of aggregated transfers. That ensures that it will take a near-constant time regardless of the number of transfers we want to optimize and the number of scenarios that we consider.

3.3 Decomposition algorithm

As we have seen it easy to solve the subproblems for fixed schedules; now the question is how do we fix the schedules' offsets? Several options are possible. For example the generalized Benders decomposition [5] is well-suited for this situation. However, since the function W in (4) is not convex, there are no optimality guarantees. A more simple heuristic can be obtained using local search techniques (see e.g. [7, 3]). The idea is to discretize the interval $[l_p, u_p]$ (by minute for example) and explore the discrete set of possible offsets, seeking schedules that reduce the waiting time. Since each

Algorithm 1 Basic local search algorithm for the optimization of the schedules

```
Inputs: Number of iterations or time limit, original schedule, scenarios, perturbation proportion
                                                                              ▷ Initialization
The best schedule is the original schedule

                                                                                   ▷ Main loop
for the specified number of iterations or until the time is up do
    Randomly select a subset of trips to offset
    Randomly select a perturbation an offset (within the adequate bounds)
    Evaluate the impact of the offset
    if the new schedule is feasible and the waiting time is reduced then
        Update the best schedule
    end if
end for
```

evaluation of an offset will be very fast, we can potentially explore a very large space (and so reach very good solutions) quickly.

We refine algorithm 1 with a few features:

- Greedy iterations: we add greedy iterations periodically in the algorithm. In these, we go through all the transfers for which the transfer time is not 0 (starting with the ones with the largest waiting time or in a random order) and we compute the optimal offset of the involved trips for this transfer. Of course, this optimal offset may increase the waiting time at other transfers or even make them infeasible. This is why the greedy iterations are not always successful. However we observed that especially at the beginning of the algorithm, greedy iterations help to quickly decrease the waiting time.
- Smart bounds: to avoid exploring offsets that would lead to infeasible schedules, we refine the original bounds (1) to take into account the current offset and the schedules of the previous and next trips.
- Parallel inner iterations: starting from a given schedule, we explore a number of different local search iterations in parallel, then we retrieve the best solution, and start again.

4 Preliminary results

In this section, we present preliminary results of applying our decomposition method to a transit system and compare it to the state of the art.

Data. We used the transit data of the city of Nancy extracted from e-card validation collection. As a first step, we only considered the two main lines (a tramway and a bus) which are the most regular and loaded. To generate the scenarios, we assumed that for similar weekdays, the distribution of ξ is i.i.d. Hence, we carefully selected 3 "standard" days and took them as equi-probable. For the sake of comparison, we test problems composed of 2 and 3 scenarios (days).

Because there are some inherent errors in the recording of the data and in the reconstruction of the trips, we performed some data cleaning and filtering to make the data coherent. For example, we filtered the observed transfers to keep only the ones for which the difference between departure and arrival time is larger than the walking time between the stops. Moreover, for each route, we only kept the trips that are common to the three days. An alternative would be to consider the number of trips of a route per day as stochastic and add that as an additional random parameter. We ended up with a few hundreds transfers per day to optimize.

Results. We compare our results to the approach of [8]. In this latter, the "trip window" parameter allows to give more options for passengers to optimize their transfers according to the schedule shifts. The larger this parameter, the more realistic the model but also the larger the optimization problem (in terms of number of variables and constraints). In our experiments, we fixed this parameter to 15 (i.e passengers optimize their transfers within 15 trips of their original transfer). For both approaches,

Table 1: Comparison of the reduction in the expected waiting time for 2 scenarios (in %) between the MILP approach [8] and our local search approach

Offset bound	2 min	3 min	4 min	5 min	6 min	7 min	8 min	9 min	10 min	Avg
MILP	80	89	95	96	95	98	99	99	99	**94.44**
LS	70	80	85	87	89	87	90	90	89	**85.22**
Difference	10	9	10	9	6	11	9	9	10	**9.22**

Table 2: Comparison of the reduction in the expected waiting time for 3 scenarios (in %) between the MILP approach [8] and our local search approach

Offset bound	2 min	3 min	4 min	5 min	6 min	7 min	8 min	9 min	10 min	Avg
MILP	72	79	87	86	86	90	83	87	56	**80.66**
LS	62	69	76	80	78	84	83	81	83	**77.33**
Difference	10	10	11	6	8	6	0	6	-27	**3.33**

we choose a time limit of one hour. Although we may spend more computing time on this problem in a real-life setting, we believe that one hour is a good limit here since we only consider a small subnetwork and few scenarios.

