IMPROVING THE MASS-TRANSIT FLOW BY TRAIN-HOLDING IN STATION WITH A SIMPLE AND EFFICIENT HEURISTIC

RÉMY CHEVRIER & FRANÇOIS RAMOND
SNCF – Innovation & Research, France

ABSTRACT

Railway traffic management in mass transit is often disturbed by perturbations on the running and the dwell times. Then, the headways between the trains may become insufficient and due to security distances between trains the perturbations propagate in the network and amplify. In this article, we propose a train-holding strategy in order to reset sufficient headways between trains. The mechanism is based on a heuristic which is simple and responsive. The prototype developed based on this concept was experimented in a subpart of the SNCF Transilien network: the RER A and L lines. During a 2-week experimentation, the train drivers were advised to depart after a dwell time that was accurately computed. The obtained results and the experimental feedback confirm the relevance and the efficiency of such a mechanism in the mass-transit context with important flows.
Keywords: driving, dwell time, traffic management, train-holding.

1 INTRODUCTION

Real-time railway traffic management in mass-transit systems becomes more and more crucial in large and dense urban areas all over the world. It is the case of the Paris area which is one of the most dense mass-transit areas in Europe. Every day, SNCF Transilien (TN) operates 6,000 trains and carries 3.2 millions passengers in a close area, and this number continues growing. With such amounts of trains and passengers, the traffic suffers from a lot of various perturbations which may disturb the operations. Their origin is multiple and some of them are due to human behaviour in peculiar. Indeed, the running times may be very variable because it is almost impossible for a train driver to respect accurately the running time defined in planning. Likewise, the dwell times in station depend largely on the passenger flows which can impose the train to stop longer than planned.

As a consequence, these factors of variability in running and dwell times may make the headway insufficient between two consecutive trains in such a way that the second train meets closed signals. In Paris area, signalling is paired to speed control systems for safety and the most largely deployed speed control system is KVB (French acronym for beacon-based speed control), which imposes a train meeting a warning signal to slow down to 30 km/h at 200 m before the signal whatever its aspect. Concretely, even if the next signal is open, the train has to exit the section entirely before being authorized to reaccelerate as illustrated in Fig. 1. In mass-transit systems, such security systems have a strong negative impact on the performance of operations because the time margins, which are short, are almost immediately consumed and the remaining capacity is then insufficient.

Therefore, this variability in running and dwell times as well as the strong interaction between the trains tends to generate perturbations which propagate rapidly in the whole network and increase the delays. It may become difficult to follow the nominal planning and in the worst cases, the infrastructure manager in charge of the dispatching and the traffic
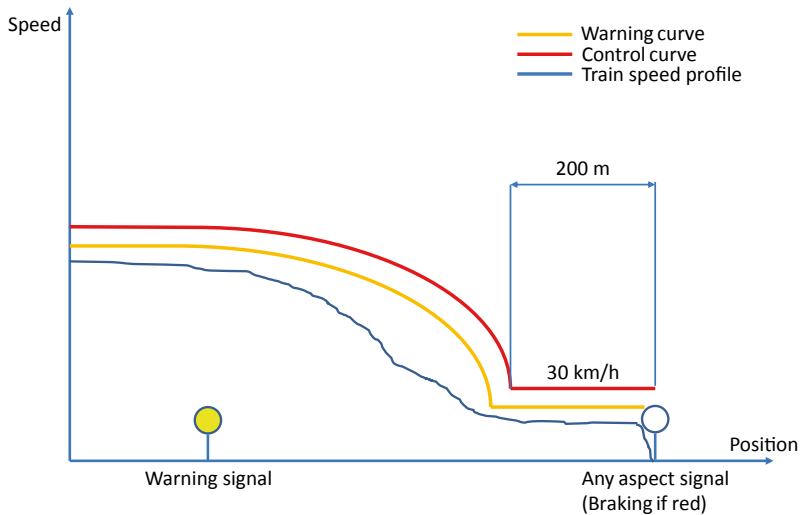
Figure 1: Speed control curves when crossing a warning signal in KVB system.

management might have to cancel trains in order to recover a nominal situation, that is, to make the traffic fluid again.

