Optimizing regular symmetric timetables: a method to reach the best modal split for railway

Maurizio Bruglieri
Dipartimento di Design, Politecnico di Milano
via Durando 38/a, 20158 Milano.
E-mail: maurizio.bruglieri@polimi.it

Roberto Maja
Dipartimento di Design, Politecnico di Milano
via Durando 38/a, 20158 Milano.
E-mail: roberto.maja@polimi.it

Samuel Tolentino
Graduate from Politecnico di Milano
Piazza Leonardo da Vinci 32, Milano.
E-mail: ing.samuel.tolentino@gmail.com

Abstract

A regular timetable is a collection of events that repeat themselves every specific time span. This even structure, whenever applied at a whole network, leads to several benefits both for users and the company, although some issues are introduced, especially about dimensioning the service. It is therefore fundamental to properly consider the interaction between the transport demand and supply, to create an effective network timetable.

In this paper, a specific cycle base to solve the Cycle Periodicity Formulation (CPF) of a symmetric timetable is proposed. This is combined with a modal choice model, adding the possibility to disable stops along lines for further increasing the transportation demand acquired by the railway system. The problem is modeled as a Mixed-Integer Linear Program (MILP) and solved through a state of the art MILP solver (CPLEX). Computational results regard a case study about a railway network in a northern Italy region.

1. Introduction

The main feature of a regular timetable, in the users' perspective, is the repetitiveness along the day, enabling the passenger to memorize the characteristics of just trip (time, duration etc) as the service will be the same every time. In the case of a regular service over the whole network, the same mechanic applies also to the interchange nodes, simplifying the journey planning. Regarding the operating company, adopting a regular structure for the service results in simplifying both planning and operating the system: every solution found in a time module is replicable during the entire service. Specific needs could anyway be faced with local singularities.

The importance of seeking the optimum while looking for a feasible regular timetable is supported by the fact that the transportation supply is basically flat over the day, both during peak and low-demand hours: a well-balanced timetable can help to produce an effective service with a fair amount of resources.

1.1   Related literature

The Cyclic Railway Timeable Problem (CRTP), regarding the planning of regular railway timetables, has been developed by Voorhoeve in (Voorhoeve, M. 1993), using the Periodic Event Scheduling Problem (PESP) introduced by Serafini e Ukovich (Serafini, P. 1989). Successively (Nachtigall, K. 1994) introduced the Cycle Periodicity Formulation (CPF), based over an auxiliary graph whose nodes represent circulation events and whose arcs model time constraints on linked events. This formulation is the basis for the following studies.

In (Kroon, L. 2003) PESP formulation is improved introducing variables associated with the travel arcs, formerly considered fixed a priori. Liebchen (Liebchen, C. 2004) studied the characteristics of symmetry in regular timetable focusing on the optimality issue.

Regarding studies with different kind of trains on the same track, a Dutch study (Goossens, J. 2006) can be cited. In that work, a model for a network featuring different ranks of trains is developed, each with specific stop pattern, aiming to minimize operational cost.

None of these studies considered the interdependency between mobility demand and transportation supply. As the goal of a public transportation service should be coping whit the mobility demand with a fair amount of resources, considering sustainability of the whole transportation system, the importance of timetable planning to improve the railway modal split is relevant.

Chierici (Chierici, A. 2004) took into consideration this matter. He proposed an optimization model integrating a discrete choice sub-model for the available transportation options, with a function to maximize the railway users. The CPF formulation is enriched with a Logit model: the resulting formulation is therefore NP-hard for the mixed-integer linear part and nonlinear for the Logit part. Cordone and Redaelli (Cordone, R. 2011) resumed Chierici's work analyzing the Logit function's form, developing a piecewise linearization to be used in ad-hoc bounding procedures. They tested two kind of branch-and bounding algorithms that proved effective for instances up to about ten lines. Liebchen and Peeters in (Cordone, R. 2011) further examined the formal structure of the problem to find better cycle choices in order to reduce the required computing time.

1.2   Original contributions

This work deals with a specific condition of periodic timetables, as it is symmetry. A minimal set of decision variables is used, taking advantage of symmetry by choosing a particular cycle basis to solve the CPF: an efficient formulation of the problem can decrease the computational time, that is a critical point in the problem solvability.

