ORIGINAL PAPER

A review on railway delay management

Eva König[1]

Abstract

Passengers traveling by train may need to change trains on their route. If the focal train of a passenger is late, the passenger might miss his connection and has to decide how to continue his trip. Delay management addresses the question whether the connecting train should wait (or not) for the delayed passengers. If the connecting train waits, delays would get transferred through the network. In literature, several works consider delays and their impact on railways and how to reschedule disturbed plans. We focus on works, aiming to minimize passenger inconvenience as it is done in delay management. In the last two decades, dozens of works considering the delay management problem have emerged, tackling the problem in different ways. In this paper, an overview on the existing literature is given, and a new classification is introduced. We provide a taxonomy scheme for railway problems at an operational level and show how the field of delay management fits to other parts of the planning process. Moreover, limitations of the delay management approaches are discussed and future research opportunities are suggested.

Keywords Delay management · Railways · Operational problems · Review

1 Introduction

The focus of this review is on delay management (DM) for railways. DM, which was introduced by Schöbel (2001) and Suhl et al. (2001), searches for the answer of the so-called wait-depart decision. Should a connecting train wait for a delayed feeder and propagate the delay in the network or depart on time so that transferring passengers will miss their connection? In the last two decades, dozens of works considering this problem have been published. Figure 1 illustrates the growth of new publications since 2001 (the numbers arise from the reviewed literature in this paper). The proposed models for DM range from simple rules of thumb to complete network optimizations. To the best of our knowledge, a survey on these models

✉ Eva König
    e.koenig@mail.uni-mannheim.de

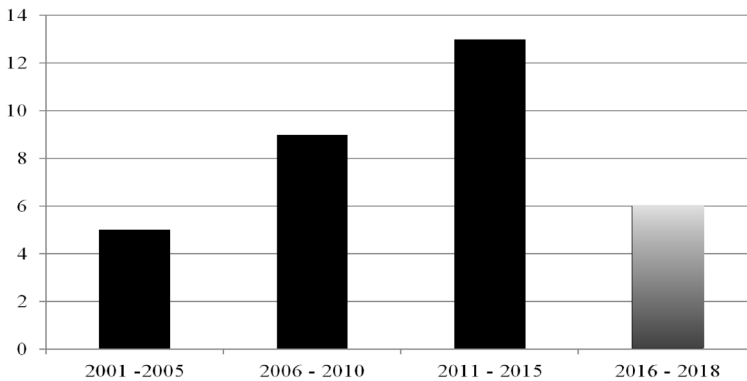1   University of Mannheim, Schloss, 68131 Mannheim, Germany

Fig. 1 Number of new publications per five year interval (with a shortened last interval)

has been neglected so far. Furthermore, a distinction between DM and other related areas is missing. The operational problems are often summarized under the term *real-time management* (Lusby et al. 2011).

DM can be seen as a strong tool to reduce delays for passengers. In several studies, the results under dispatching are compared with results where no dispatching at all was done. Usually mentioned as *never wait* strategy in literature, trains do not wait at all for each other. The *never wait* strategy performs weaker than applied dispatching, as can be seen in, e.g., Kliewer and Suhl (2011), Dollevoet et al. (2012, 2015), showing that there exists a considerable impact on delay reduction.

In practice, such as, e.g., in Germany, statistics on the punctuality only refer to trains. While the punctuality level for long-distance trains amounts to around 80%, this indicator only accounts for non-canceled trains that suffer a delay smaller than 6 min (Die Welt 2018). The delay of passengers is not reported but passengers on a canceled train might be facing transfer problems and probably also delays. The same holds for the tolerance of small delays. They are not part of the statistic but in reality they may cause connection conflicts for passengers (Die Welt 2017). For railways, as service providers, a passenger-friendly dispatching might be worth further investigation.

In 2017 a simulation tool, called PANDA (Rückert et al. 2017, see Sect. 3.2), was applied in a real-world project with Deutsche Bahn (DB), the German railway provider. The tool detects connection conflicts and simulates the consequences on the arrival delays of passengers to support dispatchers in their decision-making process (Deutsche Bahn 2017).

Planning problems for railways are manifold, beginning with long-term problems, such as building new infrastructure, to very short-termed problems, e.g., making dispatching decisions (Lusby et al. 2011). We concentrate on the operational level where railway providers have to cope with daily disturbances. Thereby, the literature often distinguishes between small disturbances leading to delays of several minutes (maybe even hours) and large disruptions that will cause a temporary break-down of the system (see, e.g., Ghaemi et al. 2017). When coping with small delays, dispatchers can

set different goals. One goal is to return as fast as possible to the original schedule and avoid further delay propagation in the network. We call it the train perspective with the objective to minimize train delays. Another goal is to minimize delays for passengers, i.e., the passenger perspective, on which we concentrate.