Our purpose is to deal with this objective of fluidity by introducing a mechanism to reset headways between trains in operations. For that, we propose a train-holding mechanism for the real-time mass-transit management. The principle is to hold trains in station whenever it is required in order to avoid restrictive signals and to display them the best moment to depart in order to make the traffic fluid. Even if other mechanisms exist to do that (e.g. by adapting the running speed), we explicitly choose to interact the least possible with the driver in inter-station, for safety reasons in particular. For that, we developed and tested a heuristic in a TN network subpart.

The article is organized as follows. In Section 2, we briefly review the literature related to the train-holding in station. The heuristic developed to accurately compute the ideal dwell time is detailed in Section 3. An experiment of the train-holding mechanism is presented in Section 4, followed by the analysis of the obtained results in Section 5. Section 6 concludes the article and mentions the next steps of this work.

2 BRIEF LITERATURE REVIEW

The principle of holding trains in station in order to let the traffic flow is well known in the literature. Indeed, it is commonly used to solve different kinds of operational problems.

For instance, one of those problems is about metro traffic management and consists of the balanced distribution of headways over the loop line to keep an appropriate frequency. In an early contribution [1], the authors use model predictive control (MPC) to foresee dwell times. The problem is formulated under quadratic form with linear constraints and the control variables represent the additional dwell times for each train for each next station. Later, this MPC formulation has been extended with speed controls in Ref. [2]. The authors use both speed and dwell time controls to spread headways over a metro line. More recently, another use of MPC appears in Ref. [3] which attempts to reduce passenger waiting times in a metro loop line, using passenger counting data.

The train scheduling problem which is known to be NP difficult can also be managed with the train-holding. For that, heuristics can be proposed to quickly solve this problem in real time, as in Ref. [4]. In relation to this, another type of application concerns the traffic management with disruptions no longer than 20 minutes which imply to hold multiple trains in stations in a network with junctions [5]. The proposed formulation is based on mixed integer non-linear programing and the solutions are obtained in around 10 seconds. As well, in Ref. [6], the authors deal with a similar problem and propose a quadratic formulation with linear constraints. The problem is solved with a heuristic which decomposes the overall problem into several sub-problems which can be solved to optimality.

In comparison to the application we deal with, the closest among those previously cited is the metro traffic management on account of the number of trains and the frequency implied. The volume of passengers is also very close to that carried in metro lines. In this kind of application, the major objective is to have a good headway spreading over the line and this is precisely one of the key issues on which we focus.

Besides, another key issue concerns the real-time aspect of the application which is crucial, given that the instructions must be given to the train driver as soon as the train arrives at the station. With regard to the literature, the real-time aspect often corresponds to some seconds of computation, to which we must in fact add the delay necessary to communicate traffic data through information systems and the delay to send the instructions. That is why we insist on the importance of the instantaneous computation in order to reduce the delay to present the information to the train driver.

3 A HEURISTIC FOR COMPUTING TRAIN DWELL TIMES

In planning, the timetables are built according to a set rules in such a way that the trains are scheduled with a sufficient headway to absorb little variations in running times. Hence, the planners add time margins to running times following the principles of the blocking time theory [7]. However, the stronger the density of trains, the shorter the time margins, and as soon as a perturbation occurs for a train, its time margin is consumed straightaway and there is no more to absorb next perturbations. Those ones are then amplified by closed signals and the driving rules imposed by the KVB. This one tends to increase the perturbations because it forces the trains to slow down and causes a capacity overconsumption.

The key idea at the basis of the heuristic proposed here is that it is preferable to waste a few seconds in station than wasting more in interstation. Thus, the heuristic aims to compute the accurate dwell time necessary to allow the train under control to run to the next station meeting only green signals. It is the principle known as the green wave as in Ref. [8].

Let T be the set of trains and $i, j \in T$ two trains such as the train j follows the train i. Let G be the set of stations and $k, l \in G$ two stations such that station k is located before station l in this way. When arriving at station k, the train j has to stop for letting passengers alighting and boarding. The minimal dwell time for the train j under control in station k depends directly on the run of train i, which precedes train j and stops in station l.

