

Analytical Robustness in Single-Line Railway Timetabling

Miguel A. Salido, Federico Barber and Laura Ingolotti

Departamento de Sistemas Informaticos y Computacion
Universidad Politecnica de Valencia, Spain
Email: {msalido, fbarber, lingolotti}@dsic.upv.es

Abstract: Over the last few years, numerous approaches and tools have been developed to compute railway timetabling. However, robust solutions are necessary to absorb short disruptions mainly in single-line tracks. In this paper, we present some guidelines to measure robustness in railway timetabling. Two analytical formula have been developed to measure robustness based on the study of railway infrastructure topology and time supplement/buffer time. Thus, each buffer time is pondered by some factors such as the weighting of each station in a running map, tightest tracks, number of subsequent trains, etc. This method is inserted in MOM¹, which is a project in collaboration with the Spanish Railway Infrastructure Manager (ADIF).

Keywords: railway timetabling, robustness, single-line, analytical measures.

1. Introduction

Railway transportation has played a major role in the economic development of the last two centuries. It represented a major improvement in land transport technology and has obviously introduced important changes in the movement of freight and passengers. Over the last few years, railway traffic has increased considerably, which has created the need to optimize the use of railway infrastructures and to maintain robustness in railway timetabling. Thanks to developments in computer science and advances in the fields of optimization and intelligent resource management, railway managers can optimize the use of available infrastructures and study important railway features such as robust timetables and railway capacity.

The term "robust" refers to the ability to resist to "imprecision".

In many decision processes it is needed to offer solutions with a certain level of robustness in order to maintain its feasibility in frameworks with incomplete or imprecise data. These data can be:

- Actual data or knowledge on the problem, which can not be known at the decision moment with the desired detail and certainty level.
- Future data, in a dynamic problem, since evolution of the problem is not exactly known in advance.

The idea of robust schedules consists in solutions that can tolerate a certain degree of uncertainty during execution. In other words, they should be able to absorb dynamic variations in the problem due to both external reasons (exogenous

events), and internal reasons (false definitions in the problem) [16].

Our aim is to obtain a measure of robustness in railway timetabling. A timetable can have the characteristic that delayed trains will lead to considerable knock-on effects, whereas another configuration of the timetable may be able to absorb such effects more readily. Ensuring that the service can recover quickly when one train is delayed not only makes traveling more reliable but also means that train operators can in some cases avoid substantial penalties. To this end, we have identified the main parameters that are directly related to robustness. Some of them have been identified by railway operators based on their experience. These parameters have been inserted in some analytical methods to give us a measure of robustness. Furthermore, these methods have been compared with a simulation method to validate the obtained measures. Fig. 1 shows our study of robustness in railway timetabling. Given a real timetable, we obtain analytical measures of robustness. These measures give us a model of robustness that we can use to identify new criteria in optimization processes. Furthermore, we insert random delays in the real timetable in order to replan and study an empirical measure of robustness. Both analytical and empirical measures will be checked in order to validate the proposed analytical methods. This study gives us the robustness degree of the real timetable.

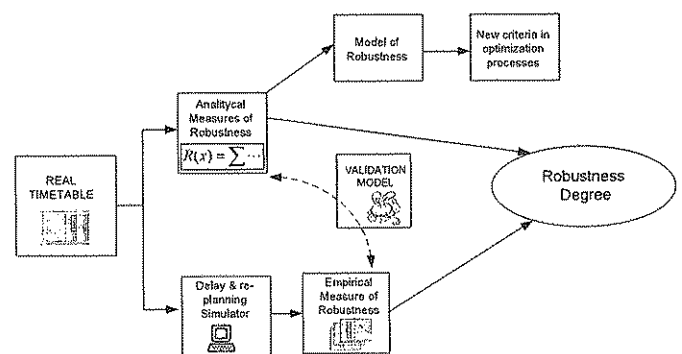


Fig. 1. Robustness degree in railway timetabling by checking analytical and empirical methods.

In this paper, we present robustness in two different frameworks: in section 2, robustness in general scheduling is presented; sections 3 shows some works related to robustness in railway scheduling. We present some factors that are directly related to robustness in railway scheduling in section 4. To measure robustness, we briefly present in section 5 some methods such as stochastic programming, analytical methods and

¹ More information can be found in MOM web page <http://www.dsic.upv.es/users/ia/gps/MOM>

empirical methods. Section 6 shows some new definitions of robustness and some analytical approaches to measure robustness in railway timetabling. Finally in section 7 we present some conclusions and further works.

2. Robustness in Scheduling

Research efforts aimed at generating solutions and quality robust schedules in combination with an effective reactive scheduling mechanism are still in a burn-in phase.

Even though different approaches have been pursued so far, the concept of robustness for scheduling solutions, as well as in other areas, remains vague and not well defined. In fact, it is possible to note different definitions of robustness in scheduling with respect the different aspects that are taken into account. Some of these definitions have emphasized the ability to preserve some level of solution quality. For instance, in [13], the robustness of a solution is defined with respect its ability to preserve the solution quality, that is, the completion time or makespan.

An intuitive approach to obtain robust schedules consists in adding redundancy to the solution. A first example is represented by [13] where a genetic algorithm for producing robust schedules in the case of job shop problems is described. The authors define an evaluation function, which is used in the algorithm to synthesize robust solutions, according to the actual makespan of the schedule during the execution and the schedule delay. A typical technique, from the general scheduling theory, consists of introducing time slacks in the execution of the tasks [7]. Another alternative is to build an explicit set of complementary solutions and in each moment to use the most suitable solution according to the current state [8]. The method proposed in [19] and [17] consists in the creation of a partial order schedule that provides a certain degree of robustness due to the fact that maintains temporal flexibility.