Table 1 presents the results for the stochastic optimization over 2 scenarios and Table 2 over 3 scenarios. For different values of the offset bounds (see constraint (1)), we report the reduction in percentage of the expected waiting time for the optimized schedules compared to the initial ones. We take the solution provided by the MILP [8] as a reference, although the solver that we used (Coin-OR CBC) did not converge within the time limit of one hour in all but one case. In the case of non-convergence, we retrieve the best value found by the solver.

We can see that the local search heuristic leads to a large reduction of the expected waiting times in all cases (at least 62% and up to 90%). We observe that the reduction is generally larger when the offset bound is larger, which is expected since the model offers more flexibility. We also notice that for the same offset bound, the reduction is smaller for three than for two scenarios, which is also intuitive: the more scenarios, the more various transfers to optimize and the more chances of getting "conflicting" ones.

When we compare the LS solution to the MILP one, we see that we loose in the worst case 10% of reduction in the waiting time. Moreover, this loss in smaller for three than for two scenarios (Table 1 versus 2) . We even get 27% more reduction in one 3-scenarios instance. These observations are encouraging because our ultimate goal is to handle much more scenarios to have a stochastic model that is as general and realistic as possible.

5 Conclusion

In this paper we proposed a new MIQP model for the stochastic optimization of the transfer waiting times in a public transport system. We highlighted a separability property of the model that allowed us to design an efficient two-stage optimization strategy. We proposed a simple and easy-to-implement parallel local search heuristic. We presented some experiments on a subnetwork of the city of Nancy (France) where our approach gave a considerable reduction of the expected waiting times. Compared to previous approaches to the stochastic optimization of the schedules, our method does not rely on generic optimization solvers and is parallel by construction. This should allow to handle much larger problems, which is the goal of our future work.

References

[1] Kittelson & Associates. Federal Transit Administration, Transit Cooperative Research Program, and Transit Development Corporation. *Transit capacity and quality of service manual*, volume 100. Transportation Research Board, 2003.

[2] John R Birge and Francois Louveaux. *Introduction to stochastic programming*. Springer Science & Business Media, 2011.

[3] Thomas A Feo and Mauricio GC Resende. Greedy randomized adaptive search procedures. *Journal of global optimization*, 6(2):109–133, 1995.

[4] R. Gallotti and M. Barthelemy. Anatomy and efficiency of urban multimodal mobility. *Scientific reports*, 4, 2014.

[5] A.M. Geoffrion. Generalized Benders decomposition. *Journal of Optimization Theory and Applications*, 10(4):237–260, 1972.

[6] 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.

[7] Jan Karel Lenstra. *Local search in combinatorial optimization*. Princeton University Press, 1997.

[8] S. Michel and B. Chidlovskii. Stochastic optimization of public transport schedules to reduce transfer waiting times. In *2016 IEEE International Smart Cities Conference (ISC2)*, pages 1–6, 2016.

[9] R. Nair, C. Coffey, F. Pinelli, and F. Calabrese. Large-scale transit schedule coordination based on journey planner requests. *Transportation Research Record: Journal of the Transportation Research Board*, 2351(1):65–75, 2013.

[10] V. Poorjafari and W.L YUE. Synchronisation of public transport services: Comparison of methods. In *Proceedings of the Eastern Asia Society for Transportation Studies*, volume 9, 2013.

[11] Richard H Pratt, Neil J Pedersen, and Joseph J Mather. *Traveler Response to Transportation System Changes: Transit Scheduling and Frequency*, volume 95. Transportation Research Board, 2004.