

# A multi scalable model based on a connexity graph representation

L. Gély, G. Dessagne, P. Pesneau & F. Vanderbeck  
*SNCF, Innovation and Research Department,  
University of Bordeaux I, France*

## Abstract

Train operations will be greatly enhanced with the development of new decision support systems. However, when considering problems such as online re-scheduling of trains, experience shows a pitfall in the communication between the different elements that compose them, namely simulation software (in charge of projection, conflict detection, validation) and optimization tools (in charge of scheduling and decision making). The main problem is the inadequacy of the infrastructure's monolithic description and the inability to manage together different description levels.

Simulation uses a very precise description, while the optimization of a mathematical problem usually does not. Indeed, an exhaustive description of the whole network is usually counter-productive in optimization problems; the description must be accurate, but should rely on a less precise representation. Unfortunately, the usual model representing the railway system does not guarantee compatibility between two different description levels; a representation usually corresponds to a given (unique) description level, designed in most cases with a specific application in mind, such as platforming. Moreover, further modifications that could improve performances or precision are usually impossible.

We propose, therefore, a model with a new description of the infrastructure that permits one to scroll between different description levels. These operations can be automated via dynamic aggregation and disaggregation methods. They allow one to manage heterogeneous descriptions and cooperation between various tools using different description levels.

This model is based on the connexity graph representation of the infrastructure resources. We will present how to generate corresponding mathematical models based on resource occupancy and will show how the aggregation of resources



leads to the aggregation of properties (e.g. capacity) that can be translated into mathematical constraints in the optimization problem.

*Keywords: modeling, optimization, railway operations, traffic management, infrastructure representation.*

## 1 Introduction

This paper defines an innovative way to represent infrastructure and a methodology allowing the use of different description levels. This is a theoretical prerequisite to any system, which will help experts to address the very heterogeneous problems encountered within the re-scheduling operations.

## 2 Classic representation

### 2.1 Origins of the classic representation

Railway studies arose during the 1970s. Planning problems have been treated since the 1990s and rescheduling is a rather recent topic of interest. The most important developments in the last decade are summarized in the surveys in [1–4]. One can have a look at [5] for earlier studies.

However, although many mathematical models and techniques are presented, modeling issues are scarcely debated. With the exception of some formal exercises, such as [6] and an interesting discussion on implicit choice of description and its consequences in [1], the importance of infrastructure representation has been barely mentioned before [7]. Indeed, most studies naturally re-use the same kind of representation designed for industrial purposes, where the adequacy of description strictly corresponds to (only) one application.

### 2.2 Examples

The range of representations goes from an exhaustive one, as in figure 1, to more synthetic representations, where only the main railway nodes and main lines remain, as is the case in figure 2.

### 2.3 A formal definition for the classic representation

Considering the most elementary description level, the infrastructure of the railway network consists of basic track sections (e.g. block or routes sections). These sections join at special points (switchings, joints, stations, etc). The real railway network is usually represented in the same way via an undirected graph (since directions are given by itinerary definitions and signaling; parts of the infrastructure are not dedicated to one-way usage).



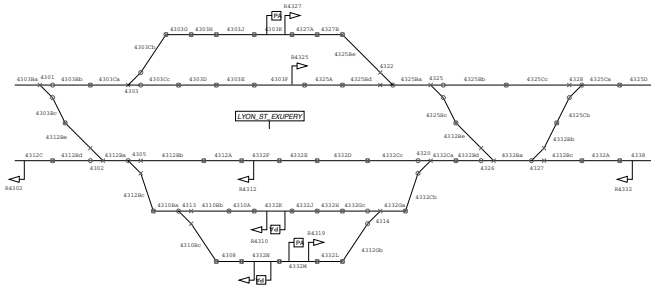


Figure 1: Lyon St Exupery TGV Station.

