Public Transport (2019) 11:1–25
https://doi.org/10.1007/s12469-019-00196-y

CASE STUDY AND APPLICATION

Delay and disruption management in local public transportation via real-time vehicle and crew re-scheduling: a case study

Federico Malucelli[1] · **Emanuele Tresoldi**[2]

Accepted: 24 January 2019 / Published online: 11 February 2019

Abstract

Local public transport companies, especially in large cities, are facing every day the problem of managing delays and small disruptions. Disruption management is a well-established practice in airlines and railways. However, in local public transport the approaches to these problems have followed a different path, mainly focusing on holding and short-turning strategies not directly associated with the driver scheduling. In this paper we consider the case of the management of urban surface lines of Azienda Trasporti Milanese (ATM) of Milan. The main issues are the service regularity as a measure of the quality of service, and the minimization of the operational costs due to changes in the planned driver scheduling. We propose a simulation-based optimization system to cope with delays and small disruptions that can be effectively used in a real-time environment and takes into account both vehicle and driver scheduling. The proposed approach is tested on real data to prove its actual applicability.

Keywords Delay management · Disruption management · Local public transportation · Real-time optimization · Real-word scenario · Big data · Vehicle scheduling · Crew scheduling

✉ Emanuele Tresoldi
emanuele.tresoldi@unimi.it

Federico Malucelli
federico.malucelli@polimi.it

[1] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy

[2] Dipartimento di Informatica, Università degli Studi di Milano, Milan, Italy

1 Introduction

One of the key elements affecting the perceived quality of the local public transport is the *regularity* of the service. The service provider, in accordance with the municipality, relates the so-called regularity of the service to how much the actual service adheres to the planned one. This adherence is computed considering the difference between the actual and the planned headways at all stops and the actual distance covered by on-duty vehicles. This regularity, however, is continuously threatened by disruptions or small disturbances. These inconveniences induce further disruptions when the service provider has to comply with the regulations of driver shifts. The final effect is a loss of reliability of the transportation system and additional costs due to penalties to be paid to the municipality and to extra allowances for drivers.

The daily operations of transit companies are often monitored by an operation central office taking advantage of automated vehicle monitoring (AVM) systems and mobile telecommunication devices. In the case of ATM of Milan, each operator visually controls the operations of one or more lines on a screen reporting in real time the vehicle positions on a map. The operators can detect from the screen delays or anomalies that may generate disruptions, or they collect information from drivers about troubles on the line such as vehicle breakdowns, accidents or medical and safety emergencies. In the presence of a disruption involving a vehicle or part of the infrastructure, the operator decides the actions to be taken coordinating the behavior of the drivers involved in the disruption or in the recovery actions. The basic actions that an operator can adopt in case of disruptions or disturbances are, for example, to perform vehicle short-turns, to delay a vehicle, or to cancel portions or complete trips. In exceptional cases the operator may also decide to use spare drivers or vehicles.

Even though operators are very good in recovering disruptions or preventing them, they usually have not the intuition on the effect of their interventions on resource management. Indeed they are not enabled to have a full understanding of the effect of their actions on the vehicle and crew scheduling. Moreover, the task of manually adjusting the planned scheduling goes beyond the duty of the operators. Thus the effect of recovering a disruption in the morning may have an even worse disruptive effect later on when drivers duties come to an end. Another key issue is to recognize potential disruptive conditions as early as possible, when the recovery actions can have a smaller impact on the service management.

These issues call for a thorough analysis of the problem and the study of optimization methods to be included into a decision support tool to assist operators in taking decisions. This decision support tool can fruitfully take advantage of the huge quantity of real-time data made available by the AVM.

Disruption management in transportation is a very well studied subject in airlines (Clausen et al. 2010) and in railways (Jespersen-Groth et al. 2009; Cacchiani et al. 2014). Disruption management in local public transport has followed a different path. The first approaches date back to the early seventies, see for example Newell (1972). These works, and the following ones, see Ibarra-Rojas et al. (2015) for a comprehensive literature review, are mainly focused on the headway regularity and the bus bunching phenomenon, proposing a variety of control strategies. Very limited attention has been devoted to the crew rescheduling issues that are actually critical in our case study, especially when they are connected with the regularity of the service. This makes the disruption management problem of our case closer to those arising in airlines and railways, though the presence of additional degrees of freedom opens for different types of approaches.

Our contribution is twofold. On the one hand, we analyze the disruption-delay management of public transport integrating the three components of service regularity, vehicle scheduling and crew scheduling. On the other hand, we propose a tool that, exploiting the real-time data, helps the operators in facing the problem with the ultimate objective of increasing the quality of service. The solutions produced by the tool use the available resources to propose in real-time adjustments of vehicle and crew scheduling with a clear assessment of their impact both on the service quality and on the operational cost.

The paper is organized as follows. In Sect. 2 we briefly classify disruptions in public transport and we analyze the related works. In Sect. 3 we present the case study and the current approach of ATM. In this section we also point out the main weaknesses of the current way of detecting disruptive conditions, based on the way the regularity of the service is measured, and we propose some alternatives. In Sect. 4, we introduce the overall framework to detect and manage disruptions and disturbances. The framework is based on a discrete event simulator, used to forecast the evolution of the service, and on a tabu-search algorithm for the real-time vehicle re-scheduling. In Sect. 5, we present a column generation method used to tackle the real-time crew scheduling re-optimization. Extensive computational experience on real data is presented in Sect. 6 for both single and multiple lines.

2 Delay and disruption management

2.1 Classification of disruptions and measures of regularity

According to Clausen et al. (2010), a disruption is “a deviation from the original plan sufficiently large to require a substantial change in operations”. In this definition, two important concepts are considered: the *planned service* usually produced by the off-line execution of an optimization procedure, and the *observed service*. When the observed service is too different from the planned one, we say that we have a disruption. However, not every deviation from the original plan requires a *substantial* change in operations, and indeed identifying such cases becomes a delicate task.

An alternative and complementary term used in case of minor disruptions is *disturbance*. We say that a line has a disturbance in case some vehicles are operating with limited differences with respect to the planned timetable. Disturbances can, in general, be reduced and solved by carefully holding, speeding up, or slowing down vehicles. Nevertheless, if the disturbances are not controlled, they can easily evolve into severe disruptions giving rise to the so called “snow-ball effect”.

Disruptions are usually classified depending on their causes (e.g., accidents, medical emergencies, vehicle breakdowns, drivers arriving late at the relief point, severe traffic congestion, critical weather conditions, etc.), and on their medium and long-term effects. In this work we will disregard the so-called *Planned disruptions*, that is, changes in the service caused by planned works, as well as the so-called *Long-term disruptions* which means those arising in the presence of severe weather condition forecasts (Mároti 2013). The type of disruption involved in our study are the *disruptive events*: when a portion of the network is blocked, as for example due to demonstrations, accidents or line breakdowns, vehicle breakdowns or driver illness. All these events involve a heavy adjustment that implies rerouting lines or canceling trips, trying to guarantee a service to all passengers. In this case the main concern is to provide alternative services to passengers. There are some contributions as for example Jin et al. (2015), Kepaptsoglou and Karlaftis (2009), Kiefer et al. (2016) and Cadarso et al. (2013). In the case of ATM these events are managed manually by officers. Their main interest arises once the blocking event has been removed and the service has to be recovered in the shortest possible time by rearranging the vehicle and crew scheduling. However, we will consider in particular the disruptions caused by minor events or by the accumulation of small delays (*minor disruptions or disturbances*). In the presence of disturbances, most of the times, it is possible to maintain and/or recover the regularity of a single line by suitably regulating vehicle speeds and managing holding times at the terminals. If the rest times at the terminal are not mandatory, and the location where the vehicle stops technically allows it, they can be shortened or lengthened in order to improve the regularity. These problems are often referred as the *bus holding*, see for example Hernández et al. (2015) and Hickman (2001). They can be approached either in a centralized way or in a self-regulating way (Cats et al. 2012; Xuan et al. 2011) and can also be addressed in real-time taking into consideration dynamic running times and demand (Sánchez-Martínez et al. 2016). Moreover, the *endogenous disruptions*, that are not due directly to external causes, but are related to the personnel shifts, will play a central role in our analysis. Indeed even if the service is apparently regular, as far as headways are concerned, the accumulated delays may disrupt the driver scheduling. Some drivers may result excessively late at their relief point or at the end of their shift. This type of anomaly is typically solved by abruptly interrupting the trip so that the driver can reach the relief point in time. This is one of the main causes of the regularity degradation at ATM and is not currently managed by the officers of the operation room. Thus, we will mainly focus on minor disruptions, disturbances and endogenous disruptions, and the recovery phase of disruptive events. In particular, we will focus on methods to detect at an early stage the potential disruptions and to adapt in real-time the vehicle and the crew scheduling.