Let $t_{k \to l}^j$ be the running time of train j from station k to l. Let d_k^j, d_k^j be, respectively, the departure times of trains j and i from stations k and l. In order to avoid closed signals, a minimum time interval I_l^i has to be respected after the departure of train i from station l. The hard constraint mandatory for the traffic fluidity is therefore stated by the following inequality:

$$d_k^j + t_{k \to l}^j \geq d_l^i + I_l^i \tag{1}$$

That means the train under control must arrive after the departure of the previous train from the next station plus the time interval necessary to the previous train to clear the next sections and let the signalling to get back to green aspect.

Let a_k^j, w_k^j be, respectively, the arrival time of train j in station k and the minimum dwell time of train j in station k.

$$d_k^j = a_k^j + w_k^j \tag{2}$$

It is then straightforward to compute the minimum dwell time for the train under control by inserting eqn (2) in eqn (1):

$$w_k^j = d_l^i + I_l^i - a_k^j - t_{k \to l}^j \tag{3}$$

As one can easily imagine, this heuristic may imply uncertain data, in particular the running times and the departure times of the previous trains if they have not departed yet from (or even arrived at) the next stations. For the former, we have chosen to keep the reference running times, given that it is our aim that the train could run as planned. For the latter uncertainty, we use the statistical dwell times observed over a certain period of time to foresee the departure times. Based on this principle of time projection and with the dispatching decisions, it is then possible to compute dwell times all over the line, even if the accuracy decreases with the projection horizon.

4 EXPERIMENTS IN THE TRANSILIEN NETWORK

In this section, we describe the lines and the area on which the train-holding principle was experimented. We also explain how we chose the stations in which trains could be held during accurate dwell times. Finally, we explain how the information was presented to the train driver, from the arrival of the train in the station to its presentation under a countdown form through the computation on a dedicated server.

4.1 Description of the experimental area

The heuristic was experimented on a TN network subpart: the RER A and L lines, as depicted in Fig. 2. The RER A line is jointly operated by SNCF and RATP (the other major operator in Paris area). The SNCF part of the line has two departure stations: Cergy and Poissy towards Paris through the Nanterre-Préfecture station, from which the line is operated by RATP. The L line also starts from the Cergy station and shares its infrastructure with the RER A branch from Cergy to Maisons-Laffitte and with the two RER A branches from Maisons-Laffitte to the Nanterre divergence. From this divergence, the L trains run towards the Paris Saint-Lazare station (terminus).