Nicola in [16] aims at increasing the robustness by introducing flexibility in the scheduling generation phase. Flexible solutions consist of a set of possible schedules that can be followed during the execution and, at the same time, that are guarantee of an easy and fast recovery of the current situation if necessity. In fact, they are concerned with the generation of schedules that offer some degree of robustness in the face of a dynamic and uncertain execution environment.

In [20], authors provide a distinction of robustness according to the following two characteristics: quality robustness and solution robustness. The first is a property of a solution whose quality does not change much from the initial one when small changes in the problem occur. The latter, instead, occurs when in the same situation (small changes) a solution does not deviate much. Then, the authors conclude that the two robustness concepts can be viewed in two different spaces: the objective function space and solution space respectively for quality and solution robustness.

3. Robustness in Railway Timetabling

As we pointed out before, there exist some definitions about robustness in scheduling and also robustness in railway timetabling, which are understood in many different ways. Norio in [21]

gives another definition of robustness: "A timetable is robust if we can cope with unexpected troubles without significant modifications". Furthermore, this author consider several robustness indexes for each level (see Fig. 2). In most cases, the main disruptions occur in level 0 due to little disruptions. Higher levels do not occur frequently and they need railway rescheduling.

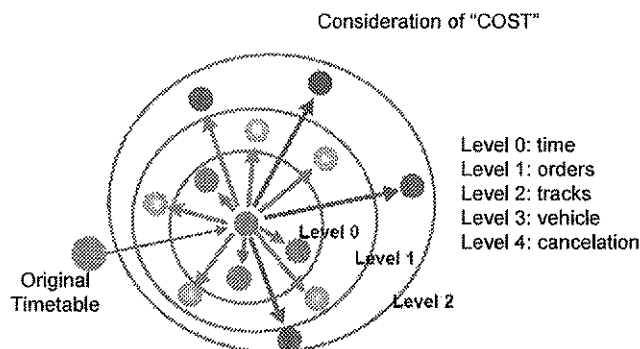


Fig. 2. The main levels of robustness.

In [4], Carey describes various heuristic measures of stability that can be employed at early planning stages. In [5], Carey and Carville present a simulation model used for testing schedule performance regarding the probability distribution of so-called secondary delays (knock-on effects) caused by the primary delays, given the occurrence of these and a schedule. The model is used for evaluating schedules with respect to the ability to absorb delays. The super-models were introduced in [9] as an scheme to measure the stability degree of a solution. This idea is extended to constraint satisfaction problems in [10] with the super-solution concept.

Nevertheless, due to the fuzzy definition about robustness in scheduling in general and in railway timetabling in particular, some simulation studies have been generated to evaluate robustness. In [22], Vroman, Dekker and Kroon present concepts of reliability in public railway systems. Using simulation, they test the effect of homogenizing lines and number of stops in timetables. Mattsson [14] presents a literature study on how secondary delays are related to the amount of primary delay and the capacity utilization of the rail network. Hoogheimstra and Teunisse [11] presents a prototype of a simulator used for robustness study of timetables at the Dutch railway network. The simulation prototype is called the DONS-simulator and is used for generating timetables. Similarly, in [15] Middelkoop and Bouwman present a simulation model, Simone, for analysing timetable robustness. The model simulates a complete network and is used to identify bottlenecks.

Railway timetabling is one of the main stages in railway management (see Fig. 3), where many feedback loops may be necessary:

Particularly, the timetabling generation phase has two inter-related steps:

1. *Operator phase*: This process is performed by railway operators in order to obtain optimal timetables for their own trains with the objectives of: to minimize waiting times among passenger connections, optimize rolling-stock and crew scheduling, optimize customer requirements, etc. This is a user-oriented timetabling.

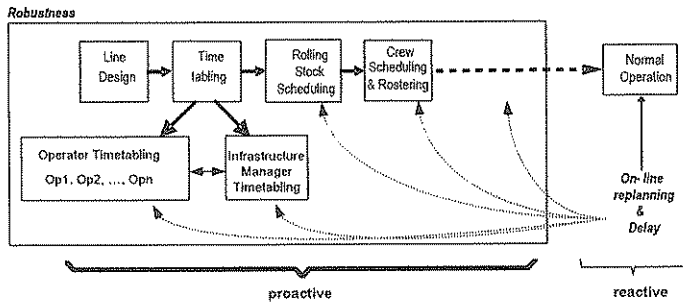


Fig. 3. Main railway management phases

II. *Infrastructure manager phase*: The request of all operators should be optimized satisfying robustness, operational and traffic constraints with the goal of optimizing service demands.

The concept of robustness should be considered in relation to each one of the above phases. In general, there are two approaches to prevent and manage the incidences that can occur in execution phase of any plan:

I. *Proactive*: it is directly related to the concept of robustness. The goal is to obtain robust plans. A commonly admitted definition of robustness is: "A robust plan is a plan that, given typical incidences in the environment, maintains its feasibility and as much as possible the optimality of the initial solution, at least with few meaningful changes". However, due to the fuzzy definition, some questions are straightforward:

- Which are 'typical incidences'? According to railway manager experience, it is very difficult to characterize causes of incidences and disruptions, except for the fact that the most common disruptions are between 1 and 5 minutes long.
- Which are 'few meaningful changes'? They are needed changes to restore the plan to the initial plan. Then, possible measures of robustness could be the required time interval until the initial plan is restored, or how many trains are affected, or the average delay of the trains.