2.2 Related works

Disruption management has been studied in the airline industry in the last 20 years, see for example Johnson et al. (1994); Lettovskỳ et al. (2000) and Clausen et al. (2010). In this case the high operational costs and the dramatic consequences of disruptions on the management as well as on passengers gave a strong motivation to the research.

Also railways have put some attention on the problem (Jespersen-Groth et al. 2009). In this case, the most diffused approaches take advantage of some results obtained in job scheduling. Vehicles correspond to jobs, and trips/infrastructures correspond to machines (Raheja and Subramaniam 2002; D'Ariano et al. 2008). In this context, very often, the complexity of the constraints, mainly due to the shared infrastructure, suggested hybrid approaches conjugating optimization techniques with simulation, as for example in Berger et al. (2011).

The disruption management of local transport systems has initially focused on the headway regularity, dealing with the so-called bus holding problem. That is, devising strategies for holding or pacing vehicles in order to increase the regularity (see Daganzo 2009; Xuan et al. 2011). In Eberlein et al. (1998, 2001), the real-time problem is studied in a deterministic setting and the objective is to minimize the passenger waiting time at stops. The case of a single line with self-coordinating vehicles, knowing the headway with respect to the previous and the next vehicles, is proposed in Bartholdi and Eisenstein (2012). In the bus holding problem, the issues deriving from crew scheduling are not considered. A brief survey of the contributions of this type is summarized in Yang and Li (2012).

Vehicle rescheduling problems are also studied in connection with passenger logistic issues. In the presence of a disruption involving vehicles or portions of the network, vehicles must deviate providing an alternative service to passengers with enough capacity (Li et al. 2009). In other cases new alternative lines have to be designed as in Kiefer et al. (2016), or portions of the network must be reconnected by "bridging" services (extensive examples in Cadarso et al. 2013; Kepaptsoglou and Karlaftis 2009; Jin et al. 2015), or delays must be managed in order to guarantee the connections (see Giovanni et al. 2014).

As for the solution methods, the most common approach is to apply the same techniques used in the off-line planning phase to the so-called *core problem*. That is a problem restricted to the disrupted services and a few more. The core problem is then tackled by a variety of methods ranging from integer programming (Nissen and Haase 2006), column generation (Huisman 2007; Potthoff et al. 2010), multicommodity flow (Wei et al. 1997), or constraint programming (Jespersen-Groth et al. 2009). All proposed methods stress the fact that problems must be solved in real-time since they arise at the operational level.

Usually, the vehicle re-scheduling and the crew re-scheduling problems are tackled separately. In Visentini et al. (2014) the contributions on vehicle re-scheduling problems are surveyed and classified depending on the type of transportation. In many cases, the attention is devoted to the crew schedule recovery and the modified vehicle schedule is taken as an input. This happens when the service has a particular structure (e.g. hub and spoke in case of airlines) or when the alternative solutions for vehicles are so few that vehicle re-scheduling can be easily managed manually. Or vice-versa, the attention falls on the vehicle re-scheduling (as in railways or some road transportation cases) because it is by far more complicated than the crew re-scheduling one. Some papers, as Walker et al. (2005), deal with the simultaneous recovery of disruptions in both vehicle and crew scheduling; however, they refer to a peculiar single track railway setting, where train takeovers and crossings must be considered carefully. Also, Huisman and Wagelmans (2006) tackled the vehicle and crew rescheduling problem jointly, showing the benefits of a combined approach in terms of effectiveness. However, the high computational times do not allow for a real-time application in practical cases.

One of the main issues is to detect unrecoverable disruption as early as possible when recovery actions can be taken at a reasonable cost and with a low impact on passengers and crews (see Abdelghany et al. 2004).

All the above-mentioned works constitute an important source of inspiration, both from the methodological and practical standpoint. However, none of them addresses completely the main aspects of the disruption management problem faced by ATM. In particular, the regularity of the service in connection with the crew scheduling, the early detection of potential disruptive conditions with special attention to the endogenous ones and the real-time vehicle and crew rescheduling.

2.3 Disruption identification: evaluating the regularity of the service

The main challenge in managing disruptions is to be able to distinguish between the events or conditions that may have a negative impact on the service regularity and those whose effect is limited and needs no special attention. In order to infer this information, we formalize the concept of service regularity and relate it to measurable indices. Therefore, whenever we identify a condition where the index of regularity can be substantially improved, we are in the presence of a potential disturbance/disruption. The concept of regularity depends on the type of service offered by public transport. We distinguish two ways of offering public transit:

1. Timetabled service: In this case, the frequencies are usually low. At each stop of the line the precise time when the vehicle will arrive and depart (timetable) is specified.
2. Frequency-based service: In this case, the frequencies are higher, thus the headway is rather small. At each stop, the headway is specified instead of the timetable (e.g., a vehicle every 7 min). Note that, in this case, even in the presence of generalized delays, the service is not perceived as disrupted by passengers, provided that frequencies are regular and aligned with the planned ones.

The threshold on the headway that distinguishes these two types of service falls between 10 and 12 min. In this paper, we concentrate on frequency-based services. Indeed our system takes full advantage from the fact that a disrupted service may be perceived as regular, if the headways are evenly spaced. However, due to changes in the vehicle scheduling, it might be required to adjust the driver scheduling in real-time, in order to avoid to generate endogenous disruptions. Our approach could be applied also to a timetabled service, though the effects would be much more limited.

There are many ways to estimate the regularity of the service. The most regular service is obviously that reproducing exactly the planned timetable. Thus the regularity measure should consider the adherence of the provided service with the planned one. In the case of a timetabled service, the measure will consider the planned timetable, while in the frequency-based service this measure can be relaxed, and only the headways will be accounted for. In the literature many proposals are present (see Barabino et al. 2013 for a brief survey). Among the many indices of the literature we will consider the following ones, that will be compared with our proposed ones.

HR Headway ratio: the ratio between the observed headway and the planned one.
HS Headway standard deviation: standard deviation of the difference between the observed and the planned headways.
PR Percentage regularity deviation: the percentage average ratio between the deviation of the observed headway from the planned one and the planned headway.

Refer to Sect. 5.3 for the formal definition.

In the next section we consider, for our case study, how ATM is currently measuring the regularity of the service, enlightening possible pitfalls. Moreover, we propose alternative indices that are better adaptable to the real-time re-optimization of vehicle and crew scheduling.

3 The ATM case

The Azienda Trasporti Milanesi (ATM) is one of the largest local public transport companies in Italy. It serves an area of about 657 km^2 with about 2.5 million inhabitants and an overall volume of 691 million of passengers and 147 million of traveled kilometers per year. The surface services involve 103 bus lines, 19 tram lines and 4 trolley bus lines covering a network of more than 1000 km and managing a fleet of about 2000 vehicles.

The service regularity index used in the agreement with the municipality works as follows. For the sake of simplicity, let us examine a single line and let n_e be the number of planned stops during the whole service period.

