ABSTRACT: In order to achieve higher capacity on existing infrastructures with dense traffic, railway companies are deploying Communication Based Train Control (CBTC) systems. These systems would considerably facilitate the implementation of new ways for managing traffic, such as headway-based control. However, for commercial reasons, traffic management systems using the timetable adherence principle have to be maintained on low traffic areas; both traffic management systems should hence have to coexist on some lines. The aim of this paper is to present a simulation method for assessing the robustness of operations on such a line, located both in low and dense traffic area. More precisely, the capability of a headway-based control system to resist to disturbances occurring out of its control perimeter is evaluated. The simulation tool reproduces train traffic evolution, disturbances and dispatching decisions, allowing to estimate the quality of service for various disturbance scenarios. Numerical experiments are carried out on one of the most saturated lines in the Paris suburban area. The obtained results prove the practical feasibility of such a mixed traffic management up to a critical level of disturbances.

KEYWORDS: Communication-Based Train Control, headway-based control, railway operations, real-time traffic management

1 INTRODUCTION

1.1 Railway operations in dense urban areas

Most railway companies operating services in densely populated areas have been experiencing an increase of demand for passenger transportation during last years. As a result, railway networks reach saturation, making it difficult to answer the demand. In addition, as the number of trains grows, operations become more sensitive to disturbances, leading to a poor quality of service for passengers. The so-called three-aspects signaling system, used in many countries for ensuring the operations safety, plays a particular role in disturbances propagation in the network. This system allows to maintain a sufficient headway between two consecutive trains. It works the following way. The track is divided into blocks, each of them protected by a signal. A green signal indicates to the train driver that the next two blocks are free, and that he can proceed at the same speed. A yellow signal means that the next block is free but the following one is not, and the driver must therefore prepare to stop before the limit of the current block. Finally, a red signal imposes the driver to stop as the next block is occupied.

Figure 1 shows how this system may lead a delay to propagate as knock-on delays along a line. If a train dwells in a station for a longer time than the nominal one (due e.g. to a jammed door), the following trains are likely to catch up and meet a yellow signal. This
will force them to slow down and result in even higher delays for these trains.

Figure 1 – Propagation of delays by a fixed signaling system

When designing timetables, railway companies usually insert buffers between consecutive trains to prevent the propagation of delays. But this comes at a price, as it has the effect of reducing the capacity of an already saturated network while not guaranteeing against delay propagation. Operating companies seek therefore to improve the system capacity and reliability without resorting to heavy and costly infrastructure modifications.

1.2 Communication-Based Train Control

Communication-Based Train Control (CBTC) systems are intelligent signaling and vehicle control systems that help to reduce the effect of the phenomena described in the previous section (IEEE Standard 1474.1 (2004)). These kind of systems have been widely implemented on metro lines (such as line 14 in Paris or line L in New York). Vehicles equipped with CBTC are precisely located on the network thanks to telecommunication systems, independently of the track detection devices. CBTC also integrates an Automated Train Operation (ATO) system, that controls speed and acceleration of the train during its driving phases, and an Automated Train Supervision (ATS) system that performs automatic traffic management decisions. As a result, trains are able to automatically adapt their speed depending on the position on the preceding train. Train protection is hence ensured by a moving block signaling system. That allows much shorter headways than the fixed three-aspects signaling system, and reduces the phenomenon of delay propagation described in the previous subsection, resulting in higher capacity.

1.3 Headway-based traffic control

Most railway companies operate trains according to a timetable, even in dense and highly circulated areas. When a disturbance occurs that results in delays, train drivers and traffic managers are required to be

have in such a way that the trains arrival and departures occur as close as possible to their planned times. However, this is not always the most efficient way to operate commuter trains. Indeed, surveys show that when the train frequency gets sufficiently high, passengers do not pay attention to schedule and present themselves at the station with the purpose of boarding the first train available (Luethi et al. (2007)). In addition, this operation mode does not guarantee a good quality of service to passengers. This is particularly true at points where a good synchronization is required, such as junctions.

