

Improvement of a railway planning tool using Genetic Algorithms

J. de D. Sanz Bobi², J. Gómez Ramírez², J. Garzón Núñez¹
& R. Galán López²

¹*CITEF Research Centre on Railway Technologies,
Universidad Politécnica de Madrid, Spain*

²*Universidad Politécnica de Madrid, Spain*

Abstract

In this paper a new approach to a traffic planning process is presented. A new railway line planning tool is being developed inside a more complex system that encloses simulation and train control tools as well. Genetic Algorithms (GAs) are used in this tool in order to lessen the time perturbations obtained by a previous algorithm when a number of train units are retired from the line to increase the time interval between circulations. The way these units are removed from the line is mainly affected by two factors: the number of available retiring points in the line and how fast this interval between train units is increased.

In the case that the interval between trains is reduced, no major problem is presented as adding train units to the line quickly decreases the interval between circulations. On the contrary, during the interval when trains can be retired as the increasing of time passing between trains in the stations, significant perturbation in the schedule of trains are produced and as is shown in the paper, GAs are able to handle the problem in a more efficient way than linear programming models or heuristic algorithms.

Keywords: railway planning, decision making, evolutionary techniques, Genetic Algorithms.

1 Introduction

A traditional railway service is based on regular timetables which are modelled as satisfaction constraint problems. Algorithms, implemented as constraint-based local search, aims to solve problems of realistic size can be found in [3–5].



Anyhow, train timetabling is a hard problem that needs of different techniques rather than the recurrent analytical and heuristic approaches, in order to deal with critical situations.

This paper presents the planning capabilities of RPLM, a tool for Railway Planning and Line Management. RPLM integrates decision making processes and advanced simulation, including electrical, dynamical and operational aspects.

Intelligent planning and traffic regulation is carried out with the help of a simulation module that provides valuable information about how the planning and regulating tools react to the control orders that are managing the railway line [1]. The simulator gives information of both mechanical and electrical behaviour of the system so electrical consumption can be taken into consideration in order to obtain an operational plan

In this study we focus on decision making processes for railway scheduling. For a general vision of RPLM with details in simulation aspects, including further analysis of the regulation actions consult [2,3].

2 Railway schedule module

Before giving a general vision of the scheduling module of RPLM, we define the problem domain and its main characteristics.

The domain is a closed single-track section for metro network, characterized as follows:

- The route of each train is fixed
- Trains movement start and finish at the stations
- Trains passing in opposite direction can meet and pass
- Trains passing in opposite direction can meet and pass each other on the tracks or in the stations.

We outline next the basic characteristics of the planning module integrated in RPLM. The planning module has as inputs a Data Base with topological information of the line and the data introduced by the train operator

2.1 Data Base parameters

- *track time* $t_{A,B}$, the time necessary for a train to cover the distance between two consecutive stations. *Minimum dwell time* t_{\min}^* , is the strict minimum time in seconds that a train can dwell at a station
- *Stations* where the trains can be inserted or retired (i.e: $st_R = [st_1, st_2, st_9]$, $st_I = [st_1, st_5, st_8]$) then the trains can be retired from stations 1,2,9 and inserted from 1,5,8

2.2 User parameters

- Spatial distribution
 - *Number of available trains* that can be dispatched n_T



- *Position of the trains in the line*, that is, if the stations are busy or free of trains and how many trains are in the deposits in every significant time
- Time distribution (for each interval)
 - Requested scheduling frequency T , one for each interval (i.e a scheduling of three intervals is $T = [T_1, T_2, T_3]$)
 - Initial time departure (of the first train) and final instant of the interval (i.e: 06:00 - 2:00)

Besides a set of rule based restrictions or security principles are also taken into account in order to get a schedule solution that represents a spatio-temporal distribution of the trains in a closed single-track section for metro network line (Figure 1).