II. *Reactive*: it is usually performed on-line when incidences can not be absorbed by a robust plan. Therefore, the goal of this approach is the re-planning of the previous plan according to the incidences in order to minimize the effects of incidences such as secondary delays, and recover to the normal plan/operation as soon as possible. Thus, the reactive approach is related to the on-line re-planning and delay management concepts.

4. How to Obtain Robust Timetables

Several methods can be applied in order to obtain robust timetables. Some of them are:

4.1 To Decrease Optimality

The robustness of the timetable against small disturbances can be improved by optimally allocating time supplements and buffer times in the timetable. This is a frequent method that implies adding additional buffer times in the travel times of

each train in each track section of its journey. For instance, Fig. 7 shows a real timetable for a train in ADIF.

Estación	Dis	Tmin	Ttip	Toc	Tco	Hlle	Tpa	F	Pt	Hsal
MADRID-ATOCHA										08023
AGUJA KM. 1.4	1.4	0049	0223		023	08003	002			08050
MADRID-SANTA C	3.9	0216	0219		023					08073
VILLAVERDE BAJO	1.8	0102	0106		013					08090
SAN CRISTOBAL D	1.5	0039	0049		010					08100
SAN CRITOBAL I	1.4	0036	0040		010					08110
EL CASAR (APD)	2.1	0048	0053		010					08120
GETAFE-INDUSTR	1.5	0033	0037		010					08130
PINTO	7.0	0237	0238	020	050					08180
VALDEMORO	6.3	0242	0247		030					08210
CIENPOZUELOS	6.5	0248	0253		030					08240
SESÉÑA (APD-CGD)	7.3	0244	0245		030					08270
ARANJUEZ	7.4	0247	0334		040	08310	001			08320

Fig. 4. Real timetable from a Spanish train

Where:

- T_{min} (in mmss) is the minimum travel time for each track section (it is obtained from physical data and dynamic models of tracks and trains).
- T_{tip} (in mmss) is the basic travel time. T_{tip} introduces a user-defined slack (it is obtained from manager's experience) for robustness purposes. Infrastructure managers call it 'security margin'.
- T_{oc} (mss) is an extra-time for current maintenance operations in tracks.
- T_{co} (in mms) is the final rounded travel time for the train in the track (in this example, operational constraints require travel times rounded to 30").

The added slacks in T_{tip} give some robustness $R(x)$ to the timetable. However, these slacks increase the global travel time $F(x)$. Thus, there exist a direct relation between the function $F(x)$ to optimize and the timetable robustness $R(x)$ [3]: robustness is increased $R(x)-R(x^*)$ at expenses of a loss of optimality $F(x^*)-F(x)$.

Currently, slacks are introduced according to historical experience of infrastructure manager. Therefore, some questions are:

- Where should the slacks be introduced in order to maximize $R(x)-R(x^*)$ versus $F(x^*)-F(x)$?
- Should they be uniformly introduced?
- Which relation exists between $R(x)$ and $F(x)$ in each timetable?, etc.

4.2 To Decrease Capacity

Robustness is obtained decreasing the theoretical capacity of the infrastructure. It is especially useful in the case of double tracks. Decreasing capacity is also other usual method applied by railway infrastructure managers, as the UIC methods proposed by the International Union of Railways [UIC96, UIC04] do.

Railway managers know that scheduling as many trains as theoretical capacity says is not viable. However, they have to make the best use of the expensive railway infrastructures. Trade-off between capacity and reliability/robustness, in other words, between the 'physical maximum' level of capacity and the 'economically optimal' level of capacity is a key point in operational management contexts [1].

For instance, Fig. 5 is a well-known graphic for railway managers. Clearly, as greater capacity is used, greater will be

the risk of secondary delays due to incidences. Thus, the idea is not to use the infrastructure at the maximum level of capacity (theoretical capacity), but with a practical capacity limit. Thus, Robustness is increased $R(x)-R(x^*)$ at expenses of a loss of capacity $C(x^*)-C(x)$.

However, there are no clear standards about how this trade-off can be modeled. For this purpose, application of simulation methods seems to be an appropriate way to evaluate the trade-off between capacity and reliability. A stochastic model relating delays and capacity can be found in [12].

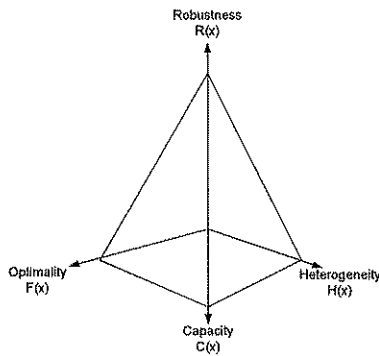


Fig. 5. Main parameters related to robustness

4.3 To Decrease Heterogeneity

Railway traffic is considered to be homogeneous if all trains have similar characteristics, especially the same average speed per track segment, resulting from the running times and the stopping times. However, for large railway networks, railway traffic cannot be fully homogeneous due to freight and passenger trains share the same infrastructure. If there are large differences in the timetable characteristics of the trains on the same track, then the railway traffic is called heterogeneous [22]. Due to heterogeneous trains share the same infrastructure over large distances, timetabling becomes very complicated. Heterogeneity usually leads to many small headway times, which may increase delay propagation in the operations so that robustness decrease. It is generally accepted in practice that the heterogeneity resulting from the line plan and the timetable has a negative influence on the punctuality and the reliability of a railway system (UIC, 2004). Thus, homogeneous trains increase robustness and heterogeneous trains decrease robustness.