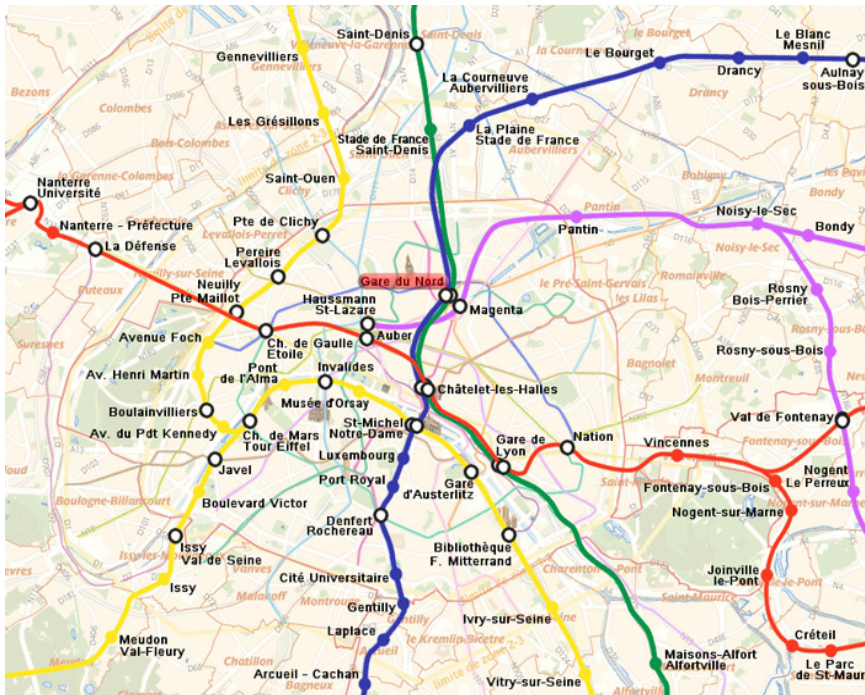


Figure 2: Railway network around Paris.

Thus, a classic representation on the field corresponds to the graph:

$$G_{\text{classic}} = (S, P), \text{ with } S = \{\text{arcs}\} = \{\text{Track Sections}\} \tag{1}$$

$$\text{and } P = \{\text{nodes}\} = \{\text{Special Points}\}$$

If the intention of the description is to be exhaustive, as shown in figure 1, arcs correspond to block sections and nodes represent limits between adjacent blocks. In the case of figure 2, on the other hand, arcs correspond to complete

lines (and might include some (minor) stations), whereas nodes represent stations or junctions.

## 2.4 Microscopic versus macroscopic

### 2.4.1 Why are there different description levels?

Railway infrastructure descriptions may require different levels of details, each one corresponding to different objectives, cf. [8]. The following set of online problems (more detailed in [9]) illustrate studies of different size and requiring different precision:

1. fluidification of a complex junction with speed fine-tuning around a limited local area.
2. re-scheduling with intermediate precision of the representation, but modeling interactions between surrounding areas.
3. re-routing of trains along new itineraries (succeeding a major breakdown, for example) within a macroscopic description of vast areas (typically involving different lines).

The bigger the area to consider (i.e. spatial distribution and time window of the incident consequences), the less precision in the description (although more time is usually available for computation). Indeed, in practice, if bottleneck areas require precise description, an exhaustive description of the whole network with maximum precision may be counter productive, especially with online applications, since calculations are usually exponential in terms of the number of elements (and mathematical variables).

On the other hand, when experts build a timetable off-line, solutions are mostly a guideline; only a moderate level of precision is required. However in operations, the solution must be immediately applicable then precise enough to ensure real feasibility. Consequently, the online rescheduling problem usually requires a more precise description, for the same area.

In other words, there is always some kind of trade-off between accuracy, size and available time. This trade-off is hard to balance with a one-size-fit-all description.

### 2.4.2 Definitions of macroscopic and microscopic representations

A microscopic description is a representation where all the elements correspond to the most basic resources; only one train can be affected on each one of them, e.g. a block section.

A macroscopic description contains elements that can be aggregations of basic elements. Resources do not necessarily have the capacity of one train. For instance, a resource can represent, as an example, a set of (connected) platforms, block sections, or even a whole complex station.

Remarks:

- a description is not necessarily homogeneous: several parts can be described at macroscopic level, others with more precision, i.e. with more disaggregated resources. Usually, lines correspond to the first one, while junctions, stations and other bottleneck areas correspond to the latter, e.g. [10, 11].



- the macroscopic description covers every level above the elementary microscopic one. Thus, every resource aggregation in a macroscopic description can be disaggregated or aggregated again into other aggregated resources.

## 2.5 How to choose a good description level?

### 2.5.1 Academic point of view

Given a specific re-scheduling problem, with a given (type of) incident and a given size of area to cover, one is able to choose an adequate description level. Hence, most algorithms use a dedicated (unique) description; some studies consider a representation of infrastructure based on the most detailed (microscopic) level: block sections level on a small area cf. [12]. However, studies consider more often more simplistic (macroscopic) descriptions [13, 14] on a larger scale; usually a main line joining two main stations with some stops (smaller stations) in between. Once again we refer to [1] for a comparative study on the size (and the precision level) used in the main recent studies.