In this review, we give a comprehensive overview on DM literature but we do not claim completeness. Therefore, we explain the characteristics of DM and distinguish it from other research areas on the short-term level and show how the planning process of Lusby et al. (2011) can be adjusted to the new categories that have arisen. Then we review the literature in the field of DM by developing a taxonomy scheme for operational problems containing five different attributes.

The train perspective is usually the goal in real-time rescheduling (RTR) where train delays are minimized. As we will see in Sect. 2.1, DM and RTR differ in several aspects. There exist numerous reviews on RTR, but most of them contain only a part of the DM literature or neglect it totally. The term RTR has a broad scope, sometimes covering all operational problems after any type of disturbance has occurred. This paper understands RTR as timetable rescheduling in case of small disturbances. In the following, we give a short overview on existing literature reviews in related areas.

- The above-mentioned review of Lusby et al. (2011) tackles all planning problems over all levels in general and gives an overall view on the railway industry. DM or RTR are not mentioned as seperate classes.
- Cacchiani et al. (2014) give a comprehensive overview on problems in railway real-time management. Some works in DM are mentioned, but they are described briefly and not the complete existing literature is considered.
- The same holds for Fang et al. (2015) where all problems in rescheduling are addressed and compared with each other; its focus lies on solution methods.
- In Lusby et al. (2018) a review on robustness in railway planning is presented, but DM is addressed only briefly. The major part is dedicated to robust timetabling.
- Ghaemi et al. (2017) report about large disruptions and how to recover from them with rescheduling models. DM models are not considered.

The paper is structured as follows: in Sect. 2, we first define the term DM by introducing the main characteristics and then distinguish DM from other problems on the operational level. In Sect. 3, a taxonomy scheme for classifying the literature is proposed and applied to the related literature. Finally in Sect. 4, some concluding remarks and ideas for further research are given.

2 Preliminaries

In this section, we first (Sect. 2.1) highlight the key criteria to classify a model as DM model by exploiting a state-of-the-art model. In Sect. 2.2, we place the DM problem among other operational problems and illustrate the influence of DM on other related fields.

2.1 Key criteria in DM

The level of detail of a railway network can be described from a macroscopic or microscopic point of view. In a macroscopic view, the network is sketched in a "rough" way, consisting of stations and tracks connecting stations. However, details such as the number of platforms or division of tracks into block sections are neglected. In microscopic models, these details are modeled additionally, leading to blown up models with several additional constraints.

The majority of the models in DM are macroscopic models while RTR models are often modeled in a microscopic manner as the feasibility from an infrastructure point of view is more important for the infrastructure manager. We will see some exceptions in Sects. 2.2 and 3.2. In Kecman et al. (2013) macroscopic and microscopic models for railways are compared in terms of performance and run time. They find out that macroscopic models perform quite well and find also feasible solutions for the network schedule without taking a detailed view into account.

The macroscopic perspective can be modeled with event-activity networks (EAN). We first introduce a model from Dollevoet et al. (2012) which is built upon an EAN, to explain key criteria with this model at hand. The model from Dollevoet et al. (2012) is an advanced model of the earlier model from Schöbel (2001) that incorporates the opportunity to reroute passengers in case of broken connections. The explanation for EAN and the model from Dollevoet et al. (2012) are concise as we will only give an idea of them. For a more detailed explanation we refer for EAN to Müller-Hannemann and Rückert (2017) and for the model to the original source.

An EAN $\mathcal{N} = (\mathcal{E}, \mathcal{A})$ consists of events (nodes) $e \in \mathcal{E}$ and activities (arcs) $a \in \mathcal{A}$. Events can be categorized as arrival, departure, origin and destination events, with

$$\mathcal{E} = \mathcal{E}_{arr} \cup \mathcal{E}_{dep} \cup \mathcal{E}_{org} \cup \mathcal{E}_{dest}.$$

Arrival and departure events represent the arrival and departure of a train at a station with arrival and departure times. Origin and destination events are related to passenger types $p \in \mathcal{P}$, which are characterized by a unique combination of the origin-destination (OD) pair that passengers want to travel and their desired departure time, $time_p$. Then, for each type $p \in \mathcal{P}$, an origin event $Org(p) \in \mathcal{E}_{org}$ and a destination event $Dest(p) \in \mathcal{E}_{dest}$ is introduced as the start and end point of its path through the network together with $time_p$. Furthermore, we assume to know the size w_p of each passenger type $p \in \mathcal{P}$.

