Train Timetabling Problem for Complex Railway Systems

Maite Peña

Institute for Research in Technology, Comillas Pontifical University
(+34) 915 42 2800 ext. 4502
maite.pena@iit.upcomillas.es

Abstract This paper deals with the design of timetables for complex railway systems. Rail transportation has become a rapid, clean and efficient way to transport passengers and freights. A good timetable design is crucial to ensure the quality of the services trying to improve the efficiency of the whole system. Only such a sustainable transportation system will be in accordance to the requirements of future cities. Underground and commuter railway systems are becoming important backbones in providing access to downtowns and supporting the main pressure to move as many passengers as possible. So the timetable design as part of the capacity problem becomes crucial in providing good performance of the system.

In this paper, a new model for solving the train timetabling problem considering general complex rail topologies and general operational policies is developed. This model overcomes some of the limitations of previous models found in the literature, such as their particularization to special topologies of the railway system, the difficulties of dealing with stations with a determined number of tracks or with train services accomplishing different routes and the difficulties of designing completely new timetables for those cases when no initial timetable is given. It also overcomes the difficulties of designing completely new timetables for those cases when no initial timetable is given. The model proposed yields to a large-scale combinatorial problem that is directly solved as a mixed integer programming problem.

Keywords Rail planning, train timetabling problem, train scheduling, complex topologies

1 Introduction

Rail transportation has been defined as the mean to move passengers and freights by way of wheeled vehicles running on rail tracks. A railway system includes three essential elements: the infrastructure (the track work, the signalling equipment, the stations, and the catenary or third railway system with power supply on electrified lines), the rolling-stock (cars and locomotives) and the operational policies: rules and procedures for a safe and efficient operation. These three elements are highly interrelated.
There are two important agents in a railway system: the infrastructure manager and the operator. The infrastructure manager controls the infrastructure (maintenance strategy, overhaul, traffic control, etc.) whereas the operator deals with the demand and the trains.

There are different kinds of railway systems: underground, commuter and long-distance trains. Underground and commuter lines are characterized by relatively complex topologies operated at the maximum capacity. Long-distance lines are not usually so congested, but they present other difficulties: passenger and freight trains usually share the lines, some high-speed trains circulate in some lines, etc.

There can be two different railway management models, those in which the infrastructure managers and the operators belong to independent companies, and those in which both belong to the same company. One example of independent infrastructure managers and operators is the Spanish railway system, where the infrastructure is controlled by ADIF (ADministrador de Infraestructuras Ferroviarias), and the rolling stock by RENFE (REd Nacional de Ferrocarriles Españoles) and other operators that could compete for the use of the infrastructure. Other Spanish metropolitan railway systems (undergrounds and commuter lines) are controlled by smaller local companies: Metro de Madrid, FGC (Ferro carrils de la Generalitat de Catalunya), etc., so the infrastructure and the operation of this infrastructure are managed by the same company.

The objective of a railway system is to provide enough capacity to satisfy a demand. The demand is the passenger or freight transportation requirement. It is usually specified as a requirement by offering some services with a determined frequency. Railway transportation demand is increasingly rapidly. An increase of 25% in its activity is expected in Europe for 2030 (Capros et al. 2007). Because of that, the long term objective is to have enough capacity to meet this demand. But the rail sector is intensive in capital. Careful capacity analysis should be carried out before investing in a determined project.

The capacity has been defined as “a measure of the ability to move a specific amount of traffic over a defined rail with a given set of resources under a specific service plan” (Krueger et al. 1999). The main difficulty to determine the capacity of a railway line is that this capacity depends on several factors related with the infrastructure of the line, with the rolling-stock and with the operational characteristics such as:
• Main infrastructure factors are the topology (siding length, track structure, speed-limits, etc.) and the block and signalling system (length of subdivision, spacing and uniformity, intermediate signal spacing, etc.).

• The most important rolling-stock factors are the characteristics of the trains (train length, power to weight ratios, speed, braking power, etc.).

• Finally, some operational factors are the policies, i.e., the operation pattern of the line (timetables, frequency, the number of trains running at the line at peak hours, mix of trains, detailed schedules, etc.), the dispatching priorities, the quality of service (travel time, headways, required reliability or robustness), etc.

In addition, the formulation of the capacity as “…a measure of the ability to move…” indicates that several measures of the capacity can be considered. From the point of view of the demand, the capacity is usually expressed as the number of trains moving through a determined point of the line in a determined time window, the frequency of the services or the headways between consecutive trains running in the line (lower headways means higher level of capacity). Usually, the infrastructure manager is more interested in the saturation level of the infrastructure (overexploited, underexploited) measured as the amount of time that the signalling system is being used, for example.

1.1 Capacity planning problems

Different problems arise when the infrastructure and the services are planned to satisfy the demand and when the operation is adjusted on-line to follow the operation plan. In his annotated bibliography, (Assad, 1981) collected and classified several articles existing in the literature about network and timetabling models. After that, rail transportation problems have been classified into three levels according distinct planning horizons: strategic planning level, tactical planning level and operational or regulation level, see (Cordeau et al. 1998; Crainic 2003; Ghoseiri et al. 2004).