In practice, the railway infrastructure description is done by human experts once and for all (nowadays such description takes days for every new study) and hence is unique. However, one cannot assume that some rules of thumb, even combined with expertise, could determine an average adequate description level that fits all incidents and covers the wide range of problems, such as those previously presented (i.e. from fluidification to re-routing problems). In conclusion, it is hardly suitable in practice and a generalization would be particularly uncertain.

### 2.5.2 Operational point of view

As previously explained, post-optimization validation (via simulation) requires the most precise description level. Consequently, the whole process uses at least two description levels, namely one for a (microscopic) simulation tool and one for a (more macroscopic) simulation tool. Needless to say, specifying (off-line) a microscopic description is unavoidable; however, we should not expect experts to provide other (every new macroscopic) descriptions all over again from scratch. Anyway, one must ensure that cooperation between at least two models would be possible.

Moreover, nowadays when real incidents occur, the impact of consequences can be hardly predictable. That is why forecasting tools are needed in the near future to help analyze, a priori, an acceptable trade-off between, on the one hand, precision and size of the description, and, on the other hand, calculation time. However, until the very end of the process, any choice will remain uncertain. We claim that any predefined fixed description level is very restrictive for optimization purposes and will probably be inefficient in many cases.

Consequently, an automatic or semi-automatic scalable representation would be of great interest if one can rely on an available microscopic description. Moreover, if we assume we can scroll easily from one level to another, why not use it dynamically in the search process itself? Finding a good trade-off would become part of the process.



We can easily imagine that switching to any level of aggregation, starting from the elementary microscopic description to very macroscopic ones, would be more suitable to deal with unpredictable incidents (and their impacts, i.e. the area to consider and the needed level of precision). This would also be easier to adapt facing the specific topology of infrastructure. To conclude, it would be more flexible and easier to generalize to an industrial tool.

### 2.5.3 Other modeling and mathematical issues

Some signaling systems do not use fixed blocks, but moving blocks where trains must respect permanent headways (e.g. ERTMS 3 signaling system). Here, an elementary resource based model is inappropriate. Accordingly, in the most general case of study, any description should be able to manage a mix between basic resources with the capacity of only one train and more aggregated resources.

Using different kinds of representations could help one to tackle some complexity issues with the mathematical problems in a counter-intuitive way, e.g. in an area with a complex network of switchings (such as the pre-entrance of main stations), a microscopic description (i.e involving resources with the capacity of one train) may lead to a more compact formulation (with regard to the number of constraints) than managing the whole set of incompatibilities between itineraries. In the first case we can aggregate occupancy (blocking) constraints (one constraint per resource), but in the second case we must deal with every couple of incompatible routes.

## 2.6 Limits of the classic representation

We will now describe why the classic representation does not conveniently fit the above requirements.

The classic representation allows elements of different nature to share the same kind of representation. In consequence, coherence may be broken, and aggregation or disaggregation operations may be difficult in practice. This will be illustrated in the following example.

In graph A, there are nodes of different nature; the gray nodes represent stations (hence infrastructure resources, such as the aggregation of platforms and ways), whereas the clear node represents only a special point (a virtual landmark that does not stand for a physical resource): the junction between three main lines, as illustrated in graph B.

Finally, graph C shows an additional representation where the same network is divided into two parts: one is composed by the high-speed line between Bruxelles and Nimes (dark color), the second is composed by the classic line between Nimes and Perpignan (clear color). The arc representing the high-speed line would correspond to an aggregation of all the nodes (i.e. a spacial point and the stations) plus the arcs (lines) that constitute the high-speed line.

All three representations would make sense from an operational point of view, with three different purposes in mind. However, if they correspond to different macroscopic description of the same network, one cannot define a common rule