Figure 2 – A junction

Table 1 – Drawbacks of schedule-based traffic control

Consider two trains running on the network described in Figure 2 according to the schedule given by Table 1. Suppose that train 1 suffers a 5-minute delay and that train 2 is on schedule. Then a conflict for infrastructure utilization will occur at the junction, and one of the trains will have to stop on the track, causing inconvenience to passengers and increasing delays. It might be better to allow train 2 to depart ahead of schedule if possible, or to hold it at Station A’ for a time long enough to prevent the conflict.

In addition, Van Breusegem et al. (1991) proved that uneven time intervals between train can damage the stability of operations if no dispatching decision is made. Indeed, a train arriving at a station a long time after the previous train will have to dwell a longer time, as more passengers will wish to board. The headway between trains will consequently grow, leading to an even higher dwell time at the next station and a higher delay, while the following train is likely to catch up with the delayed one. This phenomenon is well-known of bus passengers and operators, where it is called “bus bunching” (Newell et al. (1964)). Finally, we can note that reducing the headways variance leads to a smaller expected waiting time for passengers (Ding et al. (2001)).

For those reasons, many bus and light rail operators have given up schedule-based control to prefer headway-based traffic management. In this framework, vehicles do not have a pre-planned schedule, but are monitored in real-time in order to maintain even headways between them. Up to our knowledge,
This concept has been little applied to railway services so far; the implementation of such a framework would be considerably facilitated by the CBTC systems. However, it does not seem feasible to equip a whole suburban line, due to the high cost of CBTC systems. Nor is it practicable to operate a line with branches with headway-based control in totality. Indeed, the capacity limitations of the common section would result in lower frequencies on the branches, where a schedule is needed for commercial reasons. As a result, the option of operating a line with headway-based control coupled with a CBTC system on the dense common section, and schedule-based control with manual driving on the other parts of the line, is considered.

The aim of this paper is to assess the robustness of such a traffic management scheme. For that purpose, we present a simulation method for reproducing railway traffic under headway-based traffic management performed by a CBTC system. We specifically study the robustness of such an operations system, i.e. its capability to resist to small disturbances occurring outside the control area. The remainder of this paper is organized as follows. In section 2, we briefly review some related work. Section 3 is dedicated to the description of the modeling framework and the simulation method we developed. Numerical experiments are presented in section 4, where we apply the method to a saturated line of Paris suburban network. Section 5 concludes the paper.

2 RELATED WORK

The real time railway traffic management problem has been widely studied and many approaches have been proposed for solving it; a review of these methods is presented in Cacciani et al. (2014). Among them, some consist in recovery algorithms after minor disturbances. In that case, only rescheduling decisions (such as re-timing arrival and departures events, re-ordering trains at junctions) need to be taken. Other consider the case of recovering a feasible schedule after a major disruption, for which the rolling stock and crew have to be rescheduled as well, but this falls out of the scope of our study. Most of the designed algorithms rely on a mathematical programming formulation of the problem with a no-wait job shop scheduling problem as studied by Mascis et al. (2002). D’Ariano et al. (2007) propose a branch-and-bound algorithm for computing a conflict-free schedule in real time after a perturbation. Pellegrini et al. (2015) consider the case where decisions can be taken in a fixed control area and describe a method for reacting to disturbances occurring outside this area. However, most of the works address the case where train protection is ensured by a fixed block signaling system and trains are operated according to a sched-

ule. D’Ariano et al. (2008) investigate the concept of flexible timetables to improve real-time traffic management, making thus a first step towards headway-based control in railway systems.

Headway-based traffic management, using in particular holding strategies, have been studied for urban transportation systems such as buses (see for example Sun et al. (2008), Daganzo (2009), Berrebi et al. (2017)). The aim is to provide methods for keeping headways as even as possible so that approximately the same amount of passengers board each bus at a given stop. However, the constraints are not the same as for railway systems; the specific problem of managing junctions is not considered. This issue is addressed by Schanzenbächer et al. (2017), who propose a model based on max-plus algebra to prove that under some hypothesis, the traffic evolves toward an equilibrium, and to compute its characteristics.

The development of CBTC systems has opened a new area of research, concerning management and simulation of such systems. Ning et al. (2015) propose a control method for minimizing headway variance and energy consumption by selecting adapted speed profiles. Pochet et al. (2016) designed a genetic algorithm to reschedule trains when those are sharing infrastructure with other vehicles that are not equipped with a CBTC system. A tool for simulating traffic on a corridor is presented by Chen et al. (2010).