In the morning peak hours, the nominal timetable is cyclic and periodic and plans 12 RER A (six departing from Cergy and six others from Poissy) and six L trains departing from Cergy (see Fig. 3 for a simplified view of it in a time-space diagram). One train from each line and branch departs every 10 minutes and all the trains gather in a common flow before Maisons-Laffitte station. It has to be noted that the RER A trains from Cergy are not omnibus and do not stop at Maisons-Laffite and Houilles-Carrières stations. The stops theoretically last 30 seconds, but often longer in practice to allow the passenger flows on platforms.
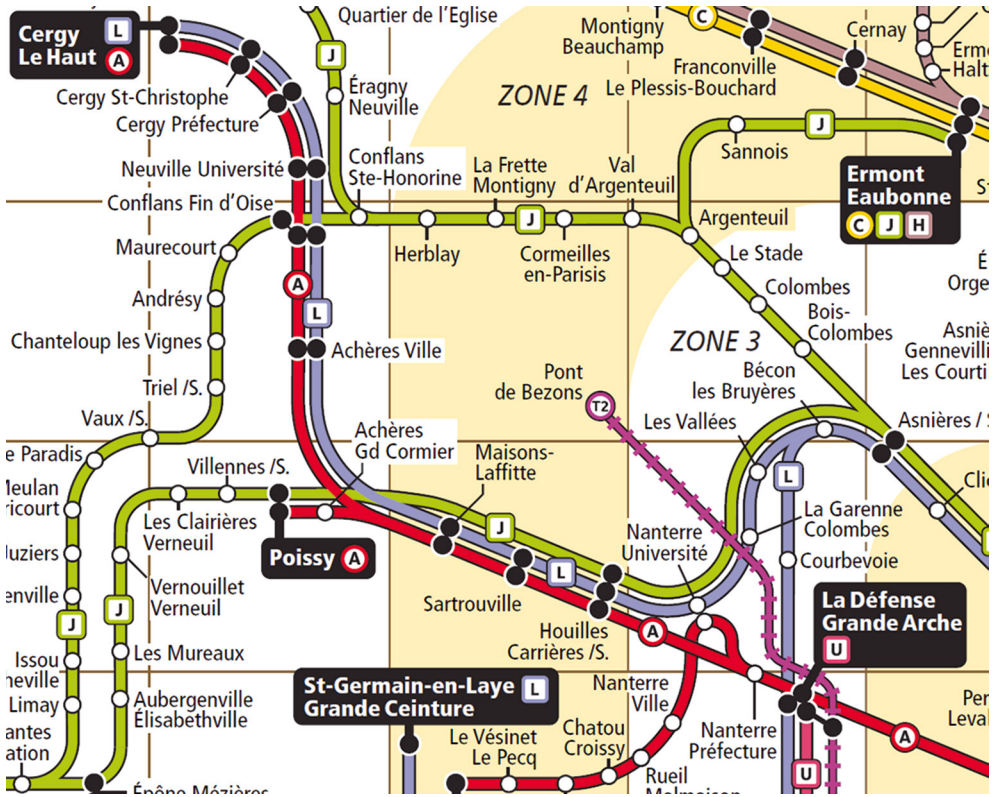
Figure 2: Schematic map of the RER A and L lines (source Transilien).
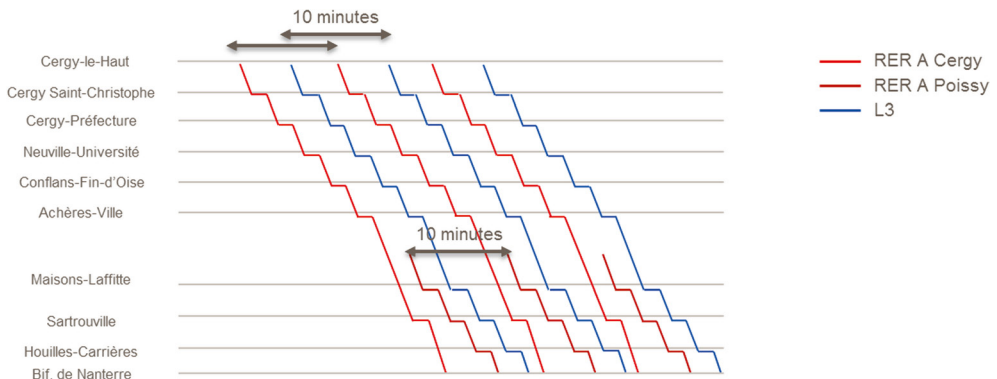

Figure 3: RER A and L lines nominal timetable (source Transilien).

4.2 Simulations for choosing the holding stations

The holding stations are pointed out in Fig. 4, which sketches the experimental area. Four stations have been identified to hold the trains and they are located around the junction: Achères-Ville and Achères Grand-Cormier before; Maisons-Laffitte and Sartrouville after.
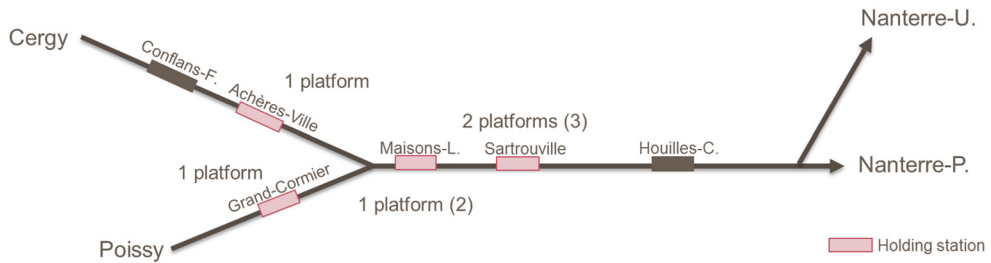
Figure 4: Experimental area with the four holding stations (up to seven platforms).