A new feature is introduced: the possibility to automatically deactivate stops along the lines in order to reduce travel time between origin-destination pairs (OD) that use that path. This might lead to an increase in the number of passenger, due to the reduced travel time, that has to be compared with the loss of the skipped station's passengers. The model can establish the optimal stopping pattern for each line considering the physical constraints (as for example in single track lines). Including this feature in the maximization function can help planners to better understand the potentiality and critical constraints of the system, to properly evaluate different scenarios.

In the last section, results of a case study are analyzed. A symmetric schedule is produced for some instances regarding a portion of a railway network, located in a northern Italy region, consisting of 9 lines with 72 stations, mainly single tracks.

2   Solution method

2.1   Railway model

We consider a railway network as a set of stations linked by one or more tracks (in this work no difference is intended between the terms "station" and "stop"). Trains circulate on rails, with a defined origin and destination, stopping at a set of stations, using a specific sequence of tracks for a given time. The network configuration is settled by the choice of the terminal stations for each line along with the intermediate stops; where two or more lines meet there is an interchange. Every station has a group of users, that is the sum of all the passengers entering or leaving the system there, plus those who possibly change line.

The railway service can be described effectively with an auxiliary graph $G = (N, A)$. Each node N represents the arrival or departure of a train; in the case of regular timetables, events happen every P minutes. A station is a set of nodes concerning the same physical place. A temporal constraint is assigned to each arc A, representing the minimum amount of time separating two linked events. In the graph we can distinguish two kind of arcs intended to model the service: travelling arcs (from a departure node to an arrival one) and dwell arcs (vice versa). To complete the passenger's paths, interchange arcs are added (from an arrival node of a line to the departure node of another line in the same station).

2.2   Modal choice model

Transport demand is defined as the quantity of actual or potential users who would use a transportation service in a given time span. It is usually represented as a flow and, once assigned, become the effective traffic on the infrastructural elements.

A transport demand model set a relation between demand flows on one side and specific characteristics of the transportation system on the other. User's trips are the result of a sequence of choices, with various complexity, spanning from very short-term ones (like the specific path for the current trip) to log-term ones (like where to live and work). Every choice is interrelated with the available transportation systems.

A widely-used modal choice analysis method (Cascetta, E. 2009) is the four-stage transportation model. We consider the modal split sub-model, assuming invariable the origin-destination matrix (OD Matrix) that supply the number of displacements between each OD pair.

$p^i[m/oshd](\pmb{SE,T})$, *modal split model*: estimates the share of users belonging to category i who travel between o and d for the s reason in the time span h, using the transport option m; *SE, T* are the socioeconomic characteristic of the territory and transportation system.

The discrete-choice model used is the Multinomial Logit, that express the probability of a specific alternative j among the k available as

$$p[j] = \frac{\exp(V_j/\theta)}{\sum_{k=1}^{m} \exp(V_k/\theta)}$$

where V is a vector of attributes for each alternative and θ a parameter, both to be estimated.

The amount of passenger λ_{od} travelling over an OD pair, using a couple of stations belonging to the set S, is therefore given by the product of the global quantity of trips Λ_{od} for that OD and the railway modal split. Considering variable only the travel time of trips by railway, for the sake of simplicity, we can write

$$\Lambda_{od} * \frac{e^{V_{od}^{(t)}(t_{od})}}{e^{V_{od}^{(b)}} + e^{V_{od}^{(b)}} + e^{V_{od}^{(t)}(t_{od})}} = \lambda_{od} \qquad o, d \in S$$

and the railway users maximization function is:

$$\max \left(z = \sum_{o \in S} \sum_{d \in S} \lambda_{od} \right)$$

The other competing transport options covered in this work are private car (overlooking the issues related to driving license and car possess) and public bus transport.

Regarding the route choice model, we consider a rigid assignment of demand flows to pre-determined path, as is typical in uncongested networks; furthermore, any path change within the railway system typically results in a consistent raising in trip time, thus becoming largely ineffective compared with the competing system. The assignment procedure for the train service starts calculating a standard travel time for each arc, then seeks for the lower trip time possible for each OD pair, assuming that trains stop at every station, finally assign the entire demand flow to the minimum time route.