Finally, note that several railway traffic simulation tools already exist. They are used for various purposes, such as ensuring that a timetable is conflict-free, passenger information or robustness assessment of a timetable. The commercial softwares OpenTrack and RailSys rely on a microscopic description of the infrastructure and signaling system, and are notably used for scheduling. However, these softwares require a high computation time and are more adapted to lines operated with according to a schedule. Li et al. (2005) propose a simple cellular automaton model for simulating railway traffic with a moving block signaling system.

3 MODELING RAILWAY OPERATIONS AND HEADWAY BASED CONTROL

3.1 Description of trains and infrastructure

In this preliminary work, we restricted ourselves to the simulation of the area where trains are operated according to a headway-based control scheme. We model the infrastructure at a mesoscopic level, by a network \((V, A)\). The set of arcs represents homogeneous segments of tracks, linking nodes that are the singular points (stopping points in stations, junctions, points where the speed limit changes). To each arc \(a \in A\), we associate a length \(\ell_a\), a speed limit \(s_a\),
and a capacity $c_\theta$ (meaning the highest number of trains that can run on this segment simultaneously). In addition, following Pellegrini et al. (2015), we allow some arcs to be reserved by a train before this train enters it, guaranteeing that no other train will enter the reserved infrastructure during the reservation. This is used for determining train orders at junctions.

The set of trains to be operated is denoted $T$. To each train $\theta \in T$ we associate its physical characteristics that are its length $\ell^\theta$, its acceleration coefficient $\alpha^\theta$ and its braking coefficient $\gamma^\theta$. The service of train $\theta$ is described by the following parameters: an itinerary $I^\theta$ that is the ordered sequence of arcs taken by the train, a servicing $D^\theta$ that is the sequence of the scheduled dwell times at each node of the network (these times being null for the nodes that do not represent a station), and the time at which the train is scheduled to enter the network $\text{start}^\theta$. The train evolution is described by: the state of the train (running, dwelling, not operating), the arc $a^\theta$ on which the head of the train is located, its position $x^\theta$ on the arc, its speed $s^\theta$, its preceding train $\text{prec}^\theta$ and its current dwell time $\text{dwell}^\theta$.

### 3.2 Eulerian scheme

During the simulation, the trains evolution is computed using an Eulerian scheme, with a time horizon $T_f$ and a time step $dt$. The core algorithm can be synthetically described in the following way:

\[
\text{t=0 ;}
\]

\[
\text{while t < T_f do}
\]

\[
\text{for all trains } \theta \in T \text{ do}
\]

\[
\text{Compute safety distance ahead of train } \theta ;
\]

\[
\text{Update position and speed for train } \theta ;
\]

\[
\text{Enforce dispatching decisions ;}
\]

\[
\text{end}
\]

\[
\text{end}
\]

**Algorithm 1:** Description of the core algorithm

We subsequently present each of the three functions performed by this algorithm. First, as train protection is ensured by a moving block signaling system, it is necessary to compute the distance between a train and the next point where it has to stop. We can then deduce the maximal speed which allows the train to stop before that point. The computation of this safety distance is performed by Algorithm 2. We assumed here the arcs of $I^\theta$ to be ordered according to the itinerary of train $\theta$.

The trains’ position is updated along the simulation using an Eulerian scheme. At each time iteration $t$, a new speed instruction is computed for each running train: that is the speed the train will reach at time $t + dt$ (assuming a constant acceleration between $t$ and $t + dt$). The position is then updated accordingly. Algorithm 3 describes the performed operations. Note that if a train cruises at speed $s$, according to the law of uniform deceleration, the minimum distance for it to stop is $s^2/2\gamma$. Therefore, if a train has safety distance $\lambda$, the maximum speed it can reach while still being able to stop before that point is $\sqrt{2\gamma\lambda}$. However, as speed cannot be adapted instantly, we compute a speed instruction for the next time step; the safety distance should be therefore reduced to $\lambda - sdt$, and the maximum speed ensuring safety $\sqrt{2\gamma(\lambda - sdt)}$. We distinguish between two cases:

- Either the train is able to speed up while still satisfying the safety constraints. Then it accelerates to the highest possible speed according to its acceleration coefficient, maximum speed of the track segment and safety conditions. Note that the maximum speed instruction satisfying the latter is the only solution of equation $s = \sqrt{2\gamma(\lambda - sdt)}$, that is $\sqrt{2\gamma\lambda + \gamma^2(\lambda - sdt)^2} - \gamma dt$.
- Or the train has to slow down, and brakes with deceleration coefficient $\gamma$.