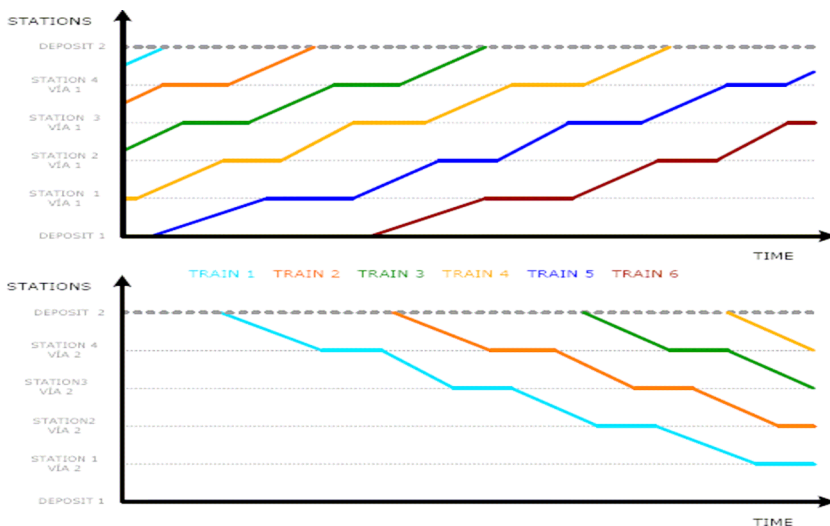


Figure 1: Distribution of the trains in a closed underground railway line.

2.3 Restrictions

- Only one train can be at any station at any time
- A train can depart the station A if both the track AB and the station B are free of trains

For the sake of space and simplicity we do not give details about the implementation of the planning module that provides the timetable for the dispatched trains. It is, however worth noting that the track time between two stations A, B T_{AB} , is not modified by the algorithm, hence it is a time constant and fixed a priori (In regulation phase $t_{A,B}$ would be different depending on the adopted ATO value.).

The scheduling, or running time that represents the timetable, that fits best with the user request, is achieved in two phases; *primo* computation of the

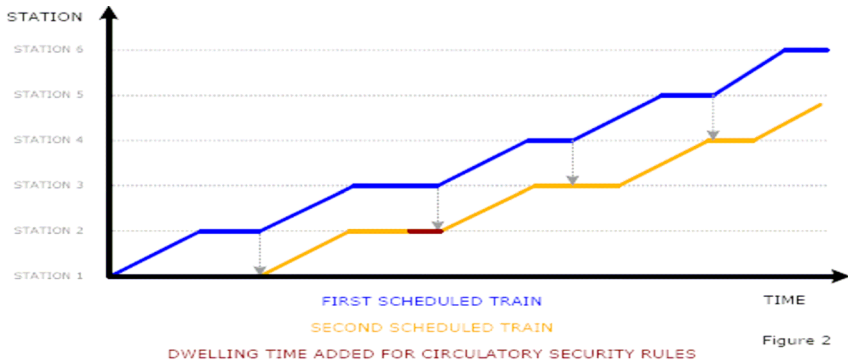


Figure 2: Two trains planned schedule.

number of necessary trains and *secondo*, the dwell time for each train every time it stops in a station. In Figure 2, the horizontal lines represent the train trajectory.

3 The problem of train retirement

RPLM has been constructed following three methodologies useful to obtain a scheduling for a whole day operation. First, linear programming obtains a train time schedule based on a set of restrictions. In the second step, heuristic algorithms are used to reduce dwell time at stations in critical points of the line. Finally, when train units are to be retired from circulation due to transitions between separate time periods with generated schedules, a GA is proposed in order to generate an optimal schedule. This paper's focus is to give an account of the last stage, that is, the application of GAs in scheduling. The interested reader in linear programming and heuristic methodologies can consult [5,6].

Of course, during the same day a train operator may fix several frequencies depending on the public flow: peak period, standard period etc. RPLM will predict the optimal number of trains in order to get the closest frequency to the required one for the different periods.

For example, during a scheduling plan, when an increment in frequency is required, a waiting time of 3 minutes turns into 6 minutes, some trains are not necessary any more and should be retired.

The number of trains to be retired, and how the traffic is affected by this perturbation in the original schedule, is the impetus for the planning module of RPLM which is presented in this paper.

In Figure 3 we can see the perturbation produced in the dwell time at station 4. On the left $T_1 = 3$ minutes and after the retirement of the train the dwelling time T_2 increases to the double (from 3 min. to 6 min). The perturbation that occurred as a consequence of the retirement of the trains that are not necessary is not severe. This is because T_2 is the double of T_1 . Graphically stated, some trains are retired (disappear from the figure) leaving a "hole", with similar length than the dwell time requested.