2.3 Cycle periodicity formulation

The classical CPF (Nachtigall, K. 1994) has been formalized as:

$$\sum_{a \in C^+} x_a - \sum_{a \in C^-} x_a = P * q_C \qquad C \in \mathcal{C}$$

$$l_a \leq x_a \leq u_a \qquad a \in A$$

$$h_c \leq q_c \leq k_c \qquad C \in \mathcal{C}$$

$$x_a \in R \qquad a \in A$$

$$q_c \in Z \qquad C \in \mathcal{C}$$

where is x_a is the main variable regarding arcs A, with lower and upper bounds on each arc (by physical characteristics and planning requirements); $\mathcal{C}$ is the set of possible cycles; q_c is an integer with lower and upper bounds (related to the arc ones); P is the period of the railnet.

2.4 New contributions

2.4.1 Symmetry

A useful condition in a regular timetable net is the symmetry of train paths for the opposite directions of each line. This is particularly effective at interchanges: in the symmetric configuration the waiting time is the same for both directions, while in any other case would differ, increasing the passenger's perception of "losing time" along the inefficient direction (Liebchen, C. 2004).

Exploiting the hypothesis of perfect symmetry, is possible to use an undirected graph whose arcs model time span constraints between two linked events π_i and π_j, independently from the events' order (i.e. under the condition $x_{ij} = x_{ji} = x_a$). For each event π we introduce two variables that represent the distance of that event from the extremity of the period: $d_{0\pi} = (t_\pi - 0)mod P$, corresponding to the time of the event, and $d_{\pi 0} = P - d_{0\pi}$, corresponding to the dual one (that is the opposite event occurring along the opposite direction).

The direction of a line is related to the order of the events: whenever two distinct events π_i, π_j present $d_{0\pi_j} > d_{0\pi_i}$ it is the outward direction, vice versa is the return direction when $d_{\pi0_j} > d_{\pi0_i}$ (except if the time exceed the period P).

Considering the events of departure from the first station and the dual of the arrive at the last station, it is possible to write the CPF condition with this alternative formulation:

$$\sum_{a \in L} x_a + d_{0\pi\,L} + d_{\pi0\,L} = k * \mathrm{P}$$

where x_a are the variables for travelling and dwelling arcs, $d_{0\pi\,L}$ and $d_{\pi0\,L}$ are the variables of the fictitious arcs representing the distance of departure and arrival events at line ends from the extremities of the period P, k is an integer multiplier in order to take into account the possibility that a line is operated in a bigger time span than the period P.

2.4.2 Suppression of stops

An innovation in this model is the possibility to evaluate the suppression of intermediate stops. A binary variable Φ is introduced, to put to zero the dwelling time x_s where it is favorable in order to reach a better modal split (junctions excluded).

2.4.3 Line periodicity formulation

The comprehensive optimization model proposed is:

$$\sum_{v \in L} x_v + \sum_{s \in L} (x_s * \phi) + d_{0\pi\,L} + d_{\pi0\,L} = k * \mathrm{P} \qquad L \subseteq A$$

$$l_v \le x_v \le u_v \qquad\qquad v \in A$$

$$l_s \le x_s \le u_s \qquad\qquad s \in A$$

$$\Phi_s \in \{0,\,1\} \qquad\qquad s \in A$$

$$x_v \in R \qquad\qquad v \in A$$

$$x_s \in R \qquad\qquad s \in A$$

$$\Lambda_{od} * \frac{e^{V_{od}^{(t)}(t_{od})}}{e^{V_{od}^{(b)}} + e^{V_{od}^{(b)}} + e^{V_{od}^{(t)}(t_{od})}} = \lambda_{od} \qquad\qquad o,\, d \in S$$

$$t_{od} = \sum_{a \in I_{od}} x_a \qquad\qquad I_{od} \subseteq A$$

$$x_a \in R \qquad\qquad a \in A$$

where A is the general set of arcs (travel x_v, dwell x_s and interchange), L the set of lines, Φ_s a binary variable accounting the station suppression; o, d the stations origin and destination of a path I whose trip time is the incoming variable in the modal choice model, where λ_{od} is the railway modal split of the general demand Λ_{od} over an OD pair.