Even if choosing these stations seems quite obvious, a set of simulation runs was initially performed with the OpenTrack simulator to confirm. During the simulations, statistical perturbations were introduced to disturb running and dwell times. Disturbing running times consists in reducing the train performance (acceleration, reference speed and braking). In order to simulate the dwell times necessary to allow the passenger flow on platform, those are uniformly randomized from statistical distribution (mean and standard deviation). Any train arriving at any station of the RER A and L lines receives a departure time calculated by the heuristic explained in Section 3. Of course, the dwell time performed in simulation might exceed the one computed by the heuristic if the statistical perturbation was greater than it. Simulation results are given in Table 1. Four benches are compared while disturbing the train performance. In this table, an interval $[p\%, P\%]$ means the train performances are randomly chosen in the interval following a uniform law. As could be assumed, the results showed that all the retentions were spread over the four stations mentioned above. Besides the determination of the holding stations, the simulations also confirmed the relevance of holding trains in stations to make the traffic fluid again. In comparison to simulations without control, holding trains was able to decrease the running times by 7% on average for the segment between Maisons-Laffitte and Houilles-Carrières. That is explained by two reasons:

- First, laisser-faire policy (no control) tends to generate conflicts at the junction and the first train come is the first to pass it. In such situations, a RER A train from Poissy, which is omnibus, could pass before a RER A from Cergy, which does not stop at all stations. As a consequence, this latter is disturbed and strongly delayed. Hence, holding long enough the second train to pass the junction prevents the traffic from being disturbed and allows the trains to run normally,
- Second, as explained in Section 1, departing at the best time from the station allows the trains to avoid meeting warning signals and slowdowning strongly, e.g., in some cases, crossing a warning signal may imply that the running time over a section can last twice longer.

The dwell times measured are, respectively, about 35 and 40 seconds for the simulations without control and with control. This result thus confirms our initial assumption according to which it is preferable to wait longer in station than to waste more time during the run between stations.

Table 1: Simulation results of holding the RER A and L trains in station (NOpt: without optimization, Opt: with train-holding optimization).

		Bench 1 [100%]		Bench 2 [95%]		Bench 3 [92–97%]		Bench 4 [90–100%]	
		NOpt	Opt	NOpt	Opt	NOpt	Opt	NOpt	Opt
#Warning signals met		15.1	0.4	46.2	1.0	52.1	0.7	52.8	0.7
Average running time	Cergy–Nanterre	1,838	1,807 −1.66%	1,979	1,859 −6.06%	2,007	1,856 −7.53%	1,995	1,849 −7.32%
	Maisons-L.–Nanterre	466	443 −5.12%	492	454 −7.67%	492	453 −7.77%	492	452 −8.05%
	Poissy–Nanterre	1,088	1,030 −5.29%	1,100	1,056 −4.04%	1,101	1,054 −4.23%	1,114	1,054 −5.35%

4.3 Operational chain for delivering the necessary dwell time

Figure 5 depicts the functional architecture and the operational chain deployed during the experimentation for delivering the train drivers the departure times when stopped in holding station. The localization of trains is done by the activation of beacons along the tracks when the trains pass. All these events are listened from a data stream by a server which updates traffic information. Indeed, whenever a train passes on a beacon, the arrival and departure times to the next stations are estimated and projected from statistical running times.

The computation of the dwell time of a train stopping in a station is triggered when the last beacon before the station is activated – if the train is not planned to stop, it is a passage event, as for the RER A trains which do not stop neither at Maisons-Laffitte nor at Houilles stations. Concretely, the arrival event is listened from a data stream by the server which updates traffic information as written above and calculates the departure time, which is transmitted to the agent's tablet of the station under consideration. This departure time is then transformed into a countdown. Before it is elapsed (around 15 seconds before the end), the agent tells the train driver that he can initiate the departure sequence, i.e. controlling the platform, closing the doors and looking at the signalling. For safety reasons, the agent on the platform also controls the signalling to avoid giving incoherent orders as a departure on a closed signal.