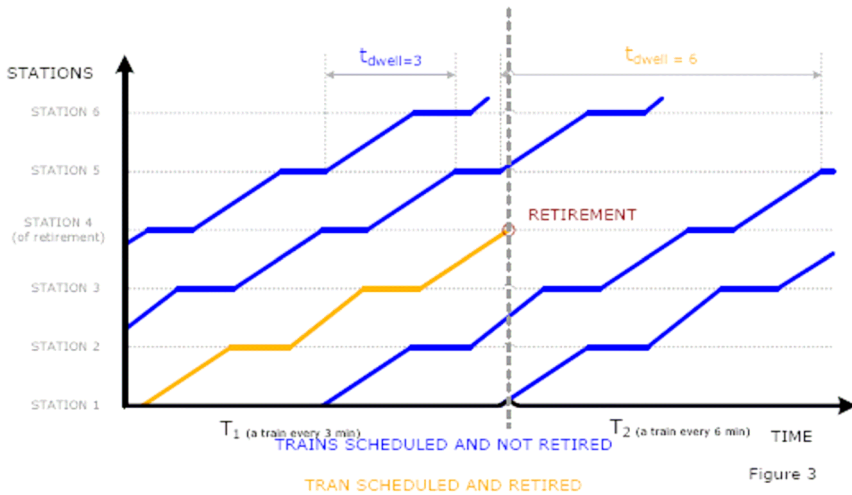


Figure 3: Change in train time interval.

Contrarily, as we have experimentally tested, in case of T_2 is close to T_1 or much bigger than $2 \cdot T_1$, the retirement of the trains produces significant perturbations in the service that can not be neglected.

In the next section we describe a suitable option to provide optimal scheduling based on GAs in order to solve the problem that arises when the number of trains retired is less than from half of the number of running trains.

It is good reminding that we are facing a space too vast to be optimised by a stochastic optimisation method.

Obviously, an analytic or heuristic solution for a problem of such complexity requires the construction of a model for a non linear system. An iterative method like GAs allows the usage of combinatorial optimisation methods that fits better with combinatorial explosion.

4 GAs for train scheduling in train retirement scenarios

In this section we detail the solution to the problem of retire trains when the dwelling time at the stations increases. The objective is to reduce the dwelling time at the stations, minimizing delays and reducing perturbations in the scheduling.

4.1 Temporal window

Let $R = \{r_1, r_2, \dots, r_N\}$ an ordered vector and r_1 the retirement time of the first train, r_2 the time for the second train and so on. The temporal window W_T is the time between the retirement of the first train and that of the last train plus minus a margin, $W_T = \{r_1 - M_0, r_2 + M_\infty\}$

Therefore W_T is the range of time where the trains are retired, and of course wherein the GA will try to reduce the perturbations in the schedule produced by the suppression of trains, specially at the beginning of the new frequency.

4.2 Restrictions

As is shown in Figure 3, the schedule can be seen as a two interval scenario, the first interval has 6 trains running from the very beginning to the retirement of the first train and the second interval, from the end of the first period until the end of the plan.

When a temporal window is included, the schedule passes to be a three intervals scenario. Of course W_T is in the middle and the two trains are retired into it. During Interval₁ and Interval₂ no train retirement is achieved. (Figure 4) Therefore the schedule is only modified into W_T , and Interval₁ and Interval₂ do not have to be affected by such modifications. We can express this formally through the next inequations that must be satisfied:

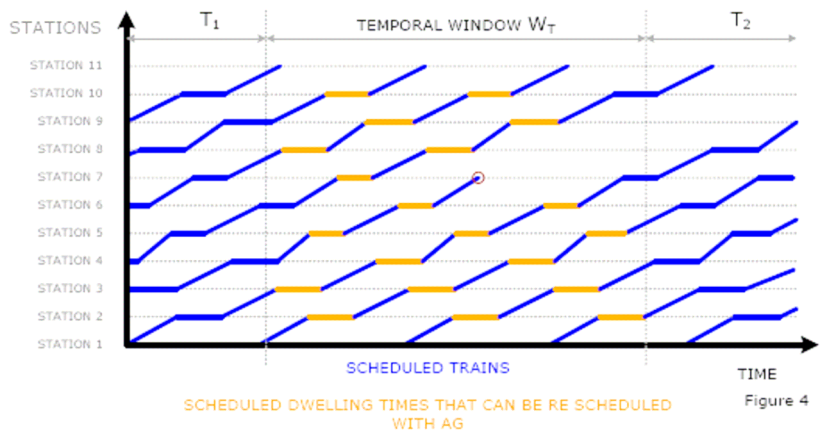


Figure 4: Rescheduled dwelling time.

Next we show a simple example that illustrates how to get the dwelling time for each station, Δt_{dwel}^i . For a line with 5 stations, $\Delta t_{max} = 10$ and $t_{dwel}^{i0}, t_{dwel}^{GAi}$ given in the table below (Table 1).