3.1 Current index of regularity and anomalies

A so-called *bad pass* occurs every time a vehicle arrives more than t minutes late with respect of the planned headway, or in case of a skipped stop. Let n_b be the number of bad passes on a line. The index of regularity of the line is the ratio between the number of regular passes and the total:

$$I(01) = \left(\frac{n_e - n_b}{n_e}\right). \tag{1}$$

Note that this type of index does not account for early passes. This is intentional, since, considering the actual headway, a vehicle anticipating its pass at a stop, while the other vehicles are maintaining their schedule, will result as an early pass with respect to the previous vehicle. However, the next vehicle will generate a bad pass considering the actual headway. Thus penalizing the late pass as well as the subsequent early one (of the same amount), would penalize the same event twice. Moreover, this type of regularity index penalizes, in the same way, an almost good pass (i.e. a vehicle missing the regularity of a few seconds) and a very bad one (i.e. a vehicle with a headway much higher than the threshold). This could have some sense in an ex-post evaluation environment, as it is currently used. While, in a setting where the service provider wishes to increase the index of regularity, this type of measure may generate some pathological behaviors.

Indeed, in the presence of a group of vehicles of the same line following each other at a short distance (*bus bunching* phenomenon), only the first vehicle gives rise to bad passes at all stops, according to $I(01)$. Moreover, the attempt of improving the regularity by detouring a vehicle of the bunch, so that it closes the large gap in front of the bunch, could worsen the index $I(01)$, though the service, as perceived by the users, would improve. An automated system that uses $I(01)$ to improve the regularity, in the presence of disturbances, would let the service converge towards bus bunching cases. In addition, from that situation it would be impossible to escape by applying single detour actions, since any single action would not be able to improve the regularity index.

3.2 Alternative ways to evaluate the service regularity

In order to overcome the pitfalls of $I(01)$, we introduce some other indices that share the basic idea of $I(01)$. However, the new indices account for both late and early passes, with the intention of using them more profitably in an automated system that aims at improving the regularity by modifying the vehicle schedule.

3.2.1 Piecewise linear function for evaluating the index of regularity

Consider the gap $x(q)$ between the planned headway of pass q ($v_e(q)$) and the observed one ($v_o(q)$):

$$x(q) = v_o(q) - v_e(q). \tag{2}$$

A negative value of $x(q)$ means an early pass, while a positive value means a delay.

Let us define the function $f(x(q))$ of the gap $x(q)$ as follows:

$$f(x(q)) = \begin{cases} -\alpha x(q) & \text{if } x(q) < -\theta_1 \\ 0 & \text{if } -\theta_1 \leq x(q) < \theta_2 \\ \beta x(q) & \text{if } \theta_2 \leq x(q) < \theta_3 \\ \gamma x(q) + \delta & \text{if } \theta_3 \leq x(q) \end{cases} \tag{3}$$

where $\theta_1, \theta_2, \theta_3 (> \theta_2)$ and $\alpha, \beta, \gamma (> \beta), \delta$ are suitable parameters. The function $f(x(q))$ is 0 if the pass is regular and it is greater than 0 (not necessarily equal to 1) if the pass is irregular. If we want to ignore the contribution of earliness on the index of regularity, it suffices to set $\alpha = 0$. Note that this index includes also the simple $I(01)$ if we set $\alpha = \beta = \gamma = 0$ and $\delta = 1$. The contribution of the function $f(x(q))$ due to values $x(q) \geq \theta_3$ intends to penalize large gaps more than the equivalent sum of small gaps.

The index of regularity based on the piece-wise function is:

$$I(PW) = \sum_{q \in P} f(x(q)) \tag{4}$$

where P is the set of all passes that actually occurred in the observed period. Note that the set P does not include all planned passes as in the ATM index. Thus, if a stop is skipped due to vehicle short-turns it is not considered in the set P.

3.2.2 Quadratic function for evaluating the index of regularity

The idea behind the quadratic penalty function stands on the intention of accounting for lateness (and earliness) more than linearly. This implies that a large lateness with respect to the planned headway will be penalized more than the equivalent sum of small delays. The quadratic penalty function is defined as follows:

$$f'(x(q)) = \begin{cases} \eta' x(q)^2 & \text{if } x(q) < \delta_1 \\ 0 & \text{if } \delta_1 \leq x(q) < \delta_2 \\ \eta'' x(q)^2 & \text{if } x(q) \geq \delta_2 \end{cases} \tag{5}$$

The interval $[-\delta_1, \delta_2]$ defines a tolerance zone: vehicles arriving at the stop not later than δ_2 and not earlier than δ_1 with respect to the headway are not penalized. The index of regularity based on the quadratic function is:

$$I(QA) = \sum_{q \in P} f'(x(q)) \tag{6}$$

3.3 Considering driver shifts in the measure of regularity

As mentioned earlier, the divergence of the service timetable from the planned one is one of the main causes of endogenous disruptions, especially when driver shifts are concerned. For this reason, the introduction of another index accounting for the regularity of the driver schedule could be of help in making decisions. Let $\mathcal{D}$ denote the set of drivers on duty, and for every driver d let R_d be the set of possible relief points. Note that in R_d a relief point may appear more than once since a driver can pass by the relief point more than once. Hence in practice, relief points in R_d are space/time occurrences. Let us assume that the initially planned relief point (original site and original time) has index 1 for all drivers. Let ET_r^d be the planned passing time at the relief point r of driver d and OT_r^d be the observed one. The adherence to the planned schedule of driver d can be measured as the smallest difference in time at any relief point in R_d:

$$y_d = \min \left\{ \theta_r^d = OT_r^d - ET_1^d \, : \, \theta_r^d > \epsilon, r \in R_d \right\} \quad d \in \mathcal{D} \tag{7}$$

where ϵ is a suitable tolerance threshold introduced to avoid too much earliness in the alternative relief points. Notice that ϵ can be also fixed to a (small) negative value. In addition, we can account for the inconvenience caused by the fact that a driver has to change the site of the original relief point, that is if the minimum in (7) corresponds to a relief point spatially different from the original one. Let z_d be a 0–1 indicator that equals one if there is such a change. A third term is the number of lost kilometers with respect to the planned service. This is evaluated considering the difference between the planned kilometers of each driver and the expected ones. This can be easily computed by considering the difference between the planned relief point and the actual one, not only considering the difference in the site but also that in time. Let s_d denote such a difference in kilometers.

A parametric measure of the adherence of the scheduling with respect to the planned one is:

$$IA = \phi \sum_{d \in \mathcal{D}} y_d + \xi \sum_{d \in \mathcal{D}} z_d + \eta \sum_{d \in \mathcal{D}} s_d \tag{8}$$

where ϕ, ξ and η are suitable non-negative trade-off parameters.

4 Disruption-delay management framework

Currently, the officers take their decisions about recovering the regular service relying upon their expertise. The effects of these actions on the service regularity are evaluated only according to their experience and intuition, and it is almost impossible for them to assess the impact of alternatives. Determining the effect of the recovering actions on the driver schedules is even more difficult, and this is actually the most critical issue for ATM management.

However, having defined a set of indices to evaluate the regularity of the service, detecting potential disruptive conditions and evaluating the impact of possible actions on the scheduling becomes possible. To do that we draw up a discrete simulation tool that forecasts the system behavior and allows us to compare the effect of different solutions.

The actions that can be taken are:

- forcing vehicle short-turns, that is, stopping a vehicle and detouring it in a suitable position in the line, in the same direction or in the opposite one. This implies that passengers must alight and take the following vehicle of the same line.
- holding vehicles at stops or terminals.
- shortening (or skipping) breaks at terminals.
- micro-adjusting the vehicle speed.

In order to feed the simulator with the complete information, all possible actions must be identified, in particular for the short-turns it must be specified between which two stops the action may take place, and how much time it needs.