4.4 Trade-off Between Optimality/Capacity and Robustness

Thus, according to the previous points, there is a clear trade-off between optimality/capacity and robustness, as the known price of robustness [3]. All these parameters: robustness, capacity, optimality and also heterogeneity are directly related (see Fig. 5). As we have pointed out, the robustness can be increased by decreasing capacity, optimality and heterogeneity.

Furthermore, different timetables with near-to-optimal travel times can be obtained due to there exist many traffic operations (commercial stops, overtaken, crossings, replacements, etc.) which are responsible for generating different timetables. Thus, we should choose the most robust timetable among alter-

native timetables with similar and near-to-optimal travel times. Here, we need some analytical or empirical functions to characterize robustness of each timetable. The goal is to obtain a better configuration of timetables, without penalizing travel time of trains.

Fig. 7 shows an example of two timetables with the same amount of additional buffer time. However, the left timetables seems to be more robust that the right one. Thus, we can generate several timetables and select the more appropriate one depending on operator requirements.

5. How to Measure Robustness?

As was pointed out, a clear trade-off among capacity, optimality and robustness exists. We can measure optimality of a function. For instance, in our case:

$$F(x) = \sum \text{travel times for all trains in the timetable.}$$

Likewise, we can obtain theoretical capacity of a rail network $C(x)$, as well as the used capacity of a given solution. However,

- How the robustness of a solution $R(x)$ can be measured?
- How can we quantify robustness of timetables?
- How can we assure that a timetable is more robust that other?

We can point out several ways:

5.1 Stochastic Programming Model

Stochastic programming is an approach for modeling optimization problems that involve uncertainty. Whereas deterministic optimization problems are formulated with known parameters, real world problems almost invariably include parameters which are unknown at the time a decision should be made. Stochastic programming models are similar in style but try to take advantage of the fact that probability distributions governing the data are known or can be estimated.

In railway timetabling, it is assumed that the process times, in the timetable generation phase, are deterministic. However, the real-time process times in the operations are stochastic. Difficulties of stochastic programming method, in railway timetabling, are pointed out in [12]. Although in simpler contexts it can be applied, we consider very difficult obtaining a model and applying stochastic programming in periodic and no-periodic timetables with very complex capacity and traffic constraints.

5.2 Analytical Methods

They consist of measuring 'certain characteristics' of timetables to evaluate its robustness.

There are several analytical methods to evaluate robustness in general scheduling problems. These methods require metrics that characterize the robustness of solutions. In [2] the flexseq metric is proposed, that consists on counting the number of activity pairs that are not directly related; In [6] a fldt metric that measures the roominess associated to each activity is defined; Nicola et al. [18] give a metric to measure the impact caused by an incidence; etc.

Although these analytical methods could provide a quicker way to measure robustness, they are not valid in railway con-

text and they do not correctly describe the robustness of a railway timetable.

In a more specific railway context, analytical measures such as SSHR and SAHR has been developed to evaluate the homogeneity of a timetable assuming the relation between this factor and the propagation of delays due to interdependencies between trains [22]. Carey [4] proposes to use dispersion measures of the inequality of headway times. These measures of robustness are especially useful in the case of double track lines. However, additional measures have to be developed in case of overtaking or in the case of single track lines. They must be based on the existing slack time with respect to the minimum reception time between two crossing trains.

5.3 Empirical Methods: Simulation

This method is related to "what-if" analysis. The method consists of simulating incidences in a timetable, next re-planning and to evaluate the effect on the final timetable (as the difference with the initial one).

Thus, the robustness of a solution $R(x)$ can be assessing by:

- I. the average delay of trains due to a set of incidences.
- II. the required interval time until the initial plan is restored.

This empirical method requires:

- I. A real-world problem simulator to generate incidences, to compute the knock-on delays and re-planning the timetable.
- II. Initial sets of timetables and data sets (benchmark) of realistic incidences. At this point, collaboration with an on-line infrastructure manager remains very important.

From these results, we could estimate (or validate) which are the main parameters that affects timetabling robustness (train velocity, heterogeneity, used capacity, configuration of signalling controls, etc.) and to asses how timetabling robustness is affected upon variation of these parameters. Perhaps then, it would be possible to obtain analytical conclusions and to draw a model of Robustness versus Optimality/Capacity.

6. Analytical Approaches for Robust Timetabling

In this section, we present two analytical approaches to measure robustness in railway timetabling. To this end, we give first two new definitions of robustness related to time. These definitions are based on the upper bound of the delay of trains in a timetable. Thus, these definitions are a function of time. The user selects an upper bound, and the function returns a percentage of robustness.

Definition 1: We define t -robustness of timetable x ($R_{robust}(x,t)$) as the percentage of disruptions lower than t time units that the timetable is able to tolerate without any modifications in traffic operations (crossing, overtaken, etc).

Definition 2: Following the definition given by Nicola [16], a timetable is (t,k) -robust if upon an disruption lower than t time units, it is able to return to the initial stage after k time units.

For instance, given a timetable x , if user selects $t = 120$ seconds, by using definition 1, $R(x, 120) = 40\%$ means that 40% of disruptions can be absorbed without any modifications

in traffic operations such as crossing, overtaken, etc. However if $t = 180$, then the percentage of robustness decreases $R(x, 180) < R(x, 120)$ due to the fact that the probability of absorbing higher disruptions is lower. Fig. 9 shows the behavior of two different timetables.