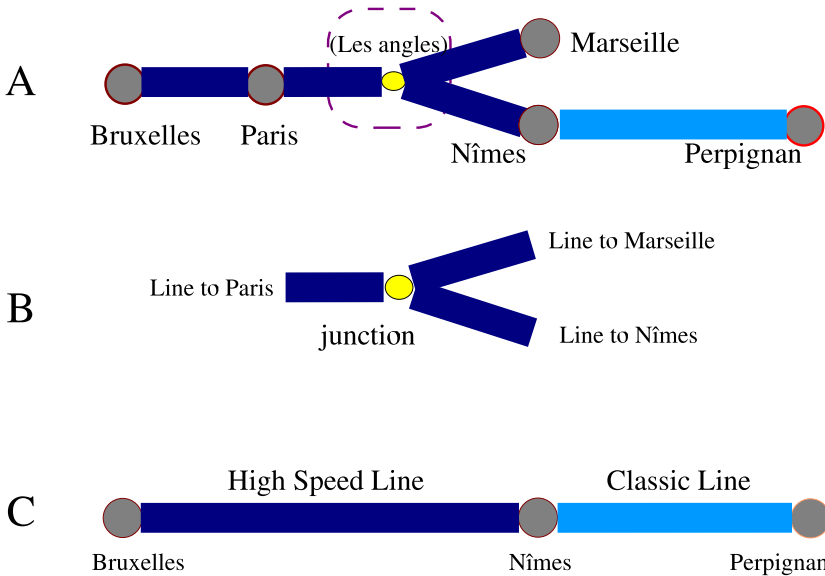


Figure 3: Illustration: TGV 9835 itinerary.

that would permit one to aggregate or disaggregate resources and then switch from one representation to another conveniently.

In practice, as we expect to implement aggregation methods, we must ensure that the following coherence rule is respected: *all elements of the same nature should share the same kind of representation* (regardless of the description level).

While classic representation is valid with a microscopic description (e.g. the exhaustive description of Lyon Saint-Exupery station, cf. 1), it cannot describe some aggregated resources and respect the coherence rule. Hence, if arcs represent sections of ways and nodes represent remarkable points, how can a station be represented in a macroscopic description? On the one hand, since a station may be connected with more than two resources (unlike arcs), one would need to represent it as a node. On the other hand it is of a different nature to a remarkable point; macroscopic resources (like a station) are aggregations composed of resources that are ways or route sections (platforms, etc . . .). They are not remarkable points (virtual landmarks); minimal duration constraints must be applied on every train crossing this resource (as for any physical resource).

In conclusion, if we want a model that respects the coherence rule (hence allowing easy implementation of an object-oriented model), the resources route section and station must share the same kind of representation, since they are of the same nature. As an aggregation of resources (such as a station) can be connected to many other resources, the most natural choice is to represent every infrastructure resource used by trains (such as ways, platform, station, . . .) with a node.

### 3 A new model representing the infrastructure

#### 3.1 Definition for a new representation

Every graph representation of the infrastructure consists of the two following elements:

1. nodes represent infrastructure resources,
2. arcs represent connections between resources.

Thus, the railway network representation with this new model is a connectivity graph:

$$G_{\text{connectivity}} = (R, C), \text{ with } R = \{\text{nodes}\} = \{\text{Infrastructure Resources}\} \quad (2)$$

$$\text{and } C = \{\text{arcs}\} = \{\text{Connectivity Relationship}\}$$

Such a formalism allows one to adapt infrastructure representation to any description level if we use an appropriate methodology.

For example, the following illustration shows four representations of a small network surrounding a junction (involving height track sections). We present, from left to right: the microscopic classic representation, its new representation, an example of aggregation (the closest four resources around the junction) and finally an equivalent of this aggregation if we had used the classic representation.

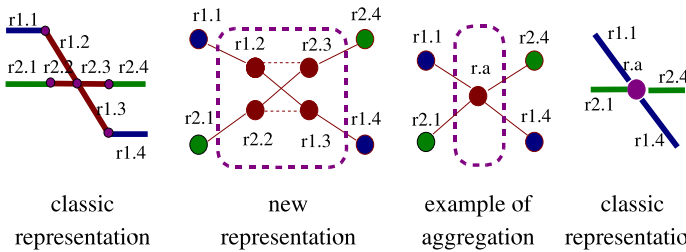


Figure 4: New representation as a connectivity graph.

Remarks:

1. If we compare the central node in both classic representation: in the disaggregated version it represents a landmark, while in the aggregated version it has become an infrastructure resource (aggregation of four sections),
2. Every arc in the classic representation becomes a node in the new representation,
3. For every relationship of a connectivity (arc), one can define a unique measure point (frontier) between two resources, as we will see later.



### 3.2 Methodology and rules of aggregations

The next question is how to treat aggregations and consequently disaggregations? Some basic rules must be pre-defined in accord with experts, once and for all. They must define how properties of aggregated resources, and hence, mathematical constraints, will evolve in an aggregation of resources, i.e. define the aggregation functions to implement.

These rules must be in relation with the properties and operational constraints one wants to model and what is considered to be relevant in regard of the addressed problem.