Arcs result from activities in the directed graph. We distinguish between driving, waiting and changing activities meaning that a train drives between consecutive stations, waits at a station and passengers can change between trains, respectively. As the EAN $\mathcal{N}$ is a directed graph, we can identify all ingoing arcs, denoted by $\mathcal{I}(e)$, and all outgoing arcs, denoted by $\mathcal{O}(e)$, of an event $e \in \mathcal{E}$. Additionally, starting and finishing activities for passengers ($a \in \mathcal{A}_{start}(p)$ and $a \in \mathcal{A}_{fin}(p) \forall p \in \mathcal{P}$) are necessary in order to start or finish a trip. For starting activities, $Org(p)$ is connected to all departure events $e \in \mathcal{E}_{dep}$ and for finishing activities, all arrival events $e \in \mathcal{E}_{arr}$ are connected to $Dest(p)$. The set of activities is then as follows:

$$\mathcal{A} = \mathcal{A}_{drive} \cup \mathcal{A}_{wait} \cup \mathcal{A}_{change} \cup \mathcal{O}(Org(p)) \cup \mathcal{I}(Dest(p)).$$

The minimum time required to perform an activity $a \in \mathcal{A}_{drive} \cup \mathcal{A}_{wait} \cup \mathcal{A}_{change}$ is declared as δ_a. Parameters for delays are denoted by Δ_e as the delay at an event $e \in \mathcal{E}_{arr} \cup \mathcal{E}_{dep}$ and Δ_a the delay during an activity $a \in \mathcal{A}_{drive} \cup \mathcal{A}_{wait}$.

To compute delays, first the starting point of a train schedule has to be included by the parameter τ_e for the planned arrival or departure times of an event $e \in \mathcal{E}_{arr} \cup \mathcal{E}_{dep}$. The earliest possible arrival time for a passenger of type $p \in \mathcal{P}$ without delays, denoted by $\bar{t}_p$, can be computed in a preprocessing step with a shortest path algorithm (see König and Schön 2020 for an explicit formulation). The preprocessing model corresponds to the DM problem where all delays are set to zero, i.e., the preprocessing model only consists of a modified objective function (Eq. (1)) and the shortest path problem (Eqs. (6)–(8) and (11)) and possesses no delay constraints.

Actual arrival and departure times are determined by scheduling decision variables, i.e., x_e for the (potentially rescheduled) time of an event $e \in \mathcal{E}_{arr} \cup \mathcal{E}_{dep}$. The rescheduled times represent a timetable that is temporarily feasible for delayed trains (a disposition timetable). Passenger delays are measured when they exit a train at their final station. For this purpose, another decision variable $t_p \in \mathbb{N}$ is introduced that denotes the arrival time of passenger type $p \in \mathcal{P}$ at the final destination.

In DM, wait-depart decisions have to be made; therefore, a binary decision variable z_a for the changing activities $a \in \mathcal{A}_{change}$ is introduced:

$$z_a = \begin{cases} 1 & \text{if connection } a \text{ is maintained,} \\ 0 & \text{otherwise.} \end{cases}$$

The routing part of the model needs an additional binary decision variable y_{ap} representing whether activity $a \in \mathcal{A}$ is included in a path of passenger type $p \in \mathcal{P}$. It is defined as follows:

$$y_{ap} = \begin{cases} 1 & \text{if activity } a \text{ is assigned to passengers of type } p, \\ 0 & \text{otherwise.} \end{cases}$$

The complete model looks as follows (see e.g., Dollevoet et al. 2012; Dollevoet and Huisman 2014; König and Schön 2020):

$$min \sum_{p \in \mathcal{P}} w_p \left(t_p - \bar{t}_p \right) \tag{1}$$

s.t.

$$x_e \geq \tau_e + \Delta_e \quad \forall e \in \mathcal{E}_{arr} \cup \mathcal{E}_{dep}, \tag{2}$$

$$x_e \geq x_{e'} + \delta_a + \Delta_a \quad \forall a = (e', e) \in \mathcal{A}_{drive} \cup \mathcal{A}_{wait}, \tag{3}$$

$$x_e \geq x_{e'} + \delta_a - M_1 \left(1 - z_a \right) \quad \forall a = (e', e) \in \mathcal{A}_{change}, \tag{4}$$

$$y_{ap} \leq z_a \quad \forall\, p \in \mathcal{P}, a \in \mathcal{A}_{change}, \tag{5}$$

$$\sum_{a \in \mathcal{O}(Org(p))} y_{ap} = 1 \quad \forall\, p \in \mathcal{P}, \tag{6}$$

$$\sum_{a \in \mathcal{O}(e) \cap \mathcal{A}(p)} y_{ap} = \sum_{a \in \mathcal{I}(e) \cap \mathcal{A}(p)} y_{ap} \quad \forall\, p \in \mathcal{P}, e \in \mathcal{E}_{arr} \cup \mathcal{E}_{dep}, \tag{7}$$

$$\sum_{a \in \mathcal{I}(Dest(p))} y_{ap} = 1 \quad \forall\, p \in \mathcal{P}, \tag{8}$$