The simulation is intended to evaluate the effect of the actions within a sufficiently wide interval of time. To this aim, we compare the regularity index output by the simulation when some actions are made with that coming from the simulation when no action is carried out.

The simulator routine is based on a set of events referring to the vehicle positions along the time. The routine considers a simulation interval Δ starting from the current moment. Initially, the set of events contains the last AVM observation for each vehicle on duty, and the events corresponding to new vehicles entering the service within the simulation interval. The new events are generated from the current ones updating the position of the vehicles. Position and time are randomly generated by using the travel time distribution. A discussion about the travel time distribution obtained from historical data can be found in Malucelli and Tresoldi (2018). Within the simulation routine, the indices of service regularity are updated according to the functions defined in the previous sections.

Ideally, the set of actions maximizing the improvement of the regularity index should be identified. To solve this problem we used a tabu-search procedure (Glover and Laguna 1997). The tabu-search starts from the current vehicle and crew schedule and explores neighboring solutions by trying all possible actions selecting them from a set $\mathcal{A}$. The output of the tabu-search is a subset $\mathcal{W}^* \subset \mathcal{A}$ of actions, which, if applied, will allow with high probability to improve the index of regularity obtained without applying any action. To produce a more reliable result the simulation is repeated several times.

Once a set of actions $\mathcal{W}^*$ that increases the regularity of service in the simulation has been identified, the driver scheduling is adapted. This is done by solving an optimization problem with the objective of minimizing the changes with respect to the planned duties and minimizing uncovered service. The description of the overall working framework of the delay/disruption management system can be found in Carosi et al. (2015).

5 Real-time driver re-scheduling

Once the vehicle schedule has been modified, both due to delays and as a result of the possible improving actions, the crew schedule has to be adapted. One possible approach is to apply the algorithm used for planning the service, though it must be applied to a very limited sub-problem, with respect to the original one. However, the characteristics of the sub-problem, in terms of size, objectives and constraints suggest an ad hoc approach. Indeed the main objective is no longer to minimize the cost, but to minimize the changes with respect to the planned service, in order to simplify the implementation of the actions. Moreover, the union regulations applied in the planning phase can be partially relaxed. Another important aspect is that in the re-scheduling sub-problem we have to account for the fact that drivers are already on duty and they expect to work in the same period as from the initial planning. Notice also that it may happen that, with the new vehicle schedule, some trips remain uncovered. These cases can be managed if the uncovered trips appear in the end of a vehicle duty or if they involve the use of the so-called “hot” spare drivers, that are officially on duty waiting for emergency calls. This implies that, though in a limited extent, the re-scheduling problem has to consider also changes in the vehicle schedule, limiting the inconvenience on the users.

5.1 Problem statement

We are given a set of drivers and a set of vehicle duties each one defined as a set of consecutive trips and some additional operations. Each vehicle duty is split into pieces of service corresponding to continuous driving periods assigned to a driver. Each piece of service has a starting time and an ending time, and a starting and ending location that may be a depot or a relief point on the line (example in Fig. 1). For each driver we are given the remaining maximum working and driving times and the maximum number of remaining pieces of service. In addition, since the decisions are taken when the service is carried out, we also know the assignment of drivers to the currently worked pieces in case they are on duty. The problem consists in finding an assignment of the pieces of service to the drivers, maintaining the assignment of the currently worked pieces so that it complies with the regulations. The regulations account for the maximum duration of the working shift, the maximum total driving time, the maximum uninterrupted driving time, the starting and ending location of the driver shift, the number, and duration of breaks between pieces of work. The objective function accounts for the cost of possible driver extra allowances and the amount of uncovered service. Notice that, in this assignment, previously planned duties may be disassembled into atomic pieces of service and then assembled again in a different way to form new duties that are more convenient to cover the service.

5.2 Graph representation

We utilize a directed acyclic graph $G = (N, A)$ to support the modeling. The node set N is partitioned into a set of drivers N_d, including possible spare ones, a set of pieces of service N_h, and some special nodes N_b used to manage breaks. The set of arcs A represents the compatibilities between pieces of service. There is an arc $(j, j') \in A$ if the pieces of service j and j' can be assigned consecutively to the same driver. Moreover, there are arcs (d, j) connecting each driver d to pieces of service compliant with the regulations that can be assigned to him/her. Nodes representing breaks are connected to any other node. One example of graph G is represented in Fig. 2).

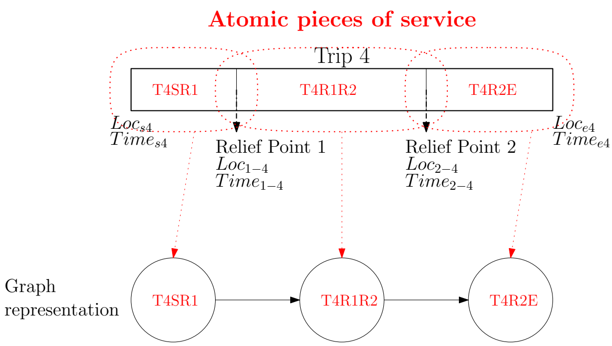

Fig. 1 Graph representation of a part of vehicle duty

Notice that, if one driver d is currently on duty, there are only arcs (d, j) such that j is the current piece of service assigned to d or its possible modifications (shortened or lengthened pieces of service derived from the original one). A path in G starting from a node d in N_d and satisfying the additional constraints deriving from regulations, corresponds to a feasible completion of duty for driver d. The driver re-scheduling problem corresponds thus to finding a path covering G using feasible paths starting from "driver" nodes.

5.3 Column generation approach

The most successful approaches to the off-line version of the crew scheduling problem are based on column generation (see for example Desaulniers et al. 2005; Gualandi and Malucelli 2013). Moreover, with this approach part of the complexity of the problem, due to the involved regulations, is transferred to the pricing subproblem. Indeed, the pricing subproblem can be easily tailored to take into account additional features and requirements.

The idea of the approach is to use two sets of binary variables x_p and y_h. Variables x_p are associated to each feasible completion (i.e., feasible path) p and are used in a set partitioning framework. Variables y_h are associated with each (atomic) piece of service h (which is a unsplittable portion of vehicle duty) and are set to 1 if h remains uncovered. Let E_{hp} be the incidence matrix of the path p, that is $E_{hp} = 1$ if and only if path p covers the piece of service h and let B_{dp} be the driver-path association matrix that is $B_{dp} = 1$ if driver d is associated with path p. Let $\mathcal{H}$ denote the set of atomic pieces of service. The crew re-scheduling optimization problem (CRP) is then formulated as follows:

$$CRP: \quad \min \sum_{p \in \mathcal{P}} c_p x_p + \sum_{hin\mathcal{H}} c'_h y_h \tag{9}$$

$$\sum_{p \in P} E_{hp} x_p + y_h = 1 \quad \forall h \in \mathcal{H} \tag{10}$$

$$\sum_{p \in P} B_{dp} x_p \leq 1 \quad \forall d \in \mathcal{D} \tag{11}$$

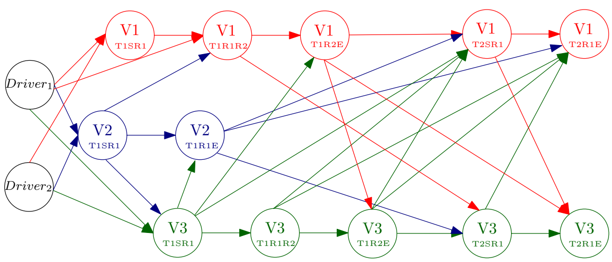

Fig. 2 Example of compatibility graph used in the pricing subproblem. Break nodes are not included for the sake of readability

$$x_p \in \{0, 1\} \quad \forall p \in \mathcal{P} \tag{12}$$

$$y_h \in \{0, 1\} \quad \forall h \in \mathcal{H} \tag{13}$$

where $\mathcal{P}$ is the set of feasible completions and c_p denotes the overall cost of a completion. This cost c_p is equal to the number of minutes of extra work required by completion p plus 0.1 for each of the following modifications with respect to the planned duty: starting location, ending location, relief point. The penalty for skipping a piece of service c'_h is equal to twice the number of minutes required to complete the piece of service h. The objective function (9) aims at the minimization of the total cost given by the sum of pieces of services not covered, extra working time and driver duties modifications. Note that the last one is usually not a direct cost for the company. The first constraint (10) states that each piece of service h is either covered by exactly one compatible completion or is skipped and consequently its corresponding y_h variable is set to 1. This ensures the feasibility of the problem even if, due to the modified vehicle schedule, some piece of service is not covered by any driver. The second constraint (11) establishes that a driver cannot be assigned to more than one completion (duty).