At the strategic level the demand is analyzed and the line is planned in order to offer a determined capacity when the infrastructure is being designed or to increase the capacity when part of the infrastructure is being renovated. The set of operational policies is also decided. At the tactical level, timetables are formulated. These timetables result on schedules for the services. After that, trains
and crew are assigned to lines (rolling-stock and crew scheduling and rostering). The main difference between tactical and strategic planning is that the investments in the rail network and in the trains are decided at strategic level while the tactical level decides how the available infrastructure and rolling-stock should be used. The level of detail in which the movements of the trains are considered at the tactical models is considerably higher than that at the strategic level. Finally, at the operational level, real-time data is used to decide how to modify the operation to keep offering, when possible, the planned services with enough quality. In the following table, the objectives and the time horizons for the three planning levels are presented.

<table>
<thead>
<tr>
<th>Planning stages</th>
<th>Time horizon</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategic level</td>
<td>5-25 years</td>
<td>Resource acquisition</td>
</tr>
<tr>
<td>Tactical level</td>
<td>1 month - 5 years</td>
<td>Resource allocation</td>
</tr>
<tr>
<td>Operational level</td>
<td>24 h-1 year</td>
<td>Daily planning</td>
</tr>
</tbody>
</table>

Table 1 Rail transportation planning levels

1.2 Train timetabling problem

Train timetabling problem involves finding a complete timetable (passing, arrival and departure times of each train at every point of the line) for a given mix of trains running under certain operational policy along a determined infrastructure. As a tactical problem, the objective is to determine medium term decisions to maximize the performance of the operations once the long-term decisions and operational policies have been adopted. Sometimes, the train timetabling problem arises also in the strategic or the operational planning level, but with different levels of detail.

The design can be performed satisfying different objectives:

- maximizing the capacity used by means of minimizing the period or maximizing train frequency,
- minimizing energy consumption by means of maximizing time slacks and using it for ecodriving, maximizing the use of regenerative braking energy, etc.,
- maximizing the robustness of the schedule by means of maximizing also time slacks,
• maximizing the quality of the service perceived by the passengers by means of minimizing dwell times, ensuring correspondences between trains, etc.,
• minimizing the rolling stock or crew requirements, etc.

Three different optimization approaches for the train timetabling problem can be found in the literature:

• The first approach models the problem as a multi-mode resource constrained project scheduling (Castillo et al. 2009; Ghoseiri et al. 2004; Zhou and Zhong 2005). In (Ghoseiri et al. 2004) a double-track train timetabling problem with a multicriteria objective function is considered. They try to reach a trade-off between energy consumption and other operation costs (operator objectives) and the timetable deviations (passenger objective). The detailed determination of energy consumption and cost proposed yields in complex non-linear equations that make the problem very difficult to solve. Zhou and Zhong (2005) consider a double-track train timetabling problem, bicriteria, with priority rules. They also try to reach a trade-off between operator and passenger objectives but using a mixed-integer linear approach. In (Castillo et al. 2009) a three-stage mixed-integer problem is proposed. First of all, a timetable is determined for minimizing the period in which the trains can be scheduled (maximize the capacity). After deciding the period, the slacks are maximized to obtain a more robust timetable. Finally, travel times are also maximized to obtain efficient timetables (allowing eco-driving).

• The second approach models the timetabling problem as a multicommodity flow problem (Caprara et al. 2002; Caprara et al. 2007; Cordeau et al., 1998). However, Powell et al. (2005) casted doubts on classical multicommodity flow models of fleet management because of its limitations for capturing realistic operation details even for strategic purposes. Powell described how an approximate dynamic programming paradigm makes it much easier to capture these details. The main trade-off of his model is that only an approximate optimum can be achieved over time.

• Finally, some contributions deal with simplified models or small instances, failing to incorporate general characteristics of real-life applications. There
are also papers that deal with other models as (Liu and Kozan 2009) where the train timetabling problem is modelled as a blocking parallel-machine job shop scheduling problem in order to use classical job shop solution algorithms, or with combinations of simulation and combinatorial optimization as (Assis and Milani 2004). Jovanovic and Harker (1991) presented a decision-support tool for designing monthly to daily schedules, or with other specific rail timetabling problems (take-overs in long distance problems, meet-pass planning, etc.). Sometimes, a trade-off between general models and solutions for specific cases is the best choice.

One limitation of these approaches is that the models are very dependent on the infrastructure of the line; very few of them consider that trains could have to accomplish different routes, very few of them consider the number of tracks and their topology, etc. Other limitation of the approaches found on the literature is that many of them need initial conditions of the line (position of the trains along the line, even initial timetables), so they usually fail to include general periodic operational policies, etc.

The train timetabling problem is known to be NP-hard, so optimal solutions are often unattainable in large-scale and complex instances (Zhou & Zhong, 2007). Consequently, although the size of the case studies presented in this paper can be managed with the algorithms already implemented in the solvers, it can be seen that for bigger sizes of the problem more research in solution algorithms is required. For real-time re-scheduling problems further research in solution algorithms is also required.