Table 1: Dwelling time for each station.

Station (i)	1	2	3	4	5
t_{dwel}^{i0}	35	30	33	41	36
t_{dwel}^{GAi}	30	25	25	35	30



$\Delta t_{dwell}^1 = [10, -5], \Delta t_{dwell}^2 = [10, -5], \Delta t_{dwell}^3 = [10, -8], \Delta t_{dwell}^4 = [10, -6], \Delta t_{dwell}^5 = [10, -6]$, and like $\Delta t_{dwell}^1 + \Delta t_{dwell}^2 + \Delta t_{dwell}^3 + \Delta t_{dwell}^4 + \Delta t_{dwell}^5 = 0$. Then, one possible solution for the new dwelling time into the temporal window is

$$\overset{->}{\Delta t}_{dwell} = \{3, 5, -6, -4, 2\}$$

4.3 The GA

The chromosome represents the increment of the dwelling time in station i Δt_{dwell}^i . Then, the initial planning (not affected by the temporal window nor GA

action) can be seen as a chromosome with $\overset{->}{\Delta t}_{dwell} = \overset{->}{0}$. A chromosome has as

many genes as stations wherein the train dwells. A gene represents an increment of the dwelling time. (Figure 5)

As the chromosomes are created randomly, is not sure that the train movement is permitted, that is, it could happen that a train departs or arrives from a station at an invalid moment. In order to avoid possible collisions we need to discern between legal and not legal chromosomes.

$$\forall t \in T, \forall i \in E, S_t^{i+1} \leq S_t^i$$

where T are the trains, E the stations, S_t^i the departure time of the train t at the station i and S_{t-1}^{i+1} departure time of the closest train to t , $t-1$, that departs from, $i+1$ which is the arrival station of t .

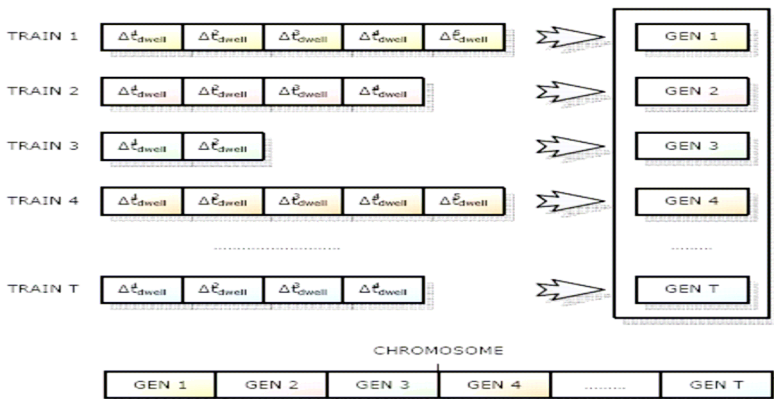


Figure 5

Figure 5: GA chromosomes and genes.

4.4 Crossover operator

As is empirically validated later, two point crossover is preferred to one point crossover for the scheduling problem with retirement of trains. Two parents (two train schedule solutions) are crossover by a point randomly chosen. Anyhow this can not be completely random because the information of the genes can not be divided. Then for a chromosome with G genes, $G-1$ possible crossover points exist.

The two children are either included into a new generation or discarded if they represent a no legal train movement.

4.5 Mutation operator

In the mutation process, the gene from the child scheduling is altered with a predetermined probability. The gene has a new randomly selected dwelling time. For our purpose, mutation helps to avoid that generations get stagnated, making generations have more variability. Indeed, characteristics that could have been abandoned in precedent generations are included in the new ones.

4.6 Fitness Evaluation

The Fitness Evaluation is based on the quadratic error between $t_{dwell}^{i,j}$, the dwell time of train j at station i , and the initially scheduled dwell time during the interval where the trains are retired (the interval with larger dwelling time).

$$F(t_{dwell}^{i,j}) = \begin{cases} 5(t_{dwell}^{i,j} - I_2)^2 & \text{if } t_{dwell}^{i,j} > I_2 \\ 5(I_2 - t_{dwell}^{i,j})^2 & \text{if } t_{dwell}^{i,j} < I_2 \end{cases}$$

5 Implementations of results

Next is shown the implementation of the GA for a scheduling with two different frequencies. The dwelling time increases from, the 3 min during the first interval, to the 5 minutes of the second (Table 2). Therefore some of the initial 18 trains will be retired.

Table 2: Time intervals.