We will illustrate this methodology with two obvious examples of aggregations where one wants to address a capacity problem, hence one considers mainly maximum flow and maximum storage capacity properties (plus a sequence property). Of course, other aggregations of properties should be eventually defined following the same kind of methodology.

#### 3.2.1 Serial aggregations, itineraries

- aggregation of infrastructure resources: serial aggregation is very close to the concept of itineraries; this is the aggregation of an ordered list of infrastructure resources (with a maximum flow property and a static capacity property).

e.g. in figure 5, where:  $r_a = \bigcup_i r_{2.i} = \langle r_{2.1}, r_{2.2}, r_{2.3} \rangle$

- aggregation of measure points: measure points (i.e. arcs, noted frontiers) connecting the aggregated resource to the adjacent resources are the same as before aggregation.

e.g.  $Fr(r_1, r_a) = Fr(r_1, r_{2.1})$  and  $Fr(r_a, r_3) = Fr(r_{2.3}, r_3)$ .

- Properties:

1.  $Flow(r_a) = \min_i Flow(r_{2.i})$

2.  $Capacity(r_a) = \sum_i Capacity(r_{2.i})$

3. re-ordering is not allowed: the sequence of entrance remains strictly the same for the clearance (with the next trains).

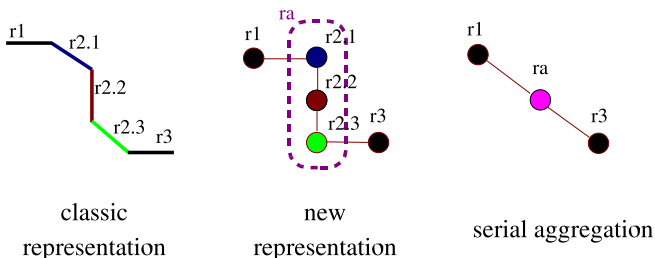


Figure 5: Serial aggregation.

**3.2.2 Parallel aggregation**

1. aggregation of infrastructure resources: the best example consists of the aggregation of parallels ways that compose a line (with a maximum flow property) or a storage (with a static capacity property).

e.g. in figure 6 :  $r_a = \bigcup_i r_{2.i} = \begin{cases} r_{2.1} \\ r_{2.2} \\ r_{2.3} \end{cases}$

2. aggregation of measure points: point of measure (arcs) should be melted.

e.g.  $Fr(r_1, r_a) = \bigcup_i Fr(r_1, r_{2.i})$  and  $Fr(r_a, r_3) = \bigcup_i Fr(r_{2.i}, r_3)$ .

3. properties:

(a)  $Flow(r_a) = \sum_i Flow(r_{2.i})$

(b)  $Capacity(r_a) = \sum_i Capacity(r_{2.i})$

- (c) re-ordering is allowed: the sequence of entrance is not necessarily the same as for the clearance.

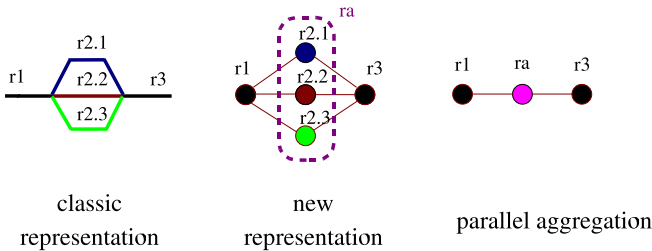


Figure 6: Parallel aggregation.

**3.3 Conventions regarding schedules**

Finally, we must define how to connect explicitly schedules with the previous representation.

**3.3.1 Convention**

Each arc in  $G_{\text{connectivity}}$  represents a unique point of measure. At this point we evaluate when the head of any train crosses the frontier between two resources (i.e. any effective entrance in a new resource). In order to construct a timetable, one horary (one variable) must be associated with the pass of every circulation on any measure point. This allows every move of every circulation to be described and a timetable to be associated with any graph  $G_{\text{connectivity}}$ .

Needless to say, every aggregation of resources yields directly to a mathematical model with fewer variables. This permits one to determine a trade-off between



precision (the more disaggregated, the more precise), size and calculations (the more aggregated, the faster the computation).

### 3.3.2 Measure of entrance and resource occupancy

The length of trains must not be neglected. One measures the date when the head of a train crosses the measure point; according to the length and the speed of the train, the whole train may completely leave the resource a long time after this measure (especially in the case of freight trains).

A circulation occupies a resource until **liberation** (or clearance), which happens a certain amount of time after its effective exit of the resource. In the same way, the resource occupies a certain amount of time called **reservation** before the effective entrance. The headway between circulations is then the sum of these amounts of time plus a buffer time (see blocking time theory, in [7], for example).