$$t_p \geq x_e - M_2\left(1 - y_{ap}\right) \quad \forall\, p \in \mathcal{P}, e = Dest(p) \in \mathcal{E}_{dest}, \ a \in \mathcal{A}_{fin}(p), \tag{9}$$

$$z_a \in \{0, 1\} \quad \forall\, a \in \mathcal{A}_{change}, \tag{10}$$

$$y_{ap} \in \{0, 1\} \quad \forall\, p \in \mathcal{P}, a \in \mathcal{A}, \tag{11}$$

$$x_e \in \mathbb{N} \quad \forall\, e \in \mathcal{E}_{arr} \cup \mathcal{E}_{dep}, \tag{12}$$

$$t_p \in \mathbb{N} \quad \forall\, p \in \mathcal{P}. \tag{13}$$

In DM the focus is on the passenger, i.e., in "classical" DM models, the objective function (Eq. (1)) minimizes the passenger-weighted delay. Similar formulations are *minimizing passenger inconvenience* or *minimizing the time spent by a passenger in the railway system*. For further objectives in DM we refer to Dollevoet et al. (2018). Another objective in the railway industry, taking the train perspective into account, is to minimize train delays, i.e., minimizing deviations from a given train schedule. This is usually the goal in railway RTR. Both objectives differ as decisions underlying a passenger perspective are not necessarily easy to operate and might cause further delays for trains. Decisions related to train delays can cause inconvenience for passengers. In Sect. 3.2, we present some *hybrid* models that concentrate on passenger and train delay simultaneously.

The central question in DM for railways is if a connecting train should wait for a delayed feeder train or depart on time. In constraints (4), the decision, if a transfer is possible, is determined with the binary variable z_a (i.e., $z_a = 1, a \in \mathcal{A}_{change}$). Passengers are only allowed to change trains if sufficient time for transferring between the arrival and departure of consecutive trains is available, with M_1 chosen large enough. While the question seems easy, the answer is not trivial. If the connecting train departs without waiting, transferring passengers miss their connection. Depending on the schedule, they might face severe delays and might have to wait a long time before they can continue their journey. When railway providers operate a cyclic timetable, a train might drive with a periodicity of one or two hours. It is

even worse if this is the last train of the evening and passengers risk stranding somewhere. If the connecting train waits, it is also delayed and on the next station other passengers are concerned with maintaining their connection. The delay can spread through the network and repercussions will get visible in other parts of the network. In König and Schön (2020), the emergence of new connections due to delays is also possible, i.e., passengers can jump on late trains for which in an undelayed case no connection was planned.

Time constraints determining new arrival and departure times including possible source delays to yield the disposition timetable are modeled in constraints (2)–(3). An event cannot be scheduled earlier than it was planned in the original timetable (Eq. (2)). The same holds for activities in train schedules (Eq. (3)). Please note, to link the objective (Eq. (1)) with the rescheduled arrival time at a passenger's destination, auxiliary constraints that transfer the arrival time of passenger streams to the t_p variables are necessary (Eq. (9)). Another way is to modify the objective by including the connection decision via the z_a variables directly therein, see e.g., Schöbel (2007), Schachtebeck and Schöbel (2010).

A specialty of this model is to assume passenger rerouting, i.e., in case of a missed or broken connection, passengers can change their route and will eventually reach their destination via a different route. In earlier models (and some more recent models as, e.g., Dollevoet et al. 2015) this option for the passengers is not incorporated. If passengers miss a connection, they have to wait a full cycle time for the next train on the line. The rerouting is included via a shortest path problem in constraints (6)–(8). An expanded set of decision variables is therefore necessary: the y_{ap} variables representing the passenger streams. To ensure that passenger changing activities are feasible only if the corresponding train connection is maintained, constraints (5) are further necessary. Finally, in constraints (10)–(13) the requirements for the variable sets are defined.

Other DM models incorporate different kinds of capacities, leading to additional constraints. So far, capacities of tracks have been taken into account in Schöbel (2009) and Schachtebeck and Schöbel (2010). In Dollevoet et al. (2015) the capacity of stations is additionally included to the capacity of tracks and in König and Schön (2020) train capacity constraints are considered. In some publications these constraints are modeled with "Big M" constraints, so the model remains a mixed integer program (MIP) (e.g., Schachtebeck and Schöbel 2010). In König and Schön (2020) the resulting model is a mixed-integer nonlinear program (MINLP) due to modeling passenger streams with continuous variables. For a further description on these models, we refer to Sect. 3.2.