The experiments were performed during two consecutive weeks (eight exploitable days) during the morning peak [7:30–9:30]. As explained before, four stations were used to hold the trains and control their headways. In the Maisons-L. and Sartrouville stations, there are, respectively, one and two platforms on which agents were required, and that corresponds to one and two tracks and up to two and three tracks, if the train paths had been changed for dispatching reasons. At all, 50 time slots had to be covered and more than 20 agents were involved in the experiments.
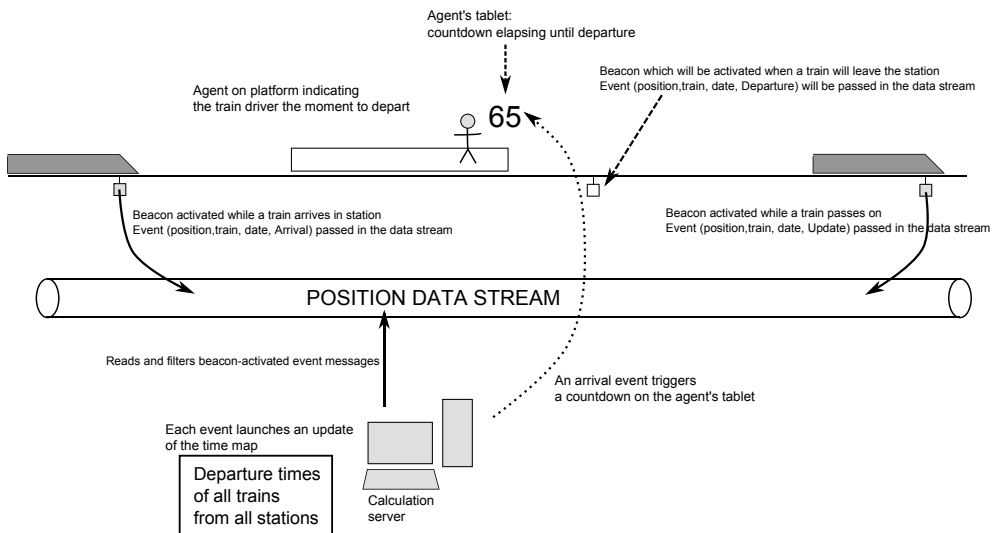

Figure 5: Description of the operational chain deployed during experimentation for updating the traffic information and delivering the necessary dwell times to the trains.

5 ANALYSIS OF THE RESULTS

In order to analyse the results of the experiments, we chose to compare them with those of the period of time directly adjacent to the experiment days (before and after them). The reference period covers 25 days.

Table 2 summarizes the results about the average running times measured on three segments of the experimental area. What emerges from this analysis is that the average running times for two segments are not really modified and that the train-holding has been quite neural for them. But the running time of the Grand-Cormier–Houilles segment was specifically reduced by 32 seconds, which is really significant in mass transit.

If we focus on the standard deviations of the running times, it appears that they are drastically reduced by more than 50% for two segments and 28% for the third. Figure 6a–c illustrates the running time distributions over the segments analysed as well as the Gaussian function drawn on the basis of the mean and the standard deviation expressed in Table 2. We can see unambiguously that the running time distribution during the experiment has been narrowed. That means train-holding has been capable of making flows more regular; in other words, there were less perturbations due to meeting of warning signals. As a consequence, the train dispatching at the junction was improved and the trains from the two branches were better synchronized so that they were able to run on open signals. That is why there is less variance between running times.

Finally, the throughput has been slightly improved by 0.25 trains/h, which is not sufficiently significant to state that it was improved by the train-holding, but at least it was not degraded.

Table 3 focuses on assessing the running times in station with regard to the running times between the stations. We distinguish them as follows. The running time in the station corresponds to the time elapsed from the beacon before to the beacon after the station. The running between stations corresponds to the time elapsed from the beacon after the origin station to the one before the next station.

It has to be observed that the running times in station during the experiments are all longer than those out of the experiments due to the lengthening of the dwell time. It is far more notable for the Grand-Cormier station, where the dwell time was lengthened by 34 seconds. Besides, if we focus on the running times between the stations, it has to be noted that, except for the Maisons-L.–Sartrouville segment, all the running times decreased. The sum of the

Table 2: Running times during the experimentation and during the reference period.