2.4.4 Interchange duration

To determine the waiting time at junctions, $d_{0\pi}$ and $d_{\pi0}$ variables are used. In the most general case, were the interchange happens in station that are not terminus, we are in the situation represented in figure (for the sake of simplicity only a couple of line is shown).
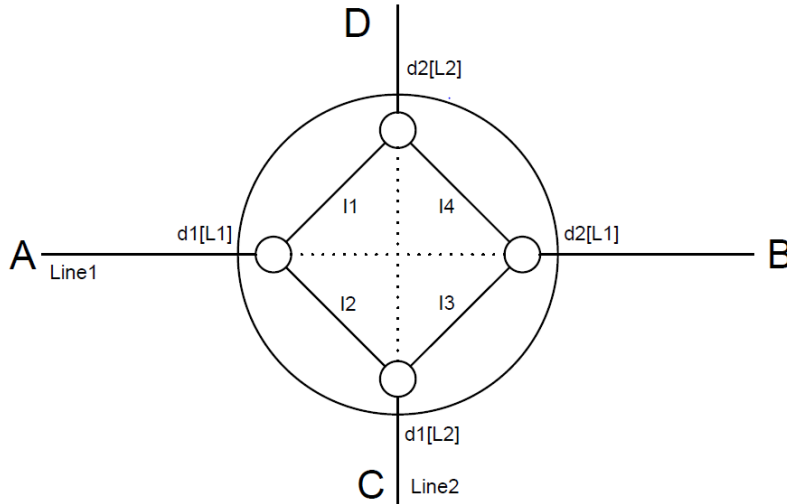
Figure 1 – Graph detail of an interchange

Considering Line1 and Line2 as in the example, there are four possible directions of interchange. The equations used to determine the waiting time have the same structure of the periodicity equation:

$$I_1 = P - (d_{1\,L1} + \frac{x_{s\,L1}}{2} + d_{2\,L2} + \frac{x_{s\,L2}}{2})mod\mathrm{P}$$

where $x_{s\,L}$ are the dwelling times of each line L in the junction station, added to consider the slowdowns of deceleration and reacceleration from the travelling speed and the boarding operations (except for the terminal station, where these are not defined).

Noticing that the graph is undirected, but lines are, is necessary to change the variables to be considered for each interchange, according to the table reported.

Line 1	Line 2	Variables
outward	outward	$d_{0\pi\,L1}, d_{\pi0\,L2}$
outward	return	$d_{0\pi\,L1}, d_{0\pi\,L2}$
return	outward	$d_{\pi0\,L1}, d_{\pi0\,L2}$
return	return	$d_{\pi0\,L1}, d_{0\pi\,L2}$

Table 1: Variables for the interchange time

2.4.5   Brief overview on constraints

2.4.5.1   Maximum speed and distancing

Variables of travelling arcs x_v have a lower limit due to the maximum speed allowed by tracks or rolling stock. A ceiling too is added to reasonably limit travel time. As the inverse relationship between speed and time is nonlinear, the Mc Cormick method is used (McCormick, G. 1972) to rewrite it.

Trains distancing is obtained imposing that in every station the variables $d_{0\pi}$ of each concurrent line should differ at least of a certain quantity (time distancing). At the scale of the model, any other circulation regime can be reduced to time distancing. Real timetable will be created on the model results adjusting these locally to the real constraints.

2.4.5.2   Dwell time

In this model, the dwell time of trains in stations x_s includes also the slowdowns of the decelerating and reaccelerating phases, that do not occur if the train doesn't stop. Dwell time has a variable lower bound, given by the proper stopping time and the slowdowns (that depend on the adjacent

arcs) and a ceiling as a reasonable value. If the station is deactivated, Φ_s multiplier variable become 0, thus annihilating the dwelling time variable.

2.4.5.3 Meeting requirements for single track lines

One of the stronger constraints in railway operation is that two trains travelling in opposite directions on the same single track have to meet only in specific stations. Under perfect symmetry, it is possible to demonstrate that the minimum number of meeting points is

$$NI = 2k - 1 - d_{0\pi\,L}\,DIV\,\frac{P}{2} - d_{\pi0\,L}\,DIV\,\frac{P}{2}$$

where k is the integer multiplier in the periodicity equation of the considered line. If at least NI meeting point aren't assured along the line, the problem is infeasible. Each additional meeting point gives more elasticity to the line, allowing the model to choose among them to provide the best timetable.

To avoid meeting of trains where aren't allowed, it is necessary to force for k events $\pi_{i_{1,..,k}}$ the following condition, with x_s bigger than a given safety parameter:

$$d_{0\pi\,i} + \frac{x_s}{2} = k\frac{P}{2}$$

Whenever the line is only partially single track, it is necessary to evaluate a posteriori the reduction of k by one or more units verifying in which part of the line the meeting happens, in order not to over-constrain the problem.