Furthermore, no comparison of the models is available, since the size of the problems are very dependent on the given infrastructure and operational policy. Because of that, the results in terms of number of trains scheduled at certain number of stations are not representative in order to compare one formulation with any other. Thus, different groups of researchers use different models instead of unifying efforts analyzing the best formulation.

Because of that, the main objective of this paper is to develop a train timetabling model able to design timetables for complex line infrastructures (topology and/or signalling system), able to design timetables for different operational policies (considering trains that accomplish different routes, periodic timetables, no need of initial timetables, etc.), able to design easily timetables for
different railway systems and comparable to previous models in terms of solution time and size of the model formulated.

2 New modelling of the train timetabling problem

In this section, a new formulation for modelling the train timetabling problem is presented. First of all, the requirements of a model coherent with the management of a railway system are presented. After that, the mathematical model is developed.

2.1 Definition of a railway line

The design of a railway line is conducted in different steps; the infrastructure manager is responsible of the topology of the line, the signalling system and the definition of the operational policy (rules and operational procedures). The operator is usually in charge of the acquisition of the rolling stock and should ask for the services required to meet a determined demand.

The first step is to define the topology of the line, the number of tracks of each element, etc. This decision is taken at the very beginning of the design of the line, and it is very difficult to change (although small changes can be conducted; new tracks, connections with other lines, etc.). Another big investment is the acquisition of the rolling stock, with its particular characteristic (length, speed, deceleration rate, etc.). When the topology of the line has been modelled, possible itineraries at each point of the line can be defined. The last step of the hardware of the railway system is the implementation of a signalling system able to authorize the movements along the line in a safety way.

Once the infrastructure and the rolling stock are defined, the operational policies should be determined. First, different train services should be defined. Each train service is associated to a determined route (topological path), a set of planned stops (every train of the route stops at the same stations) and to a determined kind of train. The operational pattern of the line should be defined (mix of trains, frequencies, etc.) as well as the boundary conditions. Finally, other operational constraints can be considered (correspondences, etc.).

The railway system has been modelled according to that structure. The information required and the objectives at each step are the following:
Step 1: Topology of the line
The main objective is to model the line splitting it into small standardized sections; nodes and connections between them. The choice of these standardized sections is a generalization of the sections proposed in (Pachl 2002; Landex 2009). The connections between the nodes are the inter-stations.

Seven kinds of nodes are required to model any topology:

- Terminal stations: Final station of a line. Every train must stop and turn-around.
- Turn-around stations: Intermediate stations of a rail line where trains can either turn-around or continue travelling in the same direction.
- Stations: sections of the line where the trains might stop.
- Overtakes: sections of the line with one track where a train can be stopped till other train travelling in the opposite direction or even a faster train travelling in the same direction overtakes it.
- Junctions: sections of the line connecting different inter-stations.
- Crossovers: sections of the line where trains can change from one track to another.
- Number-of-track change points: Sections of the line connecting two inter-stations with different number of tracks.

![Kind of nodes connected with inter-stations (line sections)](image)

Step 2: Signalling system
The objective is to determine pairwise incompatible itineraries at each section of the line. The information required at this step of the model is minimum safety headways for each kind of train travelling on the line, for each kind of event (arrivals, departures, passes) associated to each pair of pairwise incompatible itineraries.

Step 3: Train service
Train services are determined by:

- a commercial service (defined by a topologic path and by a set of planned stops)
- the kind of train offering the service
- the possible tracks of each section that these trains can use
Minimum and maximum stopping times and travel times are defined at this step.

**Step 4: Demand**

The objective of this step is to determine the number of trains offering some train service during the design time window (the frequency of each service) to meet a determined demand. Additional information required to schedule the trains is usually given in terms of:

- Operation patterns at some sections: defined by the order of events of the trains in a section during a determined period of time (time window).
- Other ways, as for example, minimum and maximum headway between two consecutive trains accomplishing the same train service in some sections of the line.

**Step 5: Boundary conditions**

Some information about the position of the trains along the line at the beginning and at the end of the time window is also required. This data can be provided:

- Defining the initial position of the trains at the beginning of the time window.
- Scheduling the first event of each train scheduled.
- Using periodic operation policies.
- Other ways, as for example, minimum and maximum headway between two consecutive trains accomplishing the same train service.

The link between one service of the train and the following one should also be defined. Dwell times for preparing the train for a new service should be considered. The possibility of having trains accomplishing different train services is decided at this step.

Further data such as the requirements of correspondences, etc. should be provided by the operator or by the infrastructure manager.

**2.2 Mathematical model**

The mathematical model formulates the topology of the line as a graph whose vertices represent the nodes and the arcs represent the connections. Each vertex and each arc have associated the number of tracks. The vertices have associated also the kind of vertex. The itineraries are modelled as successions of tracks of
nodes or connections. The routes can be modelled as a tuple of an itinerary, a kind of train, a determined plan of nodes where the train stops and a determined plan of nodes between which the travel time for the train should be determined. The information of the signalling system and of the operational policy is used to formulate the constraints of the model.