Segment analysed	Reference period (25 days)			Experiment days (8 days)		
	Running time (s)		Throughput [trains/h]	Running time(s)		Throughput [trains/h]
	Average	σ		Average	σ	
Conflans → Houilles	1,026	289	7.8	1,023 (−3)	136 (−53%)	8.06 (+0.26)
Grand-Cormier → Houilles	766	319	4.52	734 (−32)	140 (−56%)	4.75 (+0.24)
Maisons-L. → Houilles	365	101	14.06	366 (+1)	73 (−28%)	14.31 (+0.25)

(a) Conflans – Houilles segment

(b) Grand-Cormier – Houilles segment

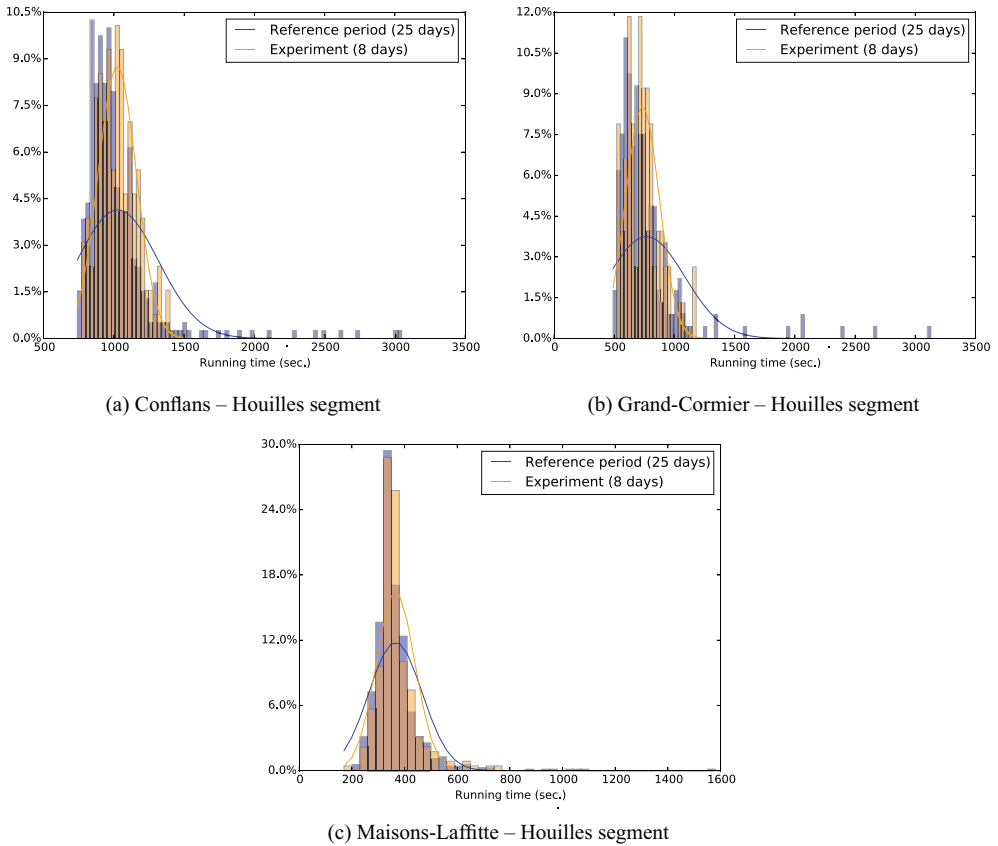(c) Maisons-Laffitte – Houilles segment

Figure 6: Distribution of the running times over three segments of the experimental area: for each distribution is drawn the Gaussian function.

differences (on the running times in stations and on the running times between stations) represents the time difference for riding from the beacon before the station to the beacon before the next station. Except for the Maisons-L.–Sartrouville segment with a 4-second increase, all the running times have been decreased despite the lengthening of the dwell times. It is really significant for the Grand-Cormier–Maisons-L. segment with a lengthening by 34 seconds in Grand-Cormier station, which is more than caught up with a 81-second reduction on the running time to Maisons-L.

The main outcome, which confirms our initial assumption, is that it is more profitable to waste time while waiting in station than to waste more while running to the next station with the risk of crossing a closed signal. The other outcome is also the good behaviour of the train-holding mechanism deployed in the experimental area.