In the unlikely case two or more lines circulate on the same single track, some additional constraints are required. To avoid trains of different lines to meet coming from opposite directions, it is necessary to guarantee that, for any single track arc, every couple of occupation window time of each couple of trains are not overlapping. Considering the events π_i e π_j that are extremities of a single track arc, this can be reached exploiting $d_{0\pi\,j}$ and $d_{0\pi\,i}$ variables for every line (analytical part is omitted for conciseness).

2.4.6 Linearization

Various elements of the proposed model are non-linear. Regarding all the modular constraints, it is always possible to define some dummy integer variable to convert them into a Mixed Integer Linear Program. The modal choice model conversely, as it is, could not be included in the MILP formulation, plus it is a not steady concave nor convex function in the feasible region.

Despite Cordone and Redaelli (Cordone, R. 2011) already proposed an operating method to seek for the global optimum implementing an ad-hoc branch and bound algorithm, in this work we overlook this part, being enough to test model's behavior with a simplified function. Studies on refined solving techniques for this model are put off to future works.

The original Logit function is replaced by a rough linearization between the extreme points of the feasible region. The upper bound for every time variable x_a leads to a feasible region for variable t_{od}, so we have λ_{od} in the extremity two point $t_{od\,MIN}$ and $t_{od\,MAX}$. As the Logit function is decreasing, λ_{od} are respectively $\lambda_{od\,MAX}$ and $\lambda_{od\,MIN}$.
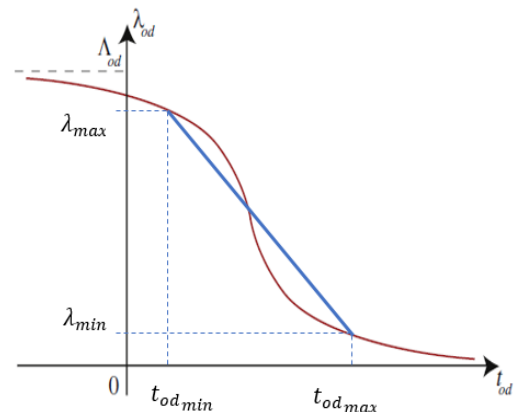

Figure 2 – Logit linearization

Dealing this way with the Logit there isn't any guarantee that the configuration found is the best solution for the original problem, however feasible and reasonably good. Unfortunately, no valid indicator of the optimum gap is available; a poor one could be the gap between the results found with the linearization and the recalculation with the original formula, weighted over the global demand:

$$\frac{\sum_{od}((\lambda_{od_{lin}} - \lambda_{od_{esp}}) * \Lambda_{od})}{\sum_{od} \Lambda_{od}}$$

3 Computational experiments

3.1 Model configuration

3.1.1 Rail net

The railway network of the case study located in northern Italy region. It consists of nine lines, four of them are entirely single track and two only partially. Furthermore, two single track lines converge in one, so all the cases are included.

There are 72 stations, of which 8 are junction between two or more routes. The catchment area for the competing transport system is sourced from the OD Matrix, resulting in more than 140000 individuals.

3.1.2 Transport options

The competing transport systems accounted in this study are the private car and public bus service. The private car mode has a set of parameters concerning travel time and fuel cost, whose value for each OD pair has been determined exploiting Google®'s API. Regarding the bus transport system, real world data are extremely difficult to collect. Parameters of travel time, fare amount and frequency has been sought for the most used routes (about 80), for the other routes a randomized set is used based on the previous likeliness.

3.1.3 Main parameters

The period P is equal to 60 minutes. In general, typical periods are 30, 60 or 120 minutes. Critical parameter are the upper bounds of x_a variables, because these affect directly the quality of the Logit linearization.

Considering travel arcs, we observed the results of variating the coefficient c used to determine the upper bound in travel time, having $x_v \in [x_{MIN}, c * x_{MIN}]$ and $c \in R$. Increasing this parameter, the space of the feasible solutions widens, but the linearization worsens, as reported in table 2. The choice for this instance is $c = 2$.