6.1 Parameters to Measure Robustness

In this section we will identify the main parameters that we will take in consideration to measure robustness in railway timetabling:

- **buffer time** is the temporal interval in which a train is stopped and it is not carrying out any traffic operation. This slack is assigned by railway operators in certain stations to absorb disruptions that have been carried out in these stations or previous stations. As the number of buffer times increased the optimality decreased, but robustness must increase.
- **Time supplement** is the additional time included in the travel time of a train on each track. Without loss of generality, we assume that time supplements are included in the travel time of a train on each track.
- **Number of trains.** In timetables composed by a high number of trains is more probably that a disruption occurs. Indeed, capacity and robustness are opposed terms, so that less capacity (less number of trains) implies higher robustness.
- **Number of commercial stops.** The number of commercial stops is directly related with the number of disruptions. As the number of commercial stops increase, the number of possible disruptions also increase.
- **Flow of passengers.** Some stations are more important than others. Stations in large cities are used by a larger number of passengers than stations in small villages. There is more probability that disruptions occur in high loaded stations. Furthermore this parameter also depends on trains. Railway operator can give us a disruption probability for each train in each station. Thus, it is very appropriate assigning buffer times in high loaded stations.
- **tuple \langle train, station, hour \rangle .** This parameter is very close to the former one. There exist rush hours where stations are collapsed and disruptions appear. Thus it is very appropriate to assign buffer times in these locations in these rush hours.
- **Tightest tracks.** The tightest track is considered as the longer distance track. Thus, if there exist a large single track between two consecutive stations s_1 and s_2 , and a train is travel from s_1 to s_2 , all trains in opposite direction must wait in station s_2 for a long time. Thus, it is convenient to maintain buffer times in both stations that compose the tightest track in order to absorb disruptions.
- **Some other parameters.** There exist some other parameters that must be taken into account. These parameters are: age of trains, infrastructure maintenance, etc.

6.2 Two Analytical Proposals

As we pointed out before, as the number of buffer times increase the timetabling robustness might also increase. Thus, a

very preliminary approach of measure of robustness would be sum all buffer times of a timetable x

$$R(x) = \sum_{i=1}^{T \times S} buff_i$$

Therefore, given two timetables x and y , if $R(x) > R(y) \Rightarrow$ timetable x might be more robust than timetable y . However, this tentative approach is very weak and it is not a good measure of robustness due to it does not take into consideration all above parameters. Fig. 7 shows two timetables with the same amount of buffer time. However, it is reasonable to say that the left timetable is more robust than the right timetable. So we must take into account the above parameter to obtain a more specific analytical method to measure robustness.

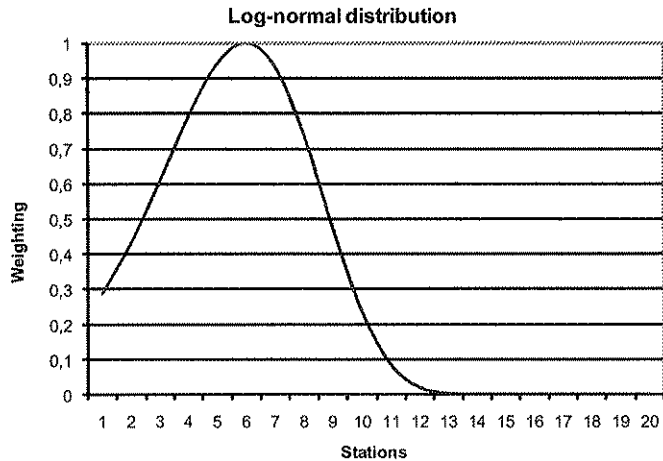


Fig. 6. $F(S)$: weighting in stations

Proposal 1. The first analytical method is given in the following formula:

$$R(x) = \sum_{T=1}^{NT} \sum_{S=1}^{NS} Buff_{TS} \times Flow_{ST} \times TT_S \times NSuc_{T_T} \times f(S);$$

where: $R(x)$: Robustness of timetable x .
 T : train.
 NT : number of trains
 S : station
 NS : number of stations;
 $Buff_{TS}$: buffer time of train T at station S .
 $Flow_{ST}$: percentage of passenger flow in train T and station S . This weight is normalized to range between 0 and 1.
 TT_S : probability of tightness of track between station S , $S + 1$. Thus, given a timetable with k tracks, the tightest track has $TT_1 = k/k = 1$, the second tightest track has $TT_2 = k - 1/k$, and so on. For instance, given a timetable with 21 stations, there exist 20 tracks. The tightest track is labeled to 1; the second tightest track is labeled to 0.95, the third is labeled to 0.9, and so on. Finally, the loosed track is labeled to 0.05.

$NSuc_{T_T}$: probability of trains that may be disrupted by train T . To this end, we will estimate a percentage of allowed delay.

$f(S)$ is a formula similar to a log-normal distribution. This function gives preference to stations where crossing are more probable to occur. This function depends of S , NS and μ that represents the average number of stations.

$$f(S) = e^{-\frac{(S+\mu/2)-\mu)^2}{NS-\mu}}$$

Fig. 6 presents the weighting of each station in a running map. It can be observed that the curve represents a log-normal distribution. The higher weightings are given to stations in the center of the running map. This is due to the fact that these stations are more probable that crossings and overtaking occur. Thus, buffer times at the beginning and center stations are more appropriate to absorb primary disruptions.

All these weightings constitute a new measure of robustness, due to each buffer time is pondered by these weightings. Furthermore, each parameter is ranged between 0 and 1 in order to give the corresponding importance to each parameter.