2.2 Placement on the operational level in the railway planning process

For the different levels in the railway planning process a scheme is provided by Lusby et al. (2011) (see Fig. 2). They divided the problems into three different levels, strategic for long-term planning (several years), tactical for mid-term (one year) and operational for short-term (one day) planning. The different problems are interrelated and plans have to be coordinated. For an explanation of the strategic

problems we refer to Lusby et al. (2011) since these are out of our scope. The main problem on the tactical level is the timetable generation (also mentioned as *timetabling*). To compute a timetable, arrival and departure times of trains for all stations on their respective line are determined. Track allocation is often part of timetabling as the timetable has to fulfill operational requirements to be feasible, such as capacity restrictions on tracks. For a comprehensive overview on timetabling, we refer to Cacchiani and Toth (2012). The allocation of rolling stock, such as trains (Cacchiani et al. 2012) and schedules for operating staff on trains, i.e., the crew (Caprara et al. 1998), depend also on the timetable. For the operational level they call these problems *real-time management* but this is only a rough classification. In the following, we demonstrate how the operational level can be structured in different problems.

In the last years, several new problems, as e.g., disruption management and DM, have arisen; some of them combining different levels and problems. These emerging problems need to be placed in the planning process (such as DM). In the literature, different understandings of the terms exist and some problems seem to depend on each other, e.g., disruption management and crew rescheduling. We suggest to sharply distinguish between the problems. In the following, we will revise the part of the operational level and provide a finer granularity of detail. Figure 3 presents an overview on different problems (written in the bubbles) on the operational planning level and related problems on the tactical level.

Arrows depict influences between different problems. Please note, in Fig. 3 only arrows concerning operational problems are included. Between the problems on the tactical level (and strategic level which is not shown here) exist also arrows, but they are beyond the scope of this paper. The size of the bubbles is representing the amount of literature, e.g., as there exists more literature for RTR than for all other operational problems, the bubble for RTR is the biggest on that level. In the following, we will briefly explain neighboring problems of DM on the operational level
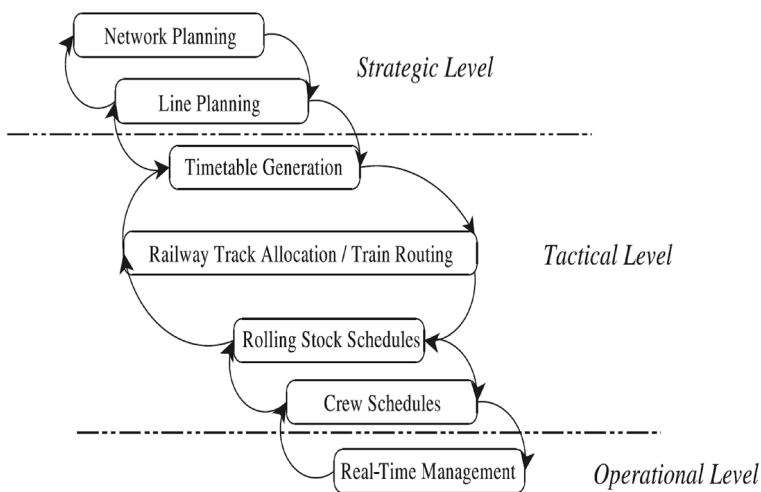

Fig. 2 Planning process in railways (Lusby et al. 2011)
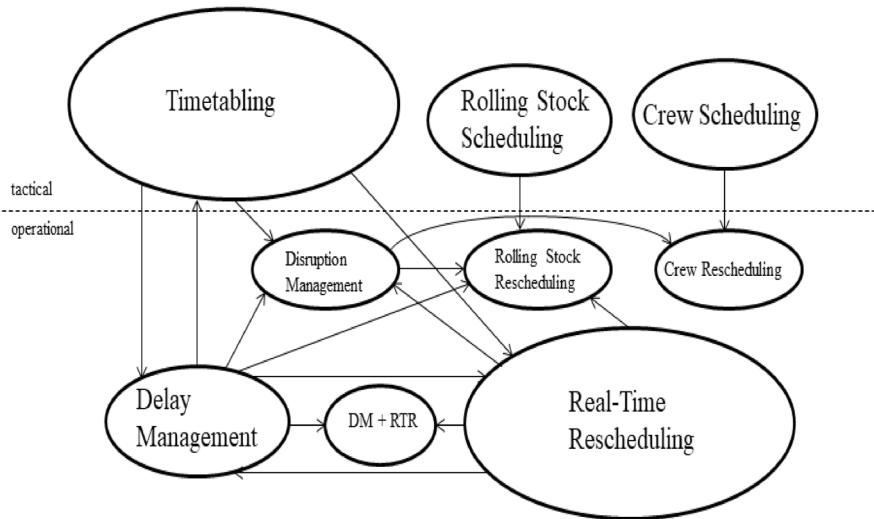
Fig. 3 Problems on the operational planning level

and interrelations. We further give examples for related literature on the neighboring problems and the arrows. DM itself has been explained in Sect. 2.1. Everything lying inside DM and DM + RTR will be discussed extensively in Sect. 3.2.