2.2.1 Notation and general formulation

In this section, the notation used in the paper is presented. Lower-case letters has been used for denoting parameters and upper-case letters for the variables. Binary variables are noted with Greek letters. The sets have been defined with lower-case letters too, but appear as sub-indexes at the constraints.

\textbf{Indexes}

The topology can be modelled with the following sets and two dynamic sets:

- \( i \) set of nodes (sta1, sta2, jct1, ...)
- \( a \) set of connections (ista1, ista2, ...)
- \( v \) set of tracks of a node or a connection (v1, v2, ...)
- \( ki \) set of kind of nodes (stations, junctions, ...)

The rolling stock can be modelled with the following two sets:

- \( j \) set of trains (T1, T2, ...)
- \( kj \) kind of trains (high-speed trains, freight trains, etc.)

And finally the routes are defined with an itinerary, with the kind of trains and with the set of stops and travel times.

- \( r \) set of train services

Other set and other dynamic set required for the formulation of the problem is:

- \( e \) set of possible events (arrivals, departures and passes)
- \( d \) direction of the train (downward, upward). This set is required to distinguish the events of the outward journey from the return journey in those cases in which the path of the outward journey and the return journey are the same.

\textbf{Parameters}

The topology of the line is defined with the incidence matrix \( MI_{v(i)v(a)} \).

\[
MI_{v(i)v(a)} = 1 \text{ if track } v(i) \text{ of node } i \text{ is connected to track } v(a) \text{ of connection } a.
\]

Two tables are required to define the kind of nodes and the kind of trains:
\( knode_{ki} \) definition of the kind of node; matrix with entry 1 if the node \( i \) is of kind \( ks \).

\( ktra_{kj} \) definition of the rolling-stock; matrix with entry 1 if the train \( j \) is of type \( kf \).

The train services can be defined with its topological path and the tracks at which each train accomplishing the train service may use (succession of tracks of every nodes and connections where the trains accomplishing the route pass):

\( \delta_{rvav'd}^- \) definition of the path; matrix with entry 1 if track \( v \) of node \( i \) precedes track \( v' \) of the connection \( a \) for direction \( d \) in the route of train service \( r \).

\( \delta_{rvav'd}^+ \) definition of the path; matrix with entry 1 if track \( v' \) of connection \( a \) precedes track \( v \) of node \( i \) for direction \( d \) in the route of train service \( r \).

And the set of planned stops:

\( stop_{rn} \) matrix with entry 1 if trains of the route \( r \) stop in track \( v \) of node \( i \). It is assumed that in nominal schedules the trains will not be stopped at an inter-station (connection).

The matrix \( asig_{rd} \) is used to define the assignation of the trains to the routes and has entry 1 if train \( j \) accomplishes route \( r \).

\( path_{jvid} \) auxiliary matrix with entry 1 if the route at which the train \( j \) is assigned passes through track \( v \) of node \( i \) with direction \( d \). It can be calculated using previous parameters: \( path_{jvid} = \sum_{r,a,v'} asig_{rj} \delta_{rvav'd}^- \)

Once the topology of the line and the train services have been defined, the dwell times required to design the schedule are:

\( h \) safety headway. It will be referred to maybe different events in maybe different tracks of one node, connection or distinct nodes or connections, the itineraries and the kind of trains. When two different itineraries are not pairwise incompatible, the constraint associated should not be formulated. Nevertheless, if it is formulated, the minimum safety headway should be equal to \( h = -m \), \( m \) being a big enough number.

\( st_{rvid} \) stopping time of trains accomplishing route \( r \) in track \( v \) of node \( i \) for direction \( d \). The following convention will be used to denote respectively, lower
and upper bounds of a parameter or variable: $s_{riv}$, $S_{riv}$. In this case, they denote respectively, lower and upper bounds of the stopping time.

$tt_{riv'v'd}$ travel time of a train travelling in route $r$ from track $v$ of node $i$ to track $v'$ of node $i'$ for direction $d$. This travel time could be different for different directions of the trains due to the grades of the line, etc.

**Variables**

The objective of this problem is to determine a train timetable, which is defined, as said in the introduction section, by the passing, arrival or departure time of each train at every point of the line. That is, the time at which every event of each train occurs at every point of the line. Thus, the decision variable will be:

$T_{vijd}$ planned time of event $e$ of train $j$ at track $v$ of node $i$ for direction $d$.

Additional parameters and variables can be defined to compute the energy consumed, the time window (period if the timetable is periodic), etc. in case they are required for the objective function.

Other binary variables will be defined when needed.

**Constraints**

As was earlier observed, there are two main groups of constraints: those which consider the signalling system (security constraints) and those which include the operational policy of the line. Some of them affect only one of the nodes or connections and some others condition the schedule at more than one node. In the following subsections, the main constraints of each kind of section are presented.

Again, additional constraints can be included in the model in order to compute the energy consumed or other variables required for designing the timetables according to energy efficiency criteria, robustness criteria, etc.