In our implementation of the column generation method the restricted master problem (RMP) is given by the linear relaxation of the CRP and it is initialized with a set of columns made up of zero cost columns corresponding to all original driver shifts. For each driver, taking into consideration the new vehicle schedule, we generate a column containing all feasible pieces of service that the driver was supposed to complete during his/her original duty. Pieces of service falling outside the original working shift are left uncovered.

Since many of the columns composing the best solution are obtained by removing or adding a single node to an original driver shift, the initial set is integrated with columns generated by a simple algorithm that is able to find basic feasible modifications of the original drivers shifts. This procedure is run once for each driver and makes use of a restricted graph (see Fig. 3). In the graph, each driver node is connected only to nodes representing the pieces of work associated with vehicles he/she is supposed to carry out in the originally planned shift. Moreover, for each vehicle, only the nodes compatible with the time and location limits of the original driver shift are considered, though a small tolerance in the starting and ending time

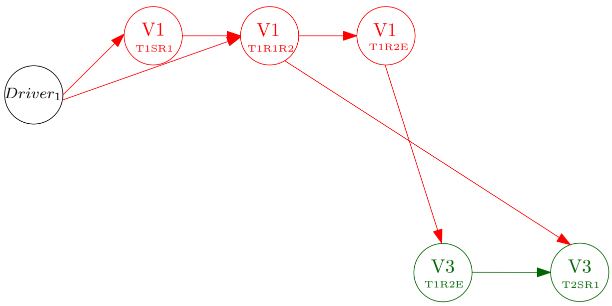

Fig. 3 Example of a reduced graph used in the initialization procedure

of each part of the driver shift can be accounted for. On this graph the algorithm, following a greedy approach, looks for a path reaching as many nodes as possible. Only two decisions have to be taken in this problem: where to start and where to change vehicle (when possible). The algorithm generates one path for each available starting location (one or two at most), it changes the vehicle in the last feasible node (compatible with breaks regulations, time and location limits) and reaches the new vehicle in the first possible location. The algorithm runs in linear time in the size of the restricted graph and generates up to $2 \times |\mathcal{D}|$ columns.

At each step of the column generation procedure the RMP is solved and, exploiting the dual representation of the problem, a set of new columns p corresponding to x_p variables with negative reduced cost is found solving a pricing subproblem and it is added to the RMP. This process stops when no useful column is found in the pricing phase.

5.4 Pricing subproblem

In details, the dual master problem of the CRP read as follows:

$$\max \sum_{h \in \mathcal{H}} (\mu_h + \tau_d) + \sum_{d \in \mathcal{D}} \nu_d + \sum_{p \in \mathcal{P}} \sigma_p \tag{14}$$

$$\sum_{h \in \mathcal{H}} E_{hp} \mu_h + \sum_{d \in \mathcal{D}} B_{dp} \nu_d + \sigma_p \le c_p \quad \forall p \in \mathcal{P} \tag{15}$$

$$\sum_{h \in \mathcal{H}} \mu_h + \tau_h \le c'_h \quad \forall h \in \mathcal{H} \tag{16}$$

$$\nu_d \le 0 \quad \forall d \in \mathcal{D} \tag{17}$$

$$\sigma_p \le 0 \quad \forall p \in \mathcal{P} \tag{18}$$

$$\tau_h \le 0 \quad \forall h \in \mathcal{H} \tag{19}$$

where dual vectors μ and ν correspond to constraints (10) and (11) while dual vectors τ and σ are associated with the additional constraints $x_p \le 1 \quad \forall p \in \mathcal{P}$ and $y_h \le 1 \quad \forall h \in \mathcal{H}$ coming from the linear relaxation of (12) and (13).

The expression of the reduced cost ζ_p associated with primal variable x_p in terms of dual variables is:

$$\zeta_p = c_p - \sum_{h \in \mathcal{H}(p)} \mu_h - \sum_{d \in \mathcal{D}(p)} \nu_d - \sigma_p \tag{20}$$

where $\mathcal{H}(p)$ represents the set of atomic pieces of service included in path p and $\mathcal{D}(p)$ is the associated driver. The pricing subproblem looks for the path p that minimizes ζ_p and satisfies the constraints. The search of a negative reduced cost completion can be modeled as an elementary shortest path problem with resource constraints. The four resources considered in the problem are the total working time, the total driving time, the uninterrupted driving time and a single break. It is worth noting that, while three resources are characterized by a monotonic consumption function, the costs and the uninterrupted driving time have a non-monotonic behavior (e.g. the uninterrupted driving time is reset when a break is consumed).

The pricing problem is solved with a standard dynamic programming method based on Feillet et al. (2004). This problem can be decomposed into independent subproblems, one for each driver, that can be solved in parallel.

When the pricing is not able to return a negative reduced cost path for any driver, there is no variable that can potentially improve the value of the linear relaxation of the re-scheduling problem (RMP); the integer CRP is solved to optimality using all columns found. This is obviously a heuristic approach. A complete Branch and Price scheme should be applied to guarantee the optimality. This choice is due to the limited time available in the real-time setting.

6 Computational results

In order to evaluate the performance and effectiveness of the procedure, we have implemented the complete *Delay Management Framework* in C++ using COIN-OR CBC 2.9.7 (Lougee-Heimer 2003) as MIP solver. All tests were done on an Intel Core i7-3770K with 16 GB of RAM running Microsoft Windows 10 Pro (64 bits).

6.1 Data set

The data set used in the experimental phase includes 215 real scenarios involving two lines, in particular, ATM tram line 14 (105 instances) and trolley bus line 92 (110 instances). The operation period covers the months of June, September and November 2015 from 07:00 to 20:00. Line 14 counts 104 stops, up to 25 vehicles simultaneously on duty, about 66 drivers, 16 relief points and up to 16 potential improving actions for every trip. The average round trip traveling time is about 2 h and 10 min, the planned headways go from 5 min during peak hours to 9 min in off-peak hours. Line 92 has an average round trip traveling time of about 1 h and 35 min and planned headways ranging from 4 min in rush hours to 9 min during off-peak hours. This line has 66 stops, up to 20 vehicles simultaneously on duty, 50 drivers, 5 relief points and up to 5 possible improving actions for each trip.

In all scenarios the simulation interval Δ has been set to 2 h, the length of the tabu list is equal to 3 and 10 rounds of simulation are considered (parameters f_{max} and k in the tabu-search algorithm). The software is expected to produce a solution for vehicle re-scheduling in less than 5 min.

In the crew re-scheduling phase tolerances are taken into account in order to simulate the real behavior of drivers. In particular, for each driver, the starting time of the shift considers a tolerance of 10 min (a driver must be at the starting location of the shift at least 10 min in advance) and the total duty time can include up to 20 min of extra work with respect to the planned one. The execution time of the crew re-scheduling is limited to 5 min.