2.2.1 DM and timetabling

DM models are highly affected by timetabling as the majority of the models aims at developing disposition timetables. We will see several models computing disposition timetables in the literature in Sect. 3.2.

The other direction, the influence of DM on timetabling, leads to so-called robust timetables which try to cover some delay cases to make the timetable robust against disturbances. DM is integrated in the computation of delay-resistant timetables in Liebchen et al. (2010). First, a timetable is computed and then evaluated in delay scenarios by solving it with DM models. The resulting disposition timetables are used to revise the original timetables. In Goerigk et al. (2014), the timetable is based on an EAN and takes a DM model into account. Cicerone et al. (2012) develop a multi-stage recovery model to integrate robustness into timetables with the aid of disposition timetables.

2.2.2 DM, disruption management and rolling stock rescheduling

As mentioned above (see Sect. 2.1), operational problems can be differentiated between minor disturbances and major disruptions. The goal of disruption management is to develop strategies for handling large-scale disruptions with a long (possibly unknown) recovery time. The remaining operational problems primarily

account for minor disturbances. Please note, we differentiate here between disruption management and rolling stock rescheduling. Usually, these problems are merged together. For further information on disruption management we refer to Ghaemi et al. (2017).

In some approaches for disruption management, several characteristics from the DM literature are used. Louwerse and Huisman (2014) aim to maximize the service level offered to passengers. In case of a disruption, a disposition timetable is computed based on an EAN, similar to common practices in DM. Further enhancements for this model are proposed in Veelenturf et al. (2016a) by modeling all phases from the beginning of the disturbance to the reconstitution of the original timetable. Binder et al. (2017) develop a multi-objective integer program (IP) where one of the objectives is passenger satisfaction. Therefore, disposition timetables considering a macroscopic view are constructed.

A different approach is developed in Schmidt et al. (2017). Alternative route choices of passengers in case of complete blockages are compared. Thereby, a decision has to be made if the passenger waits for the recovery of the system or takes another train (possibly leading to a detour). Decisions are made under uncertainty, including probability distributions for some scenarios, and dominance relations between the strategies are revealed.

The problem of rolling stock rescheduling aims for an adjusted allocation of the rolling stock after a disturbance has occurred. For rolling stock rescheduling there exist influences from DM, such as, e.g., the passenger perspective.

Kroon et al. (2014) include the passenger perspective by minimizing passenger delays when rolling stock has to be rescheduled after large disruptions. Therefore, passenger flows are simulated. In a follow-up paper of Veelenturf et al. (2017), passenger behavior and improvements for passenger service are evaluated. The model is formulated with a macroscopic view and adapts stopping patterns of trains.

2.2.3 DM and RTR

RTR determines a feasible timetable after a disturbance occurred and the actual timetable can no longer be operated. For related literature, see the reviews presented in Sect. 1. In recent years, the conjunction of DM and RTR has become stronger. There exist even combined works where both areas are merged that closely that we included a seperate bubble for these models, called DM + RTR.

One example for DM models influenced by RTR is Schöbel (2009). In Schöbel (2009) the capacity of tracks is added to the classical DM model. It was the first try to include a microscopic view, too. We will explain this model further in Sect. 3.2, see the class of [*pmadh*].

For the other influence direction, the influence of DM on RTR, there exist also some approaches in the literature. The model from Caimi et al. (2012), for example, belongs to the RTR area and proposes a predictive traffic management support system. The problem formulation is proposed as a rescheduling model with a detailed network description. But in the objective function, customer satisfaction is maximized by weighting the delay and maintained connections, which is usually done in DM literature.

Examples for mixed approaches of DM + RTR are Corman et al. (2012) and Dollevoet et al. (2014). Corman et al. (2012) integrate both the passenger and the train perspective simultaneously in a bi-objective problem with the aim of finding Pareto optimal solutions. The constraints are modeled in a microscopic perspective; therefore, we explain the model in the class of [*pmidh*]. Dollevoet et al. (2014) instead solve a macroscopic model from DM and a microscopic model from RTR iteratively. As the main objective is to minimize passenger delay, we explain this work in the class of [*pmade*] (see for both works Sect. 3.2).

2.2.4 Additional literature—slightly related to DM

Crew rescheduling determines feasible crew schedules after a disturbance has occurred. It is similar to RTR but for staff instead of trains (see Veelenturf et al. 2016b, Verhaegh et al. 2017). There exist also works in crew rescheduling influenced by disruption management and influences from timetabling on RTR. In Fig. 3 we see further arrows between RTR, disruption management, and rolling stock rescheduling. These arrows will not be described further as they are not influenced by DM. For related literature, we refer to the above-mentioned reviews (see Sect. 1).