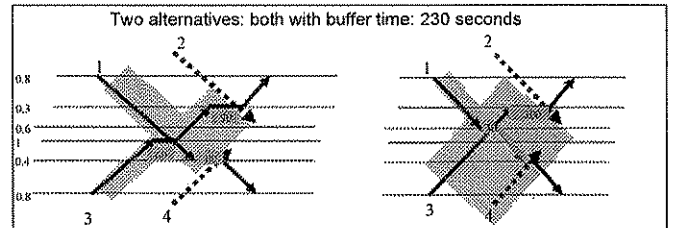


Fig. 7. Two timetables with different configurations but same amount of buffer time

Proposal 2. The second analytical method is related to the definition 1 presented above. This method gives us a time dependent indicative measure of how good the timetable behavior is when unexpected disruptions occur. As Norio presents in [21], there exist some several robustness indexes for each level (see Fig. 2). Here, we consider different percentages of robustness for different sizes of disruptions. A disruption of 3 seconds in a train can be absorbed by practically all timetables. However disruptions of 50 minutes in a train could affect a significant proportion of the fleet. Finally, disruptions of 200 minutes in several trains could generate cancelations or large disturbances in the timetable.

```

Function Nabsorbeddisruptions(x, t)
Input: timetable x, matrix of crossing/overtaken, matrix
of buffer times, time t
Output: Number of absorbed disruptions
forall train-i do
  forall station-j do
    k=j;
    while(not crossing/overtaken[train-i, station-k])
      total=total+buffer[train-i, station-k];
      k++;
    remaining[train-i, station-j]=total - t;
    if remaining[train-i, station-j]>0;
      absorbed++;
  end
end
return absorbed;
    
```

Algorithm 1: Number of absorbed disruptions

Thus, a second analytical method is given in the following formula:

$$R(x, t) = 100 * \frac{Nabsorbeddisruptions(x, t)}{T * S}$$

where, $Nabsorbeddisruptions(x, t)$ studies all crossings and overtaken and returns the number of disruptions that can be

absorbed with the available buffer times. The pseudo-code of this algorithm is presented in Algorithm 1. Given a timetabling, a matrix of crossing/overtaken, a matrix of buffer times and a time point, the algorithm calculates the number of absorbed disruptions by traversing each station for each train.

7. Evaluation

In this section we evaluate our robustness approaches on real timetables from the Spanish railway infrastructure. To this end, we carry out the following steps on each railway timetable.

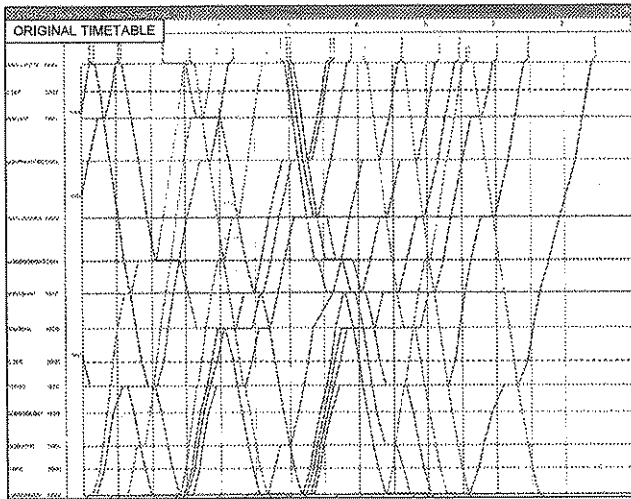


Fig. 8. Original Timetable

- I. Select a random timetable with single line, a significant number of stations and heterogeneous trains in order to study the robustness of this timetable. This timetable will be defined as *Original Timetable*.
- II. Add buffer times randomly on the timetable. Thus, we select a significant amount of global buffer time to distribute over the timetable in different buffer times. This timetable will be called *Timetable A*.
- III. Repeat the above step to generate another timetable with the same amount of global buffer time but distributed in other stations. This timetable will be called *Timetable B*.
- IV. Execute our robustness simulator to study the behavior of both *Timetable A* and *Timetable B*. The simulator assigns randomly a disruption of size d every t minutes and it returns important information such as the time needed to recover.
- V. Apply the analytical above formulae to both *Timetable A* and *Timetable B*.
- VI. Compare obtained results over both *Timetable A* and *Timetable B*.
- VII. Conclude which timetable is more robust.

Let's see these steps over a real Spanish timetable (see Fig. 8). This timetable is composed by 10 stations and 4 halts. There are 11 trains in down direction and 16 trains in up direction. We have inserted several buffer times in regions of this timetable. Thus, from the *Original Timetable*, we have generated two different timetables: *Timetable A* (see Fig. 10) and *Timetable B* (see Fig. 11). *Timetable A* was generated by inserting, from the *Original Timetable*, buffer times on the beginning stations.

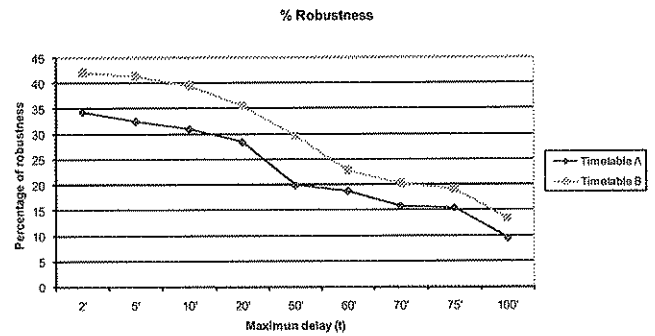


Fig. 9. Percentage of robustness of *Timetable A* and *Timetable B* using formula (2).