6.2 Column Headers

In Tables from 1 to 9 the column headings have the following meaning:

- PF: service regularity index used for the optimization.
  - 01 ATM regularity index.
  - *PW*: Piece-wise linear regularity index.
  - *QA*: Quadratic regularity index.
  - *HR*: Headway ratio-based regularity index, percentage ratio between the observed headway and the planned one. See Sect. 5.3 for a formal definition.
  - *HS*: Headway standard deviation-based regularity index: standard deviation of the difference between the observed and the planned headways. See Sect. 5.3 for a formal definition.
  - *PR*: Percentage regularity deviation-based regularity index: the percentage average ratio between the deviation of the observed headway from the planned one and the planned headway. See Sect. 5.3 for a formal definition.
  - $HY - *$: Hybrid regularity index composed by the combination of the function estimating the impact of the actions on the drivers scheduling and the $*$ regularity index.
- %I: percentage improvement in the regularity, with respect to IR_0, measured with the regularity index used for the optimization.
- Avg %I: average percentage improvement in the regularity, with respect to IR_0. This is the average value obtained evaluating both the starting situation and the solution found with all available penalty functions.
- %No: number of instances where no improving actions were found.
- N Ac: number of improving actions in the solutions found (max. allowed improving actions: 10).
- KM-Ac: lost Km due to actions, e.g. portion of the itinerary left uncovered due to a short-turn or a trip limitation.
- V T: computational time for the vehicle re-scheduling algorithm (tabu search), in seconds.
- KM-D*: lost Km due to the driver scheduling adjustment.
- %D-D*: percentage difference between the length of the optimized driver duties and the planned ones.

Table 1 PF comparison: line 92

PF	%I	Avg %I	%No	N Ac	KM-Ac	V T
01	3.281	0.849	12.727	3.314	2.209	5.073
PW	59.219	1.705	4.545	5.874	1.758	5.172
QA	1.359	2.897	4.545	5.600	1.695	4.745
HR	0.687	− 19.624	0.000	9.630	2.690	5.214
HS	6.392	2.089	1.818	8.061	2.387	5.049
PR	4.104	1.704	0.000	7.600	2.180	5.115

- %M-D*: percentage of drivers having changes either in the relief points, or in the starting/ending location, or the duty extended by more than 10 min.
- Extra*: extra working time, in seconds, required in the optimized scheduling. Each individual driver duty can be extended by at most 20 min as requested by ATM.
- D T: computational time for the crew re-scheduling algorithm (column generation), in seconds.

6.3 Comparison of regularity functions

In this section, we analyze the performance of the disruption-delay management framework and crew rescheduling using different regularity functions. In particular, we consider PW and QA, introduced in this paper, 01 by ATM and HR, HS and PR, taken from the literature. Since PW and QA functions have several parameters that can affect the behavior of the system, a preliminary computational experiment to determine suitable parameter settings is described in Malucelli and Tresoldi (2018).

Let $v_o(q)$ and $v_e(q)$ denote the observed and the expected headways in stop q and let P be the set of all observations. The index of regularity for the literature functions is computed in the following way:

$$
\begin{aligned}
HR =& 1 - \left|\frac{\sum_{q\in P}\frac{v_o(q)}{v_e(q)}}{|P|} - 1\right|;\\
HS =& 1 - std(\{v_o(q), \forall q \in P\})/avg(\{v_e(q), \forall q \in P\});\\
PR =& 1 - \frac{\sum_{q\in P}|v_e(q) - v_o(q)|/v_e(q)}{|P|}.
\end{aligned}
$$

The ATM 01 regularity function has $t = 180[s]$.

The results obtained on line 92 and 14 are reported in Tables 1, 2 and in Tables 3 and 4, respectively. We tested the system on single lines for two main reasons. On the one hand, we wanted to set up the system and verify its efficiency in a limited context, even though the lines that we studied are among the most complex and critical of the service network. On the other hand, ATM usually manages resources (vehicles and drivers) independently on each line, thus generating data on the crew scheduling would have been artificial. We will extend the experiments to a multi-line case in Sect. 5.5.

Considering columns *%I* and *Avg%I* regularity functions 01 and *HR* perform very poorly and generate solutions that actually decrease the regularity of the service when the results are evaluated with any other function. Moreover, function 01 is not able to find any improving solution in almost 13% of the instances. *HR* always proposes a large number of actions, even if it provides very little improvement in the regularity of the service. All the other indices are able to improve the overall quality of the service. They provide comparable performances from the regularity of the service point of view and propose a similar

Table 2 PF comparison: line 92

PF	KM-D*	%D-D*	%M-D*	Extra*	D T
01	0.628	0.792	4.288	295.872	0.295
PW	0.253	1.014	6.584	406.147	0.372
QA	0.095	0.980	6.912	337.379	0.435
HR	1.080	1.171	9.690	970.430	1.249
HS	0.469	0.860	8.218	379.551	0.816
PR	0.390	0.803	7.062	364.800	0.829

Table 3 PF comparison: line 14

PF	%I	Avg %I	%No	N Ac	KM-Ac	V T
01	5.690	– 4.927	13.33	4.821	17.791	24.647
PW	53.181	4.760	1.905	6.937	7.468	24.338
QA	6.715	6.928	4.761	6.750	9.697	24.481
HR	3.302	– 15.218	0.000	9.741	36.531	24.835
HS	23.812	3.180	0.000	8.580	11.679	24.586
PR	6.575	5.438	0.000	8.296	10.407	24.540

Table 4 PF comparison: line 14

PF	KM-D*	%D-D*	%M-D*	Extra*	D T
01	8.761	2.240	11.867	1684.746	33.160
PW	4.139	0.918	7.535	1469.456	37.656
QA	4.395	1.074	7.465	1507.289	34.806
HR	14.877	1.723	24.884	3797.926	60.142
HS	6.790	1.591	13.501	1631.630	58.597
PR	6.519	1.016	11.484	1553.914	58.349

number of actions. However, there are differences when the cost of implementing the solution is taken into account. The actions proposed in solutions provided by *PW* and *QA* have a lower impact on the cost than those found with *HS* and *PR* (see column *KM-Ac*). As for the evaluation on the impact on the driver scheduling (columns *KM-D**, *%D-D**, *%M-D** and *Extra**), the solutions proposed when using *PW* and *QA* have a better performance with respect to *HS* or *PR*. Finally, it is worth noting that, due to this fact, the computational burden to re-optimize the driver scheduling for the outcome of *HS* and *PR* is almost double than that of *PW* and *QA*. In conclusion, the disruption-delay management framework that uses *PW* or *QA* functions provides solutions that are comparable with those of *HS* and *PR* in terms of regularity, while the solutions are by far more efficient in terms of cost and difficulty of implementation.

6.4 Hybrid function

Combining ideas shown in Sects. 2.2 and 2.3 a new hybrid index *HY*, taking into account simultaneously the index of regularity and the index IA (see Sect. 2.3) measuring the adherence of the crew scheduling to the planned one, can be defined as follows:

$$HY = PF + \phi \sum_{i \in \mathcal{D}} y_i + \xi \sum_{i \in \mathcal{D}} z_i + \eta \sum_{i \in \mathcal{D}} s_i$$

where *PF* is the basic index of regularity computed with any function presented before. This hybrid function has been tested on all available data with parameters: $\phi = 0.10$, $\xi = 0.05$ and $\eta = 0.15$. Values for y_i, z_i and s_i are estimated from a simple solution of the problem described in Sect. 4. For this estimation, no pricing is executed and once the master problem is initialized with the heuristic procedures described it is solved considering all generated columns and including integrality constraints. The solution of the master problem in this phase can require up to 0.1 seconds.

The results obtained on all available instances using all penalty functions described in this document are reported in Tables 5, 6, 7 and 8 grouped by line.