Malucelli and Tresoldi (2019) propose a simulation-based approach for disruption management combining crew and rolling stock rescheduling that can be used for small disturbances as well. The regularity of service relies not on a timetable but on a frequency based service, i.e., the headway of two vehicles is determined instead of deviations from a given timetable. Crew and rolling stock rescheduling are performed if the regularity of the service cannot be guaranteed any longer. The performance of the approach is measured with different regularity functions reflecting the reliability of the service.

In Goerigk et al. (2013), the concepts of DM, timetabling and line planning are combined. DM is included by generating some delay instances and evaluating robustness for the timetable and the planned lines. On the other hand, it is analyzed if the line concept and the timetable facilitate the emergence of delays. This arrow is not included in Fig. 3 since the area from the strategical level is not part of the figure.

3 DM literature

In Sect. 3 the literature on DM is reviewed and classified. In Sect. 3.1 we will explain a taxonomy scheme for different attributes of models and methods on the operational level. Section 3.2 contains the literature review structured with the proposed taxonomy scheme. We further give an overview of applications in the real world.

3.1 Taxonomy

The existing literature often distinguishes between a macroscopic and a microscopic view to differentiate between DM and rescheduling. Sometimes it is mentioned that DM models get their information for the decision-making process in an *online* or *offline* manner, see, e.g., Schmidt (2013), Rückert et al. (2017). However, not all information statuses are covered with that, i.e., stochastic models are not considered. Some attributes were already mentioned above, in Sect. 2.1, when describing key criteria for DM.

The focus of the objective can either be on the passenger—we mark these works with attribute level [*p*]—or on the train, marked with [*t*]. In DM the focus is usually on the passenger; therefore, the majority of the works will be classified as [*p*]. Nevertheless, there exist some mixed or *hybrid* models where passenger and train focus are combined (as mentioned above in Sect. 2.2).

Delays can arise due to different causes and lead to disturbances of different length. As explained above (Sect. 2.1), different approaches are necessary for either coping with minor disturbances [*m*] or large disruptions [*l*]. DM usually considers delays in a smaller time window; the discussed works will all assume minor disturbances [*m*].

The perspective on the railway network can be macroscopic [*a*] or microscopic [*i*] depending on the level of detail. For DM it is common to model the network in a macroscopic view, as shown in Sect. 2.1. Nevertheless, we will see some exceptions, taking a microscopic perspective into account in the following Sect. 3.2.

Another category of attributes is the available input to solve the problem. We will differentiate in three attribute levels for the information at hand, according to Jaillet and Wagner (2010). If all delays are assumed to be known so that the decision can be made under full information, we call the model deterministic [*d*]; in literature often described as *offline models*. For less available input, where no exact delay is given but delays follow a known distribution function, we will describe the model as stochastic [*s*]. Moreover, the input data may be incomplete [*n*], i.e., no information on the future is available at the point in time the decision has to be made. Future information is obtained dynamically as time goes by; therefore, this type is often called *online models* in literature.

Finally, we differentiate between exact solution methods [*e*] and heuristics [*h*]. With the macroscopic view on a railway network, the modeling is done in simplified terms with fewer constraints and binaries than RTR. The older, more basic models, can often be solved exactly in a deterministic setting, as we will see in Sect. 3.2. For more developed models, heuristic methods are often applied additionally to shorten computation times.

A certain type of heuristics are so-called dispatching rules, where wait-depart decisions are made through rule-based strategies. With these dispatching rules, the decision can be made quickly and easily as only partial information is necessary. A rule called regular waiting time (RWT) determines the amount of time a train is allowed to wait for a delayed feeder train depending on the train type. Other dispatching rules prescribe that trains do not wait at all (NW) and that all trains wait for all (delayed) feeder trains (AW). In the past, some of them were

used for practice in Germany, e.g., RWT: a long-distance train is allowed to wait up to 3 min for a delayed long-distance train (Stelzer 2016). Dispatching rules, especially NW, RWT and AW, are, therefore, often used in numerical studies for comparison with optimization models that are harder to solve, taking advantage of the fast and easy computation; see, e.g., Dollevoet et al. (2012), Dollevoet and Huisman (2014), Dollevoet et al. (2014), Bauer and Schöbel (2014), Schön and König (2018).