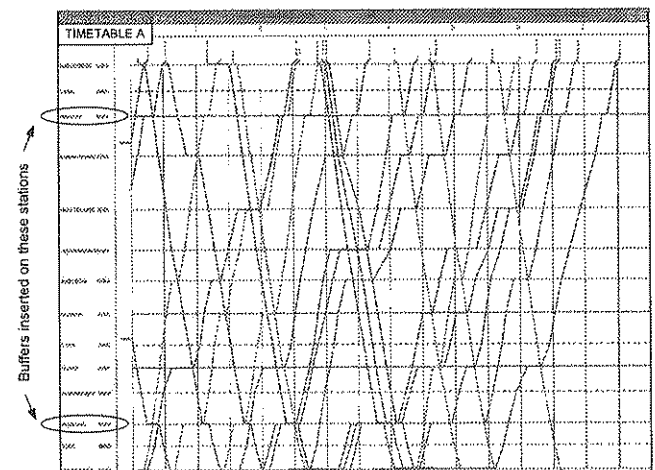


Fig. 10. *Timetable A* with random buffer times on the beginning stations

Timetable B was generated by inserting, from the *Original Timetable*, buffer times on the middle stations.

Let's see these steps over a real Spanish timetable (see Fig. 8). This timetable is composed by 10 stations and 4 halts. There are 11 trains in down direction and 16 trains in up direction. We have inserted several buffer times in regions of this timetable. Thus, from the *Original Timetable*, we have generated two different timetables: *Timetable A* and *Timetable B*. *Timetable A* was generated by inserting, from the *Original Timetable*, buffer times on the beginning stations. *Timetable B* was generated by inserting, from the *Original Timetable*, buffer times on the middle stations.

Then, *Timetable A* and *Timetable B* were evaluated by using formulae (1) and (2). Formula (1) gave the following results.

- $R(A) = 25.3$
- $R(B) = 32.7$

So, we can conclude by using Formula (1) that *Timetable B* is more robust than *Timetable A*. However this formula does not return a measure of robustness for a timetable. It only compares two timetables and returns which is more robust.

Formula (2) is a function over time t to measure the percentage of robustness. We have evaluated this function by increasing t from 2 minutes to 100 minutes. Fig. 9 shows the behavior of both timetables. It can be observed that as the maximum delay t increased, the percentage of robustness decreased in both timetables. However, *Timetable B* was more

robust than *Timetable A* in all instances. Thus, by using this formula, we can also conclude that *Timetable B* is more robust than *Timetable A*.

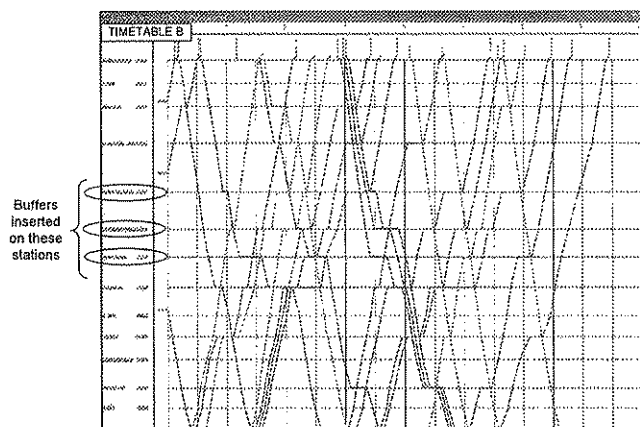


Fig. 11. *Timetable B* with random buffer times dispersed on the middle stations

To validate these results, we have simulated the behavior of both timetables in order to study the global delay of both timetables. To this end, the simulator randomly inserted a disruption every one hour over the timetable. The size of disruptions was increased from 1 minute to 100 minutes. Fig. 12 shows the total delay on *Timetable A* and *Timetable B*. For short disruptions, both timetables had similar behaviors, but for higher disruptions, the total delay of *Timetable A* was higher. Thus, we can assume that *Timetable B* is more robust than *Timetable A*.

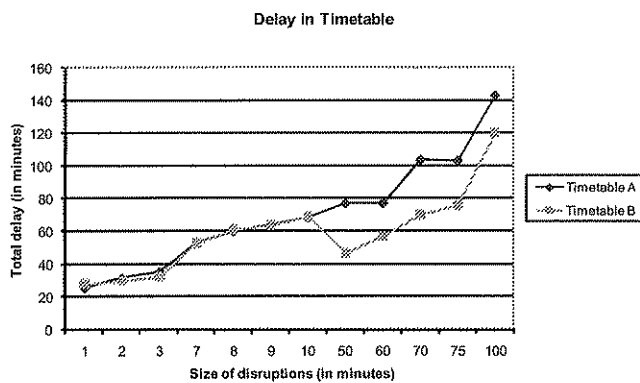


Fig. 12. Delay in *Timetable A* and *Timetable B* caused by disruptions using simulation.

These formulae have been applied over several Spanish timetables with different topological characteristics (number of trains, number of stations, train heterogeneity, etc.). The obtained results have been compared with the simulation tool to validate the above formulae.

8. Conclusions and Further Work

Robustness of a railway timetable is an indicative measure of how good a timetable is due to robust timetables allows the railway operators to cope with unexpected disruptions which normally occur on a daily basis. A timetable can have the characteristic that delayed trains will lead to considerable knock-on

effects, whereas another configuration of the timetable may be able to absorb such effects more readily. In this paper we have presented the main parameters that can be directly related to robustness in single-line tracks. Railway operators (based on their expertise) help us to determine the most important parameters and they have been included in our analytical formulae to compare and measure robustness. We have checked these approaches with a simulation tool to validate the obtained results. We will study some other features that can be involved to measure robustness. To this end, we must analyze the historic data about disruptions (given by railway operators) to obtain useful statistics to improve our formulae.