It is worth noting that the relative performance differences between penalty functions remain the same since the performance of the hybrid function strictly depends on the basic function. Poor functions (*01*, *HR*) generate bad results even inside this hybrid framework. However, comparing these tables with the corresponding results detailed in the previous sections (see Tables 1, 2,3, 4) a few considerations can be drawn. The impact of the hybridization on the service regularity improvement (%I) is minimal and it is greatly compensated by the increase in the robustness of the generated solutions (*Avg %I*). This means that the solutions found using any basic function in the hybrid system are more likely to be considered good solutions even when the regularity of the service is re-checked expost with a different function. The number of improvement actions used in hybrid and regular solutions is similar, but the type of improvement actions is different. In particular, solutions generated with hybrid function favor low-impact actions such as shortening/lengthening breaks instead of route deviations and short-turning trips. This produces a strong reduction (30% on average) in the kilometers lost due to improving actions (*KM-Ac*) and it is achieved with, on average, less

Table 5 Hybrid PF comparison: line 92

HY-PF	%I	Avg %I	%No	N Ac	KM-Ac	V T
HY-01	3.210	0.555	16.363	5.866	1.622	5.286
HY-PW	57.090	3.438	4.545	6.800	1.537	5.265
HY-QA	1.347	2.954	14.545	6.238	1.131	5.021
HY-HR	0.540	− 14.575	0.000	7.390	0.440	5.031
HY-HS	6.225	3.055	7.272	7.837	2.381	5.101
HY-PR	3.879	2.228	4.545	7.463	2.126	5.199

Table 6 Hybrid PF comparison: line 92

HY-PF	KM-D*	%D-D*	%M-D*	Extra*	D T
HY-01	0.366	0.368	3.182	387.793	0.724
HY-PW	0.095	0.761	6.327	413.411	0.823
HY-QA	0.000	0.470	3.890	398.702	0.768
HY-HR	0.000	0.647	3.325	773.490	0.772
HY-HS	0.261	0.632	6.406	365.880	0.972
HY-PR	0.158	0.551	4.632	333.779	0.807

Table 7 Hybrid PF comparison: line 14

HY-PF	%I	Avg %I	%No	N Ac	KM-Ac	V T
HY-01	5.346	− 7.541	19.048	6.852	11.082	23.691
HY-PW	50.680	8.680	1.905	7.962	6.620	25.411
HY-QA	5.702	7.099	5.714	7.787	4.440	24.743
HY-HR	2.188	− 10.552	0.000	9.099	12.617	23.514
HY-HS	20.005	6.957	0.952	7.975	7.188	24.104
HY-PR	5.672	6.128	1.905	8.025	8.127	24.352

Table 8 Hybrid PF comparison: line 14

HY-PF	KM-D*	%D-D*	%M-D*	Extra*	D T
HY-01	1.475	0.543	4.597	1324.180	27.617
HY-PW	3.304	0.509	6.075	1650.405	41.960
HY-QA	1.240	0.360	3.635	1539.787	25.044
HY-HR	6.049	0.682	12.270	3008.481	46.598
HY-HS	2.175	0.483	4.939	1464.463	44.468
HY-PR	2.734	0.444	7.243	1614.911	46.614

than one second increase in the computational time of the tabu search algorithm (column *V T*).

Furthermore, the best improvement is obtained in the reduction of the implementation cost of the solutions. Indeed, three of the four main cost indicators (columns *KM-D**, *%D-D**, *%M-D**) are reduced by more than 50% on average while the average extra working time employed does not change (column *Extra**). Finally, it is worth noting that with a hybrid function, since the number of improving actions in the solutions is lower, the associated crew re-scheduling problems are easier to solve and require on average 20% less computational time.

6.5 Multi-line optimization

Since the delay-disruption management system is effective and efficient in dealing with single lines, we consider now the case where multiple lines can be optimized at the same time. In particular, we focus on cases where drivers are shared between different lines with some relief points in common. In this case, we could not obtain real data. Thus, starting from lines 14 and 92 (line structure, frequencies, relief points, driver duties, etc.) we have randomly generated ten instances in order to simulate three different lines with independent vehicles and shared drivers.

The results obtained on these instances, using the PW regularity function with the best settings defined in Malucelli and Tresoldi (2018), are reported in Table 9. We have considered two different types of optimization:

Single: the three lines are optimized independently. *V T* and *D T* report the sum of the three independent computational times required to find the solution for the three lines; *KM-D** and *Extra** are the sums of all km lost and extra seconds used on the three lines; *%D-D** and *%M-D** are computed taking into account all drivers duties for the three lines.

Multi: vehicles are optimized independently while all drivers are considered as part of a single super-line encompassing all three lines.

The results highlight a huge cost improvement when drivers from multiple lines are shared. This is mainly due to the great reduction in lost kilometers, 48% on average. Moreover, the average increase in driver duties length and in the percentage of driver duties to be modified are also reduced by 19% and 15%, respectively. The extra time, on the other hand, is increased by 14% on average. However, this

Table 9 Multi-line optimization

Scenario	V T	KM-D*	%D-D*	%M-D*	Extra*	D T
1-Single	291.49	64	1.83	63.90	6923	616.87
1-Multi	320.85	14	0.11	3.08	3621	231.15
2-Single	45.03	10	0.47	0.00	18	1.52
2-Multi	48.20	0	0.17	0.00	18	16.36
3-Single	15.47	25	1.56	27.31	5063	0.15
3-Multi	15.96	21	1.54	21.15	6923	0.30
4-Single	60.30	28	1.22	8.33	3464	6.16
4-Multi	69.67	18	0.74	19.05	6705	60.21
5-Single	14.18	41	1.98	58.25	6119	0.15
5-Multi	14.43	32	2.11	29.41	7227	0.43
6-Single	10.99	15	1.53	44.61	3345	3.24
6-Multi	11.32	0	1.45	29.76	3270	21.00
7-Single	15.73	41	1.98	58.25	6119	0.25
7-Multi	15.88	32	2.11	27.45	7227	0.43
8-Single	14.87	25	1.56	27.31	5063	0.16
8-Multi	15.07	21	1.54	21.15	6923	0.40
9-Single	9.08	15	0.57	29.89	3409	2.86
9-Multi	9.13	0	0.53	28.57	3270	17.23
10-Single	44.78	0	0.14	0.00	126	12.26
10-Multi	50.99	0	0.14	0.00	126	43.58

is expected since we are able to cover more kilometers and the contribution in the objective function for extra time is smaller with respect to that of lost kilometers.

As for computational times, on the one hand, in the vehicle optimization procedure times are similar. There is a little overhead when resources must be split among different lines, but it is usually smaller than 10%. On the other hand, the size of the driver re-scheduling problem is definitely increased when multiple lines are optimized together, requiring on average double the time used for independent lines.

7 Conclusions

In this paper, we tackled the real-time disruption-delay management problem in local public transportation considering the real case of ATM. The paper presents the following main contributions:

1. We defined new indices of regularity that can overcome the pitfalls of the index currently used by ATM when applied within an automated re-scheduling system with the purpose of improving the performance.
2. We introduced and implemented the *disruption-delay management framework* integrating vehicle and crew real-time re-scheduling for one local public transport line.
3. We performed an extensive testing campaign on more than 200 real-world scenarios using our newly introduced indices of regularity as well as a few indices taken from the literature. The results obtained demonstrate the feasibility and effectiveness of our approach.
4. We introduced a hybrid index of regularity that takes into account vehicles and crews simultaneously. We demonstrated that this hybrid index provides an improvement in the regularity of the service almost as good as with the other ones and it is able to greatly reduce the cost due to the crew re-scheduling.
5. We proved the flexibility and scalability of the presented framework by solving some multi-line scenarios with three different lines sharing drivers.

This study shows the feasibility of vehicle and crew re-scheduling in real-time on real scenarios and represents a solid starting point for the implementation of an actual online real-time decision support system to help control room operators in their daily work. A preliminary version of this support tool has been implemented and it is currently tested on some lines of ATM.