Scale parameter for the upper bound on travel time	1.8	1.9	**2**	2.1	2.2
Railway users (choice model linearized)	35562	35726	35858	35979	36098
Railway users (original formula with linear optimum variable set)	33717	33779	**33803**	33789	33797
Linearization gap	1.31%	1.38%	1.46%	1.55%	1.63%

Table 2: Results varying the upper bound on travel time

Regarding β and θ parameters of the Logit formulation, values given by Maja in (Maja, R. 1999) have been used. Although this solution for real applications should be avoided, needing a new calibration of the modal choice model related to the specific area, considering the demonstrative purpose of this work any plausible value would have do.

3.2 Numerical results

Several instances of the same case study have been analyzed. Aside from the base scenario (I), results are presented for the following: (I-bis), having some railnet parameters variation along a critical line; (II), featuring the insertion of two new no-stop lines between two couples of cities along existent routes; (II-bis), with the assignment of some stops to these fast lines.

Besides the mentioned above, results for some other benchmark scenarios based on the (II-bis) are shown: one with station deactivation inhibited, another with only double tracks, and a last one featuring both simplifications.

Despite there are minor differences between the scenarios, they can be considered really different instances, because the modal choice model is quite sensitive to the infrastructural changes, as can be seen both in numeric results and computational time.

In the following table and chart are collected the results for the mentioned scenarios.

| Scenario	I	I-bis	II	II-bis	All On	Double Track	All On & D.T.
Lines	9	9	10	10	10	10	10
Deactivated stops	6	5	4	4	0	2	0
Users (linear solution)	35858	35935	39267	39444	39388	40139	40066
Users (original formula)	33803	33880	37585	37777	37697	38812	38723
Individuals in catchment area	140909	140909	142228	142228	142228	142228	142228
Railway modal split	23,99%	24,04%	26,43%	26,56%	26,50%	27,29%	27,23%
CPU time [s]	171	2306	332	1477	197	23	8
Linearization gap	1,46%	1,46%	1,18%	1,17%	1,19%	0,93%	0,94%

Table 3 – Computational results
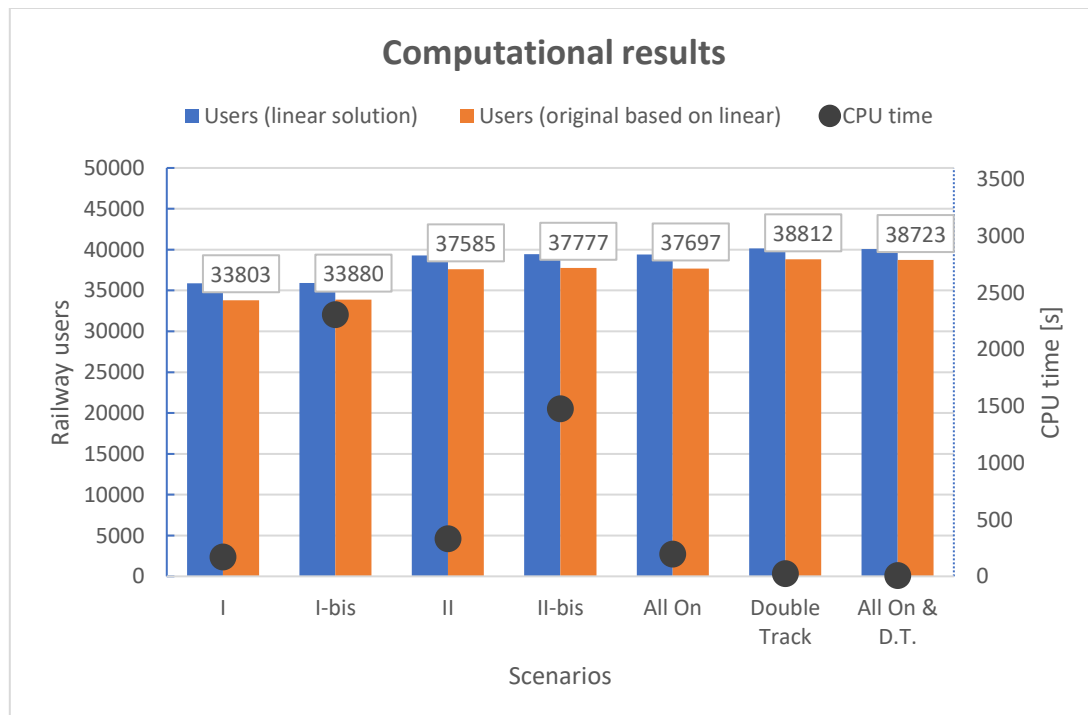
Figure 3 – Computational results