Acknowledgements

This work has been partially supported by the research projects TIN2007-29666-E and TIN2007-67943-C02-01 (Min. de Educacion y Ciencia, Spain-FEDER), GV/2007/274 (Generalidad Valenciana) and by the Future and Emerging Technologies Unit of EC (IST priority - 6th FP), under contract no. FP6-021235-2 (project ARRIVAL).

We thank the assistance of Jose Estrada of the Spanish Railway Infrastructure Manager (ADIF) for his expertise in robust timetabling. He helped us to identify the main parameters to measure robustness and give us many real timetables where many disruptions occur.

References

- [1] M Abril, F Barber, L Ingolotti, M Salido, P Tormos and A Lova, An assessment of railway capacity. *Transportation Research - Part E*, doi:10.1016/j.tre.2007.04.001, Vol. 44, 2008.
- [2] M Aloulou and M Portmann, An efficient proactive reactive scheduling approach to hedge against shop floor disturbances. In *Proc. of the 1st Int. Conf. on Scheduling, MISTA 2003*, 2003, pp. 337-362.
- [3] D Bertsimas and M Sim, The price of robustness. *Operations Research*, Vol. 52, 2004, pp. 35-53.
- [4] M Carey and S Carville, Exact heuristic measures of schedule reliability. *Journal of Operational Research Society*, Vol. 51, 2000, pp. 666-682.
- [5] M Carey and S Carville, Estimating schedule performance and reliability for train stations. *International Transactions in Operational Research*, Vol. 11, 2004, pp. 382-394.
- [6] A Cesta, A Oddi and S Smith, Profile based algorithms to solve multiple capacitated metric scheduling problems. In *Proc of the 4th Int. Conf. on Artificial Intelligence Planning Systems AIPS-98*, 1998.
- [7] A Davenport, C Gefflot and C Beck, Slack-based techniques for robust schedules. *Proceedings of 6th European Conference on Planning*, 2001.
- [8] M Drummond, J Bresina and K Swanson, Justin-case scheduling. *Proceedings of the 12th National Conference on Artificial Intelligence*, 1994, pp. 1098-1104.
- [9] M Ginsberg, A Parkes and A Roy, Supermodels and robustness. *Proceedings of AAAI-98/IAAI-98*, 1998, pp. 334-339.
- [10] E Hebrard, B Nich and T Walsh, Robust solutions for constraint satisfaction and optimization. *Proc. of the 16th European Conf. on AI (ECAI 2004)*, 2004, pp. 186-190.
- [11] J Hoogheemstra and M Teunisse, The use of simulation in the planning of the dutch railway services. In *Proceedings of the 1998 Winter Simulation Conference*, 1998.
- [12] T Huisman and R Boucherie, Running times on railway sections with heterogeneous train traffic. *Transportation Research Part B*, Vol. 35, 2001, pp. 271-292.

- [13] W S Leon V. and R Storer, Robustness measures and robust scheduling for job shops. *IIE Transactions*, Vol. 26, 2004, pp. 667–682.
- [14] L Mattson, Train service reliability: A survey of methods for deriving relationship for train delays. Research Report, Unit for Transport and Location Analysis, Royal Institute of Technology, Stockholm, Sweden, 2004.
- [15] D Middelkoop and M Bouwman, Large scale train network simulations. In *Proceedings of the 2001 Winter Simulation Conference*, 2001, pp. 1042–1047.
- [16] N Policella, Scheduling with uncertainty: A proactive approach using partial order schedules. PhD Thesis, Università degli Studi di Roma: La Sapienza, 2005.
- [17] N Policella, A Oddi, S Smith and A Cesta, Generating robust partial order schedules. *Lecture Notes on Computer Science (LNCS) Vol. 3258*, 2004, pp. 496–511.
- [18] N Policella, S Smith, A Cesta, and A Oddi, Steps toward computing flexible schedules. *Proc. of Online'03 Constraint Solving: Handling Change and Uncertainty*, 2003.
- [19] N Policella, S Smith, A Cesta and A Oddi, Generating robust schedules through temporal flexibility. *Proceedings 14th International Conference on Automated Planning and Scheduling*, 2004.
- [20] M Sevaux and K Srensen, A genetic algorithm for robust schedules in a just-in-time environment. Technical Report LAMH/SP-2003-1, 2002.
- [21] N Tomii, Robustness indices for train rescheduling. *1st International Seminar on Railway Operations Modelling and Analysis*, 2005.
- [22] J Vroman, R Dekker and L Kroon, Reliability and heterogeneity of railway services. *European Journal of Operational Research*, Vol. 172, 2006, pp. 647–665.

constrained optimization and distributed methods for railway scheduling.

Author Bios

- Miguel A. Salido is an associate professor in Computer Science at the Technical University of Valencia, Spain. Most of his research is focused on techniques for constraint satisfaction techniques and railway scheduling problems. He is the recipient of some national and international awards. He is author of more than 70 papers published in international journals and conferences. He is the editor of several books and guest editor of some international journals. He is a member of several Scientific and Organizing Committees.
- Federico Barber is Full Professor with the Department of Computer Science, where he leads an artificial intelligence research team. He has worked on the development of temporal reasoning systems, Constraint Satisfaction Problems, planning and scheduling. He is the author of about 90 research articles which have been published in several journals and conferences. His research has produced several tools for solving real-world optimization combinatorial problems. He has participated in and led several national and European research projects related to these areas. He is currently president of the Spanish Association for Artificial Intelligence, and member of several scientific associations.
- Laura Ingolotti is an associate professor in Computer Science at the Technical University of Valencia, Spain. She has an industrial background with the National Network of Spanish Railways. Her work has focused on the development and study of robustness and railway capacity. Her current research interests include metaheuristic methods,