However, there is room for improvements and further developments. In particular, the reliability of the discrete event simulator can be improved taking explicitly into account additional elements such as expected traffic flow, weather conditions, and drivers behavior to dynamically modify the empirical distribution used for the generation of traveling times. This can be very beneficial, especially when the evolution of the transportation line is simulated over a time horizon of several hours.

Acknowledgements We would like to tank Stefano Gualandi, Fabrizio Ronchi and Stefano Scotti for their contribution that greatly improved the quality and completeness of this paper. MAIOR s.r.l. had a fundamental role in the analysis of the problem and in defining the data retrieval tools. We are gratefully indebted with the reviewers for their constructive comments and helpful suggestions. This work has been awarded the “Best OR application 2016”, by the Italian Association of Operations Research (AIRO).

References

Abdelghany A, Ekollu G, Narasimhan R, Abdelghany K (2004) A proactive crew recovery decision support tool for commercial airlines during irregular operations. Ann Oper Res 127(1–4):309–331

Barabino B, Di Francesco M, Mozzoni S (2013) Regularity diagnosis by automatic vehicle location raw data. Public Transp 4:187–208

Bartholdi JJ III, Eisenstein DD (2012) A self-coordinating bus route to resist bus bunching. Transp Res B 46:481–491

Berger A, Hoffmann R, Lorenz U, Stiller S (2011) Online railway delay management: hardness, simulation and computation. Simulation 87(7):616–629

Cacchiani V, Huisman D, Kidd M, Kroon L, Toth P, Veelenturf L, Wagenaar J (2014) An overview of recovery models and algorithms for real-time railway rescheduling. Transp Res B 63:15–37

Cadarso L, Marín Á, Maróti G (2013) Recovery of disruptions in rapid transit networks. Transp Res Part E Logist Transp Rev 53:15–33

Carosi S, Malucelli F, Tresoldi E (2015) Delay management in public transportation: service regularity issues and crew re-scheduling. Transp Res Proc 10:483–492

Cats O, Larijani A, Ólafsdóttir Á, Burghout W, Andreasson I, Koutsopoulos H (2012) Bus-holding control strategies: simulation-based evaluation and guidelines for implementation. Transp Res Rec 2274:100–108

Clausen J, Larsen A, Larsen J, Rezanova NJ (2010) Disruption management in the airline industry–concepts, models and methods. Comput Oper Res 37(5):809–821

Daganzo CF (2009) A headway-based approach to eliminate bus bunching: systematic analysis and comparisons. Transp Res Part B Methodol 43(10):913–921

D’Ariano A, Corman F, Pacciarelli D, Pranzo M (2008) Reordering and local rerouting strategies to manage train traffic in real time. Transp Sci 42(4):405–419

De Giovanni L, Heilporn G, Labbé M (2014) A polyhedral study for delay management in public transportation. Proc Soc Behav Sci 108:15–25

Desaulniers G, Desrosiers J, Solomon MM (eds) (2005) Column generation. Springer, Boston, MA, USA

Eberlein XJ, Wilson NHM, Barnhart C, Bernstein D (1998) The real-time deadheading problem in transit operations control. Transp Res B 32(2):77–100

Eberlein XJ, Wilson NHM, Bernstein D (2001) The holding problem with real-time information available. Transp Sci 35(1):1–18

Feillet D, Dejax P, Gendreau M, Gueguen C (2004) An exact algorithm for the elementary shortest path problem with resource constraints: application to some vehicle routing problems. Networks 44:216–229

Glover F, Laguna M (1997) Tabu search. Kluwer Academic Publishers, Norwell, MA, USA

Gualandi S, Malucelli F (2013) Constraint programming-based column generation. Ann Oper Res 204(1):11–32

Hernández D, Muñoz JC, Giesen R, Delgado F (2015) Analysis of real-time control strategies in a corridor with multiple bus services. Transp Res Part B Methodol 78:83–105

Hickman MD (2001) An analytic stochastic model for the transit vehicle holding problem. Transp Sci 35(3):215–237

Huisman D (2007) A column generation approach for the rail crew re-scheduling problem. Eur J Oper Res 180(1):163–173

Huisman D, Wagelmans APM (2006) A solution approach for dynamic vehicle and crew scheduling. Eur J Oper Res 172(2):453–471

Ibarra-Rojas OJ, Delgado F, Giesen R, Muñoz JC (2015) Planning, operation, and control of bus transport systems: a literature review. Transp Res Part B Methodol 77:38–75

Jespersen-Groth J, Potthoff D, Clausen J, Huisman D, Kroon L, Maróti G, Nielsen MN (2009) Disruption management in passenger railway transportation. In: Ahuja RK, Möhring RH, Zaroliagis CD (eds) Robust and Online Large-Scale Optimization. Lecture Notes in Computer Science, vol 5868. Springer, Berlin, Heidelberg

Jin JG, Teo WM, Odoni AR (2015) Optimizing bus bridging services in response to disruptions of urban transit rail networks. Transp Sci 50(3):790–804

Johnson V, Lettovsky L, Nemhauser GL, Pandit R, Querido S (1994) Final report to northwest airlines on the crew recovery problem. Technical report, Technical Report, The Logistic Institute, Georgia Institute of Technology, Atlanta

Kepaptsoglou K, Karlaftis MG (2009) The bus bridging problem in metro operations: conceptual framework, models and algorithms. Public Transp 1(4):275–297

Kiefer A, Kritzinger S, Doerner KF (2016) Disruption management for the viennese public transport provider. Public Transp 8(2):161–183

Lettovskỳ L, Johnson EL, Nemhauser GL (2000) Airline crew recovery. Transp Sci 34(4):337–348

Li J, Mirchandani PB, Borenstein D (2009) A Lagrangian heuristic for the real-time vehicle rescheduling problem. Transp Res E 45(3):419–433

Lougee-Heimer R (2003) The common optimization interface for operations research: promoting open-source software in the operations research community. IBM J Res Dev 47(1):57–66

Malucelli F, Tresoldi E (2018) Delay and disruption management at ATM: technical details. Technical report, Politecnico di Milano-Dipartimento di Elettronica, Informazione e Bioingegneria. http://www.optimization-online.org/DB_HTML/2018/10/6856.html. Accessed 29 Jan 2019

Mároti G (2013) Ultra-short term rescheduling of railway rolling stock. In: Slides at EURO-INFORMS conference

Nissen R, Haase K (2006) Duty-period-based network model for crew rescheduling in european airlines. J Sched 9(3):255–278

Osunaand EE, Newell GE (1972) ontrol strategies for an idealized public transportation system. Transp Sci 6(1):52–72

Potthoff D, Huisman D, Desaulniers G (2010) Column generation with dynamic duty selection for railway crew rescheduling. Transp Sci 44(4):493–505

Raheja AS, Subramaniam V (2002) Reactive recovery of job shop schedules—a review. Int J Adv Manuf Technol 19(10):756–763

Sánchez-Martínez GE, Koutsopoulos HN, Wilson NHM (2016) Real-time holding control for high-frequency transit with dynamics. Transp Res Part B Methodol 83:1–19

Visentini MS, Borenstein D, Li J, Mirchandani PB (2014) Review of real-time vehicle schedule recovery methods in transportation services. J Sched 17(6):541–567

Walker CG, Snowdon JN, Ryan DM (2005) Simultaneous disruption recovery of a train timetable and crew roster in real time. Comput Oper Res 32(8):2077–2094

Wei G, Yu G, Song M (1997) Optimization model and algorithm for crew management during airline irregular operations. J Comb Optim 1(3):305–321

Xuan Y, Argote J, Daganzo CF (2011) Dynamic bus holding strategies for schedule reliability: optimal linear control and performance analysis. Transp Res Part B Methodol 45(10):1831–1845

Yu B, Yang Z, Li S (2012) Real-time partway deadheading strategy based on transit service reliability assessment. Transp Res A 46(8):1265–1279

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.