Every MILP problem has been resolved reaching the optimum in a reasonable CPU time. The benchmark scenarios point out that penalizing constraints, such the single track ones, prove to have quite an impact over the solution time, (in this case greater than station deactivation possibility). Whenever these complications are overlooked, the problem is solved in less than 10 seconds, showing how the model is capable to find a feasible and good solution very quickly. The possibility to deactivate intermediate stops, as it is provided, proves anyway to be effective, since the total number of users decrease in the scenario that doesn't feature it.

Though the optimal set of variables determined with the linearized model is probably sub-optimal for the original problem, the users number trend for the different scenarios is similar both in the linearization version and the original one, leading to the hypothesis that the optimality gap wouldn't be huge.

4. Conclusions

This paper deals with the optimization problem of regular symmetric timetables, considering the relationship between the characteristic of the competing transport systems and the transport demand.

The classical CPF problem is revised with a formulation based on an undirected graph, while a discrete choice model is introduced to estimate the modal split dependency on the variation of railway trip time. The maximization function aim is to find the best feasible timetable to gain the better modal split given the infrastructure characteristics.

The proposed model features the possibility to deactivate intermediate stops along lines: this automatically happens whenever the traveler's utility increase is sufficient to lead to a rise in the number of total passengers (considering the lost ones).

With the purpose of testing the model, a linearized version of the modal choice model is implemented, thus reconducting the problem to a MILP, being able to deal with it using a commercial solver like CPLEX®.

Computational results regarding a real-life case study are presented. The solving time is very reasonable for instances regarding a ten of lines, with various single track ones. The solutions found for the linear problem are feasible also for the original one, though it is not possible to guarantee the quality of them respect to the global optimum of the nonlinear problem. Results are anyway encouraging, as the trend of the solution for the different scenarios is similar both for the linear version and the original one.

Further studies will be focused on a more accurate linearization of the Logit function, studying the behavior over a wider class of scenarios, and the development of a method to solve the nonlinear problem to optimality within reasonable time, possibly using the variable set of the linearized solution to establish a lower bound on the optimum.

References

CASCETTA, E., 2009. *Transportation Systems Analysis: Models and Applications.* Springer US.

CHIERICI, A., CORDONE, R. and MAJA, R., 2004. *The demand-dependent optimization of regular train timetables.*

CORDONE, R. and REDAELLI, F., 2011. Optimizing the demand captured by a railway system with a regular timetable. *Transportation Research Part B: Methodological,* **45**(2), pp. 430-446.

GOOSSENS, J., VAN HOESEL, S. and KROON, L., 2004. A branch-and-cut approach for solving railway line-planning problems. *Transportation Science,* **38**(3), pp. 379-393.

GOOSSENS, J., VAN HOESEL, S. and KROON, L., 2006. On solving multi-type railway line planning problems. *European Journal of Operational Research,* **168**(2), pp. 403-424.

KROON, L.G. and PEETERS, L.W.P., 2003. A Variable Trip Time Model for Cyclic Railway Timetabling. *Transportation Science,* **37**(2), pp. 198-212.

LIEBCHEN, C., 2004. Symmetry for Periodic Railway Timetables. *Electronic Notes in Theoretical Computer Science,* **92**, pp. 34-51.

LIEBCHEN, C. and PEETERS, L., 2009. Integral cycle bases for cyclic timetabling. *Discrete Optimization,* **6**(1), pp. 98-109.

MAJA*, R., 1999. Accordo di programma per la riqualificazione e il potenziamento della linea ferroviaria Milano-Mortara. Technical Report, Politecnico di Milano*

MCCORMICK, G.P., 1972. *Converting general nonlinear programming problems to separable nonlinear programming problems*.

NACHTIGALL, K., 1994. *A Branch and Cut Approach for Periodic Network Programming.* Univ., Inst. fur Mathematik.

SERAFINI, P. and UKOVICH, W., 1989. A Mathematical Model for Periodic Scheduling Problems. *SIAM Journal on Discrete Mathematics,* **2**(4), pp. 550-581.

VOORHOEVE, M., 1993. Rail scheduling with discrete sets. Unpublished report, Eindhoven University of Technology, The Netherlands.