### 3.4 About the choice of a mathematical model

Once a multi scalable representation, as detailed here, is available, any mathematical model reviewed in [1] can be applied (adapted) on. The alternatives depend on what kind of operational problem is treated, and the types of operational constraint to consider are those that are more convenient and efficient, but in the end it should not depend on the representation nor the convention proposed.

On the other hand, different operational problems can be addressed (each with a different level of representation and a different mathematical model), as soon as a microscopic description based on this multi scalable representation is available.

## 4 Conclusion

We have defined a methodology and a representation that permits one to scroll from microscopic to any aggregated modelization. We have shown basic examples of aggregation rules that make automated aggregation possible. Finally, we have defined a convention for schedule that allows one to address the model of timetabling problem (and rescheduling problem).

Another aim of this paper was to explain why a complex software system will be needed to help online operations efficiently. We are convinced that a multi-level capable model will play a key role and is the first theoretical prerequisite towards their development.

Another prerequisite is a microscopic digital description of railway infrastructures; this would mark the entrance of railway operations in the digital age. Models of description and numerical databases are now under development in Europe, e.g. RailML [15] or Eifel (the SNCF dataset that will be compliant with the present concepts). However, compatibility issues could arise. Consequently, we encourage anyone to consider this new representation, which should enhance compatibility (at least with descriptions that are not of the same level).

Finally, as developed through an object-oriented mind, this model can be applied to any traffic management problem (involving resource allocation), and is probably



also applicable to a more general set of industrial problems outranging the scope of the railway industry, exclusively treated in this paper.

## References

- [1] Törnquist, J., Computer-based decision support for railway traffic scheduling and dispatching: A review of models and algorithms. *5th Workshop on Algorithmic Methods and Models for Optimization of Railways*, eds. L.G. Kroon & R.H. Möhring, Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany, 2006.
- [2] Caprara, A., Kroon, L., Monaci, M., Peeters, M. & Toth, P., *Passenger Railway Optimization*, Elsevier, volume Transportation, 2006.
- [3] Cordeau, J.F., Toth, P. & Vigo, D., A survey of optimization models for train routing and scheduling. *Transportation Science*, **32(4)**, pp. 380–404, 1998.
- [4] Bussieck, M.R., Winter, T. & Zimmermann, U.T., Discrete optimization in public rail transport. *Mathematical Programming*, **79(1-3)**, pp. 415–444, 1997.
- [5] Assad, A.A., Models for rail transportation. *Transportation Research Part A: General*, **14(3)**, pp. 205–220, 1980.
- [6] Bjørner, D., *The Domain Book: A Compilation of Reports and Papers on Domain Models*, Technical University of Denmark, chapter Railways, pp. 157–185, 2007.
- [7] Hansen, I. & Pachl, J., *Railway Timetable and Traffic. Analysis, Modelling, Simulation*. Eurailpress, 2008.
- [8] Lindner, T. & Zimmermann, U., *Mathematics-Key Technology for the Future: Joint Projects Between Universities and Industry*, Springer: Berlin, chapter Train Schedule Optimization in Public Rail Transport, pp. 703–716, 2003.
- [9] Gély, L., Real time train rescheduling at sncf. Robust planning and Rescheduling in Railways, 2007.
- [10] Burkolter, D., Herrmann, T. & Caimi, G., Generating dense railway schedules. Advanced OR and AI Methods in Transportation, Publishing House of Poznan University of Technology, 10th EWGT Meeting and 16th Mini-EURO Conference, pp. 290–297, 2005.
- [11] Caimi, G., Burkolter, D., Herrmann, T., Chudak, F. & Laumanns, M., Design of a new railway scheduling model for dense services. ISOR, 2007.
- [12] Brannlund, U., Lindberg, P.O., Nou, A. & Nilsson, J.E., Railway timetabling using lagrangian relaxation. *Transportation Science*, **32(4)**, pp. 358–369, 1998.
- [13] Carprara, A., Fischetti, M. & Toth, P., Modeling and solving the train timetabling problem. *Operations Research*, **50**, pp. 851–861, 2002.
- [14] Carprara, A., Monaci, M., Toth, P. & Guida, P.L., A lagrangian heuristic algorithm for a real-world train timetabling problem. *Discrete Appl Math*, **154(5)**, pp. 738–753, 2006.
- [15] RailML.org, <http://www.railml.org/>.