### 2.2.2 Security constraints

In this subsection some security constraints involving information of only one node are presented first. After that, those involving data of one connection are formulated. Finally, the structure of the constraints involving information of more than one node or connection is presented.
**Stations**

First of all, the headway between consecutive events at the same track must be greater that the minimum safety headway. Thus depending if the train $j$ or the train $j'$ are scheduled first, constraint (1) or (2) must be fulfilled:

$$ T_{ivejd} - T_{ive'jd} \geq h^1_{iivrrdd} $$  \hspace{1cm} (1)

$$ T_{ive'jd} - T_{ivejd} \geq h^1_{iivrrdd} $$  \hspace{1cm} (2)

where the safety headway corresponding to trains $j$ and $j'$ can be calculated as

$$ h^1_{iivrrdd} = \sum_{r,r'} h^1_{iivrrdd} \cdot \text{asign}_{rj} \cdot \text{asign}_{r'j'} . $$

The combination of events depends on the headway given by the signaller. If the minimum headway is referred to leave-entry (time distance between the instant in which a train leaves from the station and the instant in which another train can enter in this station), and depending if the train stops at the node or not (which depends on the route): $(e,e')=(\text{arr},\text{dep})$ if $\text{stop}_{riv}$ or $(e,e')=(\text{pas},\text{pas})$ otherwise. Similarly, $(e',e'')=(\text{arr},\text{dep})$ if $\text{stop}_{riv}$ or $(e',e'')=(\text{pas},\text{pas})$ otherwise. Analogously, headways referred to departure-departure, etc. can be formulated.

To model that one and only one of the constraints (1) and (2) holds a binary variable $\alpha^1_{iivjjdd}$ is required. Constraints (1) and (2) will be replaced by the following constraints:

$$ T_{ivejd} - T_{ive'jd} \geq h^1_{iivrrdd} - \alpha^1_{iivjjdd} \cdot m^1 $$

$$ T_{ive'jd} - T_{ivejd} \geq h^1_{iivrrdd} - \left(1 - \alpha^1_{iivjjdd}\right) m^1 $$  \hspace{1cm} (3)

where $m^1$ is a big enough number. This constraint will be formulated $\forall i \backslash \text{knode}_{i_{\text{station}}}, \forall v$, track of the node $i$, and $\forall d,d' \text{ directions and } \forall j < j'$ trains with $\text{path}_{jvid}$ and $\text{path}_{jvid'}$.

The headway between consecutive events occurring at different tracks should be greater also that a minimum safety headway. As before, this constraint can be formulated as:
\[ T_{w_{v'},j'd'} - T_{w_{v'},j'd} \geq h^2_{v',v'';d'd'} - \alpha^2_{v',v'';d'd'}m^2 \]
\[ T_{w_{v'},j'd'} - T_{w_{v'},j'd} \geq h^2_{v',v'';d'd'} - (1 - \alpha^2_{v',v'';d'd'})m^2 \]  \hspace{1cm} (4)

\[ \alpha^2_{v',v'';d'd'} \] being again a binary variable that activates one of the constraints (depending on the train that is scheduled before). \( m^2 \) is a big enough number. The constraint is formulated \( \forall i \setminus knode_{\text{station}}, \forall v \neq v' \) tracks of the node \( i \), and \( \forall d,d' \) directions and \( \forall j < j' \) trains for which \( \text{path}_{jv'id} \) and \( \text{path}_{jv'id'} \).

**Terminal stations**

For all node \( i \setminus knode_{\text{inter-station}} \) the following constraints are required:

Again, the headway between consecutive events at the same track must be greater that the minimum safety headway.

\[ T_{w_{v'},j'd'} - T_{w_{v'},j'd} \geq h^3_{v',v'';d'd'} - \alpha^3_{v',v'';d'd'}m^3 \]
\[ T_{w_{v'},j'd'} - T_{w_{v'},j'd} \geq h^3_{v',v'';d'd'} - (1 - \alpha^3_{v',v'';d'd'})m^3 \]  \hspace{1cm} (5)

\( \alpha^3_{v',v'';d'd'} \) being a binary variable and \( m^3 \) being a big enough number. This constraint will be formulated \( \forall i \setminus knode_{\text{inter-station}}, \forall v, \) track of the node \( i \), and \( \forall d,d' \) directions and \( \forall j < j' \) trains with \( \text{path}_{jv'id} \) and \( \text{path}_{jv'id'} \).

The combination of events depends again on the headway given by the signaller. The combination of directions depends on which direction is assigned to each event. Usually, arrivals to the terminal station are assumed to have the same direction of the train in the previous nodes and direction of the departures from the terminal station are assumed to be equal to the direction of the trains of the route in the following nodes. Every train stops at the terminal stations.

The headway between consecutive events occurring at different tracks should be greater also that a minimum safety headway. As before, this constraint can be formulated as:

\[ T_{w_{v'},j'd'} - T_{w_{v'},j'd} \geq h^4_{v',v'';d'd'} - \alpha^4_{v',v'';d'd'}m^4 \]
\[ T_{w_{v'},j'd'} - T_{w_{v'},j'd} \geq h^4_{v',v'';d'd'} - (1 - \alpha^4_{v',v'';d'd'})m^4 \]  \hspace{1cm} (6)

\( \alpha^4_{v',v'';d'd'} \) being a binary variable and \( m^4 \) being a big enough number. The constraint is formulated \( \forall v \neq v' \) tracks of the node \( i \), and \( \forall d,d' \) directions and \( \forall j < j' \) trains for which \( \text{path}_{jv'id} \) and \( \text{path}_{jv'id'} \).
If the terminal station has a cul-de-sac, consecutive arrivals and departs to the cul-de-sac will be controlled with constraints (5) and (6). In this case, it is usually assumed that the direction of the train changes after its arrival to the cul-de-sac.