### 3.3 Disturbances

Let us recall that our goal is to provide a simulation method for assessing the robustness of a traffic management system where a transition between schedule-based and headway-based control is operated. Our
method seeks therefore to reproduce the disturbances that are susceptible to occur in such a system. We identified two categories of disturbances:

- Disturbances occurring outside the headway-based control area. We do not delve into the details of such disturbances, and model them by an arrival of trains at the entrance of this area with uneven time intervals. This is done by adding to the nominal arrival times the realizations of independent uniform random variables, with a magnitude to be chosen as a parameter.

- Prolonged dwell times at stations located in the dense area (due e.g. to passengers blocking doors or driver waiting for passengers to board).

We introduced these disturbances as realizations of independent random variables following uniform distributions. This choice is motivated by the lack of available data for estimating the probabilistic distribution of such phenomena; this is indeed a prospective work and no railway system is currently operated in such a paradigm. We therefore used ad hoc hypothesis coming from operators’ experience. However, some refinements are possible: Yuan (2006) established that, under some hypothesis, arrival times of trains follow a log-normal distribution.

3.4 Heuristics for railway traffic management

Last, the simulation method must be able to reproduce traffic management decisions. For the sake of simplicity and performance in terms of computation time, we implemented sub-optimal rule-based heuristics rather than more elaborated methods for solving the real-time traffic management problem. The possible decisions are the following:

- All trains may depart stations as soon as possible (No Holding policy), or some trains might be held longer in stations in order to re-adjust the headways between trains in case of disturbances (Holding policy). This policy might be particularly efficient in stations located ahead of junctions, thus preventing trains to compete for the use of infrastructure and to be forced to wait for the junction to be free.

- When passenger data is unknown and traffic is homogeneous (all trains stop in the same stations), it seems reasonable to manage junctions according to a First In First Out (FIFO) policy. However, when direct and omnibus trains share the same infrastructure, our method proposes a Threshold policy: when a direct and an omnibus train are likely to compete for passing the junction, the direct train is given priority if the resulting delay for the omnibus train remains below a fixed threshold.

4 NUMERICAL EXPERIMENTS

4.1 Case study: line D of Paris suburban network

We applied our method to assess the feasibility of such a mixed traffic control on one of the most saturated lines in Paris suburban area, namely line D of the RER (Réseau Express Régional, the local express network). A map of the dense part of this line is given by Figure 3. We assumed this part to be equipped with CBTC and operated with headway-based control.

Line D crosses Paris for linking Orry-la-Ville (on the north) to Corbeil-Essonnes and Melun (on the south). This line presents several peculiarities. First, it shares
its infrastructure with line B of the RER network, another saturated line, between the stations Paris-Nord and Châtelet-les-Halles. For that reason, this common segment is one of Europe’s busiest track sections, with a traffic reaching 32 trains per hour during the peak period. Second, in order to increase the offer on the southern part of line D despite the saturation of the common section, some trains coming from the south end their journey at Paris-Lyon station and operate from there a new service towards the south. Two tracks are dedicated for these turnaround services, while two other tracks are used for those running through the station. Figures 4 and 5 present respectively a map of these track sections.

The transport offer is organized as follows in peak hours:

- 20 trains per hour on line B in each direction;
- 12 trains per hour on line D from the north to the south and from the south to the north;
- 8 trains per hour on line D from the south to Paris-Lyon and from Paris-Lyon to the south.