6 CONCLUSION AND PERSPECTIVES

In this article, we presented a heuristic based approach for dealing with railway traffic management in mass transit. The major problem is the lack of fluidity in the traffic flows due to the perturbations occurring in operations. To deal with that, two key issues have to be

Table 3: Running times in origin station (from the beacon before the station to the one after) and between stations (from the beacon after the station to the one before the next station) during and out of the experimental period (RT: running time).

From	To	Reference period (25 days)		Experiment days (8 days)		
		Average RT in origin station (seconds)	Average RT between stations (seconds)	Average RT in origin station (seconds)	Avg. RT between station (seconds)	Difference (seconds)
Achères-V.	Maisons-L.	90	388	96 (+6)	370 (−18)	−12
Maisons-L.	Sartrouville	80	55	83 (+3)	56 (+1)	+4
Sartrouville	Houilles	98	134	103 (+5)	125 (−9)	−4
Grand-Cormier	Maisons-L.	81	313	115 (+34)	231 (−81)	−47

addressed: (1) the accuracy of the computed departure time of a train to set an appropriate headway with the previous train and (2) the responsiveness of the heuristic to compute instantaneously the departure time of the train under control. The heuristic proposed here is based on the basic principles of timetabling and blocking time theory and addresses the formulated key issues for real-time applications.

The heuristic initially tested in simulation has been used in operations during a 2-week experiment, on the RER A and L lines around the Bifurcation de Dieppe. The obtained results showed the capability of the heuristic to improve the flow regularity. Moreover, the performance of the RER A Poissy branch has been rebalanced compared to the other branch, from which the RER A Cergy and the L trains come. We therefore showed that the train-holding during an accurate dwell time can improve the regularity of the flows provided that the time intervals between two consecutive trains are correctly calculated (also in function of the speed control system).

As perspectives, given that the experimental results have been judged satisfying, an industrial study is ongoing to deploy an equipment on the platforms of the experimental area in order to present instructions to the train drivers in operations.

REFERENCES

[1] Van Breusegem, V., Campion, G. & Bastin, G., Traffic modeling and state feedback control for metro lines. *IEEE Transactions on Automatic Control,* **36(7)**, pp. 770–784, 1991. DOI: 10.1109/9.85057.

[2] Fernandez, A., Cucala, A. P., Vitoriano, B. & de Cuadra, F., Predictive traffic regulation for metro loop lines based on quadratic programming. *Proceedings of the Institution of*

Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, **220(2)**, pp. 79–89, 2006. DOI: 10.1243/09544097F00505.

[3] Grube, P. & Cipriano, A. Comparison of simple and model predictive control strategies for the holding problem in a metro train system. *IET Intelligent Transport Systems,* **4**, pp. 161–175 (14), 2010. DOI: 10.1049/iet-its.2009.0086.

[4] Cai, X. & Goh, C.J. A fast heuristic for the train scheduling problem. *Computers & Operations Research*, **21(5)**, pp. 499–510, 1994. DOI: 10.1016/0305-0548(94)90099-X.

[5] Puong, A. & Wilson, N.H.M., A train holding model for urban rail transit systems. *Computer-Aided Systems in Public Transport*, eds. M. Hickman, P. Mirchandani & S. Voß, Springer Berlin Heidelberg: Berlin, Heidelberg, pp. 319–337, 2008.

[6] Sun, A. & Hickman, M., The holding problem at multiple holding stations. *Computer-Aided Systems in Public Transport*, eds. M. Hickman, P. Mirchandani & S. Voß, Springer Berlin Heidelberg: Berlin, Heidelberg, pp. 339–359, 2008.

[7] Pachl, J., Timetable design principles. *Railway Timetable and Traffic*, eds. I.A. Hansen & J. Pachl, Eurail Press: Hamburg, pp. 13–46, 2008.

[8] Corman, F., DAriano, A., Pacciarelli, D. & Pranzo, M., Evaluation of green wave policy in real-time railway traffic management. *Transportation Research Part C: Emerging Technologies*, **17(6)**, pp. 607–616, 2009. DOI: 10.1016/j.trc.2009.04.001.