**Turn-around station**

Security constraints at a turn-around station are equivalent to the security constraints of a terminal station. The are only two details that should be taken into account. First, the combination of directions depends on which direction is assigned to each event, although now there are trains which do not change their direction. Second, not every train should stop at a turn-around station, so some events could be passes of a train through the turn-around station.

**Junctions**

Security constraints at a junction are equivalent to those of a station. Usually the events will be passes of the train through the junction, since the trains rarely have a planned stop at a junction.

**Overtakes**

Again, security constraints at one overtake are equivalent to those of a station. In this case, usually the trains entering in some of the tracks will stop and will departure from the overtake after other train has passed.

**Crossovers**

Security constraints at a crossover are also equivalent to those of a station.

**Number-of-track change point**

When the number-of track change point is located after a station, there is no need to formulate this additional node, since the change of tracks has been already modelled in the station. If it is located in an inter-station, its security constraints are equivalent to those of a junction.

**Inter-stations**

The headway between consecutive arrivals or departures of a inter-station should be lower than the minimum safety headway of the inter-station. This constraint can be formulated as:

\[
T_{\text{inj}} - T_{\text{inj'}} \geq h^5 - \alpha^i_{\text{inj'}} \cdot m^5
\]

\[
T_{\text{outj}} - T_{\text{outj'}} \geq h^5 - (1 - \alpha^i_{\text{inj'}}) \cdot m^5
\]

\( \alpha^i_{\text{inj'}} \) being a binary variable and \( m^5 \) being a big enough number too. The constraint is formulated \( \forall d,d' \) directions and \( \forall j,j' \) trains, \( \forall v,v' \) tracks of the
nodes \(i, i'\) respectively from which trains are entering at certain track \(v'\) of the inter-station \(a\), that is: \(path_{jvid}, path_{jv'd'}\) and the trains are assigned to routes \(r, r'\) for which \(\delta_{riva'd'}\) and \(\delta_{riva'd'}^-\).

**Constraints affecting various nodes or connections**

Sometimes, security constraints affecting more than one station or inter-station should be formulated. The structure of these constraints will be:

\[
T_{ij} - T_{i'j'} \geq h_{ij} - \alpha_{ij} \cdot m
\]

\[
T_{i'j'} - T_{ij} \geq h_{ij} - (1 - \alpha_{ij}) \cdot m
\]

\(\alpha_{ij}\) being a binary variable and \(m\) being a big enough number. The constraint is formulated \(\forall v, v'\) tracks of the nodes \(i, i'\) respectively, and \(\forall d, d'\) directions and \(\forall j, j'\) trains for which \(path_{jvid}\) and \(path_{jv'd'}\).

**2.2.3 Operational constraints**

**Stopping time determination**

For every track \(v\) of every node \(i\) and every train \(j\) assigned to a route \(r\) with \(stop_{riv}\), the stopping time have to be bounded with the following parameters:

\[
st_{riv} \leq T_{i'v'd'} - T_{iverd'} \leq st_{riv}
\]

where the directions \(d, d'\) are equal if the train stops at a station or at a turn-around station and continues travelling without turn around. \(d, d'\) are different if the train stops at a terminal station or at a turn-around station and turns around.

That is:

\(d = d'\) if \(path_{jvid}\) and \(\sum_{v',a,r} path_{jv'ad'} \cdot asig_{n, j} \cdot \delta_{riva'd'}^-\).

\(d \neq d'\) otherwise.

That stopping time includes passenger stopping time (to get into/out of the trains), operational stopping time (to allow the drivers to change from one driving cab to the other), stopping time due to signalling system, slacks (to ensure some robustness at terminal stations).

**Travel time determination**

There are two travel times that should be determined. The first one is very important to ensure a continuity of the trains along their travel. It is a travel time
between consecutive nodes in a route. The second one is the commercial travel time offered to the passengers between important nodes (usually stations, terminal stations or turn-around stations) of the line.

In both cases, the travel time can be determined as:

\[ T_{ri\rightarrow iv} - T_{ri\rightarrow iv} \leq T_{ri\rightarrow iv} \leq T_{ri\rightarrow iv} \]

where the combination of events \((e,e')\) could be \((arr, dep)\) if the train stops in both nodes, \((pas, dep)\) if the train does not stop in the second node, \((arr, pas)\) if the train does not stop in the second node or \((pas, pas)\) if the train does not stop in the first, nor in the second node.

These constraints will be formulated \(\forall j: \text{asig}_{rj}\) for which the any of the bounds \(T_{ri\rightarrow iv}, T_{ri\rightarrow iv}\) is provided.