Interval	Dwell Time
6 am- 10am	3 min
10 am- 1pm	5 min

Besides, we set up the next parameters:

The temporal window $M_0 = 600s$ and $M_{\infty} = 1200s$; and the probabilities of mutation are $P_M = 0.2$ and the probability of crossing is $P_C = 0.7$. Executing the GA we table with the evaluation function values shown next (Table 3).

Table 3: Evaluation function values.

Parameter	1	2	3
Population	50	100	100
Cross	1-point	1-point	2-point
Generation			
Initial	100,28	99,27	102,34
F(Generation 1)	97,77	95,84	95,78
F(Generation 2)	96,44	93,75	68,18
F(Generation 7)	91,61	86,33	21,64
F(Generation 15)	89,2	83,32	9,95

Now, we can calculate the new dwell time at the stations for the best chromosomes. Using the GA the dwell time is almost a 10 % better than the original solution based on linear and heuristic programming. The dwell time passes from 409s to 383s. (Figure 6).

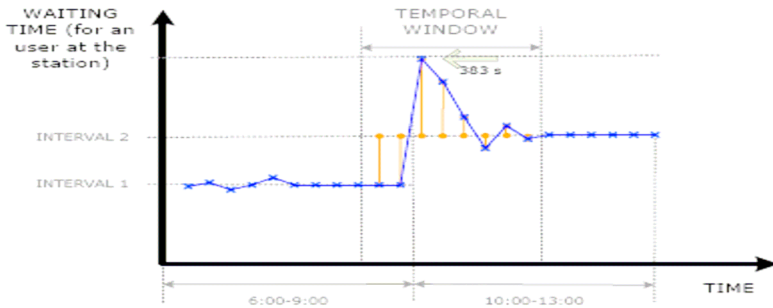


Figure 6: GA planning results.

6 Conclusions

A new method to improve railway planning has been shown. The usual timetable is affected by the capacity of the line (measured by the train-to-train time that depends heavily on the facilities, the distance between stations, the human resources and the taking-decision tools like central operation post) and social factors like number of expected passengers, strategic station and/or line.

The new method exposed in this work becomes an optimal search tool based in GA, combining the several heavily dependance variables to obtain a solution that lessens the transient time when the interval is increased. In this sense, the solution obtained provides an homogeneous time interval between trains as well as reduces the oscillations of the user waiting time, hence passengers can expect a minimum time to take the train including even when the number of trains in the line along the day is reduced.

GA gives a solution to the problem in a more efficient way than linear programming and heuristic algorithms used in this tool.



This method has a higher computing cost. This computing cost has to be reduced in order to use it to provide solutions in operation scenarios.

References

- [1] Sanz J. de D. Control de Procesos basado en eventos mediante Bases de Conocimiento Borrosas y Algoritmos Genéticos. *Ph. D. Thesis*. E.T.S. Ingenieros Industriales. Universidad Politécnica de Madrid. Spain. 2002.
- [2] Sanz J. de D., Galán R., Mera J. M., Garzón J., Reyes P. (2007). Advanced tool for railway planning and traffic control decision making. *Proceedings of IMECE2007. ASME International Mechanical Engineering Congress and Exposition*. November 11-15. Seattle, Washington, EE.UU. 2007.
- [3] Sanz J. de D., Jorreto F., Garzón J. A Simulation Tool for sizing electrical railway lines. *Proceedings of 2007 ASME/IEEE Joint Rail Conference & Internal Combustion Engine Spring Technical Conference*. March 13-1, Pueblo, Colorado, EE.UU. 2007
- [4] A. Lova, P. Tormos, M. Cervantes, F. Barber, An Efficient Adaptive Heuristic for the Resource Constrained Project Scheduling Problem. *XII Conferencia de la Asociación Española de IA*, 2007
- [5] L. Ingolotti, F. Barber; P. Tormos, A. Lova; ; M. A. Salido M. Abril. A scheduling order-based method to solve timetabling problems. *Lecture Notes in Computer Science, LNAI. Springer. vol: 4177, pp. 52–61. 2006*.
- [6] K. Fukumori, H. Sano, T. Hasegawa, T. Sakai, Fundamental algorithm for train scheduling based on artificial intelligence. *Systems and Computers in Japan*, 18(3): 52–63. 1987
- [7] A. Caprara, M. Monaci, P. Toth, P.L. Guida, A Lagrangian heuristic algorithm for a real-world train timetabling problem. *Discrete Appl. Math.* 154, 5 2006