The nominal time interval between trains in the common section is 90 seconds. Trains of line B are scheduled every three minutes, while trains of line D follow a 6 min/6 min/3 min pattern in the common section. South of Paris-Lyon, the nominal time interval is three minutes. Figure 6 represents on a time-space diagram the nominal trajectories of trains from North to South; the blue, green and red lines are associated respectively to the trains of line B, the trains of line D from the north to the south and the trains of line D from Paris-Lyon to the south. Trains of line D and line B are respectively 246 meters and 208 meters long, and we assumed both acceleration and braking coefficients equal to 0.7 m/s².

Our aim is to assess the capability of the system to resist to disturbances that may occur upstream of the headway-based control area. Figure 7 illustrates what may happen if trains enter this area without complying by the nominal interval (because of disturbances that were not recovered by traffic management). We can observe on that figure some phenomena of congestion ahead of the junction between lines B and D at the station of Paris-Nord: indeed, some trains are forced to stop ahead of the station. This results in an increase of travel time for passengers. We can also note that some direct trains (red lines) get stuck behind omnibus trains south of Paris-Lyon and are therefore slowed down as well. For example, unlike Figure 6, there is no spare time window for allowing the first red train to insert on the south branch of the line. Consequently, this train has to follow the fifth green train that is a slower one, and gets delayed.
For each level of disturbance and each policy, we performed 100 replications. The Holding policy is applied in the stations located upstream of junctions only, and according to the following rule (we reuse the notations of Figure 2). We denote by $\Delta_{AB}$ (respectively $\Delta_{AB}'$) the nominal travel time between station $A$ (respectively station $A'$) and station $B$, and $I$ the nominal time interval between both trains. If train $1$ leaves station $A$ at time $dep_{A}$, train $2$ will not be allowed to leave station $A'$ before the time $dep_{A'} = dep_{A} + (\Delta_{AB} - \Delta_{AB}') + I$, thus ensuring that train $2$ will not arrive at station $B$ with a shorter interval than the nominal one. Results are plotted on Figure 8.

The junctions of Châtelet les Halles and Paris-Nord are managed according to a FIFO policy while the junction of Paris-Lyon follows a Threshold policy (priority given to the fastest trains up to a 30s consequent delay for the other train). Indeed, both direct and omnibus trains use the infrastructure south of Paris-Lyon while all trains are omnibus on the northern part of the line.

The distributions of delays using respectively the Holding and the No Holding policies are plotted on figure 8. We note that for both policies, delays remain limited both in average and extreme values for small levels of disturbances (up to approximately 45 seconds), but grow and become more unpredictable for higher levels. However, delays grow at a non-exponential rate: this is mostly due to the moving block signaling system that is known to alleviate the risk of delay propagation from one train to the next, compared to a fixed signaling system.

In addition, up to a certain point the Holding Policy allows to reduce extreme values of delays, but in most cases this comes at the price of a higher travel time in average. For greater levels of disturbances the Holding Policy seems to become unable to cope even with the extreme values. This work consequently does not appear conclusive on the relevancy of implementing such a policy and further studies seem to be needed.

This study also underscores the effects of stochasticity on the quality of operations in the dense area. A specific attention should therefore be paid to the schedule adherence in the upstream low traffic areas, where the system is less sensitive to disturbances, as small delays in these parts of the network can lead to instability in the common section. In case of delays occurring in the branches, traffic management decisions such as canceling a whole train or part of its service could be made to prevent propagation to the dense area, but this possibility was not investigated in this work.

5 CONCLUSION

Our approach demonstrates the practical applicability of such a mixed traffic management for a railway line crossing a dense area, the system being able to resist to disturbances of small magnitude. It also highlights the benefits and drawbacks of a Holding policy to adjust headways.

Yet, some questions remain to be answered. First, the obtained results emphasize the need of an efficient traffic management; improved control strategies should therefore be tested. In particular, the possibility of adjusting headways by controlling the trains speed between stations in addition to their dwell times could be explored. Second, random disturbances upstream of the dense area were assumed to be independent and uniformly distributed, which in practice is usually not the case: our results might benefit to be confirmed by a refined model. Last, no studies have been made so far on how to build schedules and manage traffic in real time in the upstream and downstream areas to facilitate the transition between schedule-based and headway-based control. All these topics could be matters for further research.
REFERENCES