Recall that it is very important to have this lower bound between each pair of consecutive nodes \(i,v, i',v'\) of a route \(r\):

\[ \sum_{v',d} \text{asig}_{rj} \cdot \delta_{riv}^{-} \cdot \delta_{riv}^{+} \geq \delta_{riv}^{-} \cdot \delta_{riv}^{+} \]

2.2.4 Other constraints

In this subsection, some ideas about the constraints that can be formulated to formulate a periodic operational policy and some ideas about the objective function are presented.

Periodic operation

Sometimes, when the timetable design is a peak-hours timetable, the design of a periodic timetable is more flexible than the design of a timetable assuming the initial position of every train. In order to model this operational policy, the following constraint should be fulfilled:

\[ 0 \leq T_{iv\rightarrow iv} \leq P \]

where \(P\) is the period. It has been assumed that the period is a variable. Otherwise use \(p\).
But recall that every constraint controlling a dwell time or a headway should also be reformulated to ensure that fulfills when the period end. For example, constraint (3) will be reformulated as:

\[
T_{ivejd} - T_{ive'jd'} \geq h^i_{ive'dd'} - \alpha^i_{nij'dd'}m^i
\]

\[
T_{ive'jd'} - T_{ive'jd} \geq h^i_{ive'dd'} - (1 - \alpha^i_{nij'dd'})m^i
\]

\[
T_{ivejd} + P - T_{ive'jd'} \geq h^i_{ive'dd'} - \alpha^p_{nij'dd'}m^i
\]

\[
T_{ive'jd'} + P - T_{ive'jd} \geq h^i_{ive'dd'} - (1 - \alpha^p_{nij'dd'})m^i
\]

(13)

**Objective function**

As said in the introduction (section 1.2), different objective functions can be formulated depending on the requirements of the design. One possible objective function could be to minimize the period (time window) in which certain amount of trains can be scheduled in order to determine the maximum capacity of the line:

\[
\min P
\]

(14)

But usually the timetables are designed maximizing the slacks in order to obtain more robust timetables, or minimizing the energy consumption of the trains scheduled, etc. The problem can even be solved in different stages focusing on different objectives and taking different decisions at each stage.

**2.3 Implementation**

The train timetabling problem has been stated as a Mixed Integer Problem (MIP). The model has been written in GAMS 23.3. It has been solved by CPLEX 12.1, using a PC (personal computer) at 2.99 GHz with 3.21 GB of RAM memory running the Microsoft Windows XP operating system.

The size of the problem is very dependent on the topology. Usually, it increases with the number of sections \((I\) represents the total number of nodes and \(K\) represents the total number of connections), the average squared number of tracks of each node or connection \((V(I,K))\) and the number of trains \((J)\) squared too:

\[
O((I+K)(V(I,K))^2 \cdot J^2),
\]

because the number of security constraints ensuring safety headways between consecutive events at the same tracks or different tracks is normally far higher than other figures.
3 Results

In this section, an example of the modelling of part of the line of FGC is shown. After that, a case study with a simpler topology has been chosen to compare the solution time of this model with the solution time obtained with the formulation proposed in other papers in the literature.

3.1 Example of a complex-topology line modelling

In this subsection, some ideas for modelling the topology of a railway system are presented. The ideas are aimed at a case study of part of the commuter railway system of Barcelona (FGC). Fig. 2 shows part of the lines of FGC. This is an example of a complex railway system because of three reasons. First of all, six different routes share a big terminal station (Plaza de Catalunya) and some intermediate stations. Furthermore, there are two different kind of trains; short and long trains. Not every train can enter at every track (there are short tracks at some stations too). Finally, some of the trains turn around at two intermediate turnaround stations (Rubi and Universitat Autònoma) and one junction and two complex stations separate the traffic going to or coming from different destinations.

![Fig. 2 Part of the lines of FGC](image)

Some of the intermediate stations are relatively simple, see Fig. 3.a. But other intermediate stations have complex topologies since the trains accomplishing different routes go to different connections, see Fig. 3.b and Fig. 3.c.

![Fig. 3.a Node Provença. Simple intermediate station](image)
There are also terminal stations with relatively simple topologies, see Fig. 4.b and one terminal station with many tracks interconnected and a signalling system that should be taken into account when the timetable is design, see Fig. 4.a.

When the trains go to different inter-stations far from the station, a new node (a junction) is introduced in the model. See Fig. 5.
As was earlier observed, there are also two turn-around stations: Rubi and Universitat Autònoma, where the trains accomplishing the services S5 and S55 respectively turn around. The trains accomplishing the services S1 and S2 continue till Terrassa and Sabadell respectively.

The rest of the topology of this line is relatively simple. The connections are double-track inter-stations shared by trains accomplishing different routes. The rolling stock can be modelled with a set with two kinds of trains, short ones and long ones. This is important since minimum safety headways provided by the signaller will be different depending on the length of the train.

Usually, the demand is concreted with an operation pattern of the train services in Plaza de Catalunya.

The solution time for this problem was of nearly 6 hours. The size of the problem is approximately of 350 000 variables according to the estimation of the previous subsection for 30 stations \((I+K=59)\), 2 or 3 tracks per section in average \((V(I,K)=2,3)\) and 30 to 32 trains \((J=30,32)\). In Fig. 6 the real size of the problem solved in GAMS for different number of trains of the case study is represented.

![Fig. 6 Size of problems with the infrastructure of FGC for different number of trains](image)

**3.2 Comparison of the models**

One of the objectives of this paper is to develop a train timetabling model comparable to previous models in term of solution time and size of the model formulated. As was earlier observed, these parameters are very dependent on the
topology of the problem. Thus, the best way to compare this model with previous approaches is to solve both models in the same case study. Since no detailed case study are presented for the two main approaches (the multi-mode resource constrained project scheduling approach [Castillo et al. 2009; Ghoseiri et al. 2004; Zhou and Zhong 2005] and the multicommodity flow approach [Caprara et al. 2002; Caprara et al. 2007; Cordeau et al., 1998]), a small case study has been defined.

No comparison with a multicommodity flow approach has been carried out since the computation of pairwise incompatible arcs is carried out before the optimization and the efficiency of the solution of the problem varies significantly depending on how the computation is carried out. Furthermore, general procedures for developing this computation have not been found in the literature. The objective is to compare the model proposed (MP) in this paper with a multi-mode resource constrained project scheduling approach (MMRCPS) implemented based on the formulation of the paper of (Castillo et al. 2009).

The topology and the infrastructure of the line have been modelled according to Fig. 7. The size of the model can be increased easily just changing the number of trains and stations, by changing $n$. The topology is made up of $2n$ independent lines that meet in two junctions. All the traffic has to pass through two intermediate stations. Every inter-station is a double-track inter-station. Every station except the terminal ones are four tracks stations, though only two of them are used in this problem. $2n$ homogeneous trains accomplishing $n$ different routes with stops at every station of the route should be scheduled.

![Fig. 7 Topology and rolling stock of the comparison case study](image)

The stopping time at every station is bounded between 15 and 45 seconds and the travel time is bounded between 60 and 75 seconds. The safety headway between consecutive departures from one station – arrival to the same station is 30 seconds. The operation pattern at intermediate stations is known (in the upward direction the first arrival at the first station is that of the train coming from branch 1, the second one is that of the train coming from branch 2,…). The same policy is used for the second intermediate stations and the trains travelling downward. The
initial position of the trains is those presented in Fig. 7 and each train accomplishes always the same train service.

In the following figure, the solution time of both approaches (MMRCPS and MP) is shown. It can be seen that the MMRCPS approach implemented has lower computation times for small $n$ but for $n \geq 5$ the solution time obtain with the model proposed in this paper is lower.

![Figure 8 Comparison of the models solution times for the case study with different values of $n$](image)

In the following table a detail of two of the comparisons is shown. It can be seen that the size of the MP is always greater than the size of the MMRCPS implemented. Nevertheless, for big enough problems, the solution time of the model proposed is lower since the model has been strengthened: the security constraints are formulated for the nodes and for the connections, so the feasible region when for the relaxed problem is more constrained.

In future research, methodological analysis of the structure of both approaches will be carefully analyzed to determine which constraints strength the formulation and which approach should be used depending on the problem.

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Trains</th>
<th>Stations</th>
<th>Variables</th>
<th>D.Variables</th>
<th>Equations</th>
<th>Sol. Time</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMRCPS</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>282</td>
<td>152</td>
<td>553</td>
<td>0.078</td>
<td>272</td>
</tr>
<tr>
<td>MP</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>890</td>
<td>804</td>
<td>1693</td>
<td>0.312</td>
<td>940</td>
</tr>
<tr>
<td>MMRCPS</td>
<td>6</td>
<td>12</td>
<td>26</td>
<td>1994</td>
<td>1224</td>
<td>3577</td>
<td>13.266</td>
<td>149347</td>
</tr>
<tr>
<td>MP</td>
<td>6</td>
<td>12</td>
<td>26</td>
<td>6506</td>
<td>6252</td>
<td>12757</td>
<td>4.765</td>
<td>10449</td>
</tr>
</tbody>
</table>

Table 2 Detail of the models comparison
4 Conclusions

In this paper a new approach for modelling and solving the train timetabling problem has been developed. The model developed is able to design timetables for complex railway lines, considering many realistic details that usually appear in the infrastructure of railway system (complex topologies, different signalling systems), in the operational policy (trains accomplishing different routes and sharing part of the infrastructure, different initial conditions of the line, periodic timetables, etc.). Traditionally, specific models were designed for designing timetables for these complex but commons railway systems.

A procedure for modelling the railway system has been developed. With this procedure it is possible to change easily the infrastructure and the operational policies of the line, in order to design timetables for different lines without changing the general structure of the model. The topology has been model splitting the line into eight standards line sections defined in this paper. The mathematical model has been defined for each of these possible line sections of every complex-topology railway system.

Finally, the model has been compared with implementations of other method founded in the literature. The solution time of the mixed integer optimization problem proposed in this paper is comparable for small cases with the solution times obtained with a multi-mode constrained resource project scheduling approach implemented and outperforms in larger cases.

In future research, the possibility of using solution techniques presented in the literature should be analyzed.

References