To go back to models under incomplete information, algorithms that solve these models are called online algorithms. These algorithms have to take the dispatching decision only with past and current information. They are often further evaluated with different performance measures as, e.g. competitive analysis (see, e.g., Lan et al. 2008, Agrawal et al. 2014). Determining the quality of an online algorithm is done by computing the ratio of an optimal offline algorithm (the omniscient adversary) and the analyzed online algorithm, the competitive ratio. The ratio indicates the quality of an online algorithm, e.g., an algorithm is 2-competitive if the online algorithm finds a solution that is never twice as bad as the optimal solution. Some dispatching rules need no input data at all to determine a solution, e.g., an AW implies that all trains always wait, independently from considering any information. Some works in the literature run simulations of the network, trying to include as much real-world data as possible to simulate different processes. Solutions are usually derived by applying some of the above-mentioned dispatching rules as the goal is rather on the comprehension of the processes than yielding an exact solution.

Moreover, some works pursue the determination of the computational complexity of different DM problems. The computational complexity gives a hint for the computational effort to derive a solution for the considered problem by analyzing its worst-case time requirements as a function of the size of its input. In the works reviewed, polynomial (P), non-deterministic polynomial (NP) and PSPACE problems, i.e., the set of all problems that can be solved on a deterministic Turing machine using space restricted by a polynomial in the input size, are evaluated. For a deeper explanation, we refer to Papadimitriou (2003). All categories and attributes are summarized in Table 1.

Table 1 Attributes and their corresponding attribute levels

Attribute	Attribute level
Focus	passenger [p], train [t]
Delay cause	minor disturbance [m], large disruption [l]
Perspective	macro [a], micro [i]
Input	deterministic [d], stochastic [s], incomplete [n]
Solution	exact [e], heuristic [h]

3.2 Literature classification

In the following, we will go through the different classes in DM literature. These classes are created with the attribute levels from Table 1 and we group articles together according to their attribute-level combination. Some works contain more than one attribute level in the respective attribute; these works can be considered as *hybrid* models. Other authors have developed more than one model in their works differing in their attribute level. In this case, we classify the work, to the best of our knowledge, according to the attribute level that dominates. As all the reviewed works have minor disturbances and a passenger focus (a few works focus additionally on trains) in common, we differentiate the classes according to the attribute's perspective and input. Different solution methods will be explained therein.

3.2.1 Macroscopic deterministic models [**pmade**] + [**pmadh**]

The first class, representing the largest class, considers works with a macroscopic perspective and deterministic input. Solution methods vary; we will see standard optimization models in DM, as introduced in Sect. 2.1 that can be solved exactly as well as different heuristic methods.

The model of Schöbel (2001) is formulated as a MIP based on an EAN with the objective of minimizing passenger delay. Solutions can be obtained by using standard MIP solvers. All consecutive standard optimization models are built upon an EAN. This model is further enhanced in several works considering different aspects DM is confronted with.

In Schöbel (2007) the DM model, which is based on an EAN (Schöbel 2001), is compared to a path-based formulation and shown to be equivalent. A generalized IP formulation is given by computing exact passenger weights for the objective function. In a numerical study, the generalized formulation is analyzed and the size of the EAN is reduced by deleting redundant events. The proposed DM problem can be solved with standard MIP solvers. Additionally, an exact solution algorithm is proposed that splits the problem into independent subproblems and solves them individually.

Heilporn et al. (2008) derive a variable reduction for the DM problem in Schöbel (2001). They model the DM problem in two versions by neglecting departure events and therefore reducing the number of decision variables. The equivalence of the new modeling variants and the model from Schöbel (2001) is shown. The first model can be solved with standard solvers while for the second one a constraint generation approach is proposed. The two modeling variants differ in their performance depending on the size of the network.

In Gatto et al. (2004) the complexity of the DM problem with a single delayed train is evaluated. The authors analyze the number of passenger transfers and derive a minimum cut reduction. They further examine the structure of the network by applying a dynamic program. Additionally, the NP-completeness of a problem with a single delayed train and allowance for passengers to change their route is shown.

The work of Gatto et al. (2004) is complemented in Gatto et al. (2005) where the difference for DM problems that are polynomially solvable and the ones that are NP-complete are exposed. They find out that the complexity depends on factors such as the network topology and slack times in the schedule.

Ginkel and Schöbel (2007) formulate a bicriteria DM problem, that minimizes the missed connections for passengers and train delays simultaneously. In the objective function the focus is on passengers [*p*] and trains [*t*] as well. Therefore, we can classify the model as [$\frac{p}{t}$*made*]. The aim is to find Pareto solutions for the multicriteria model. The model has similarities to project planning and can be solved exactly by adapting a project planning method. Efficient solutions can be found quickly (< 1 min), so the authors suggest that the model could also be used in settings under incomplete information. Further, a proof for the NP-completeness of the bicriteria DM problem is given.

The following works further enhance the DM problem by adding different restrictions making the optimization models more realistic. First steps towards considering the capacities of tracks in a DM model have been proposed in Schöbel (2009). The infrastructure constraints are modeled in a microscopic view and the model can therefore, be seen as a special case. The model is a *hybrid*, consisting of the attribute levels [*a*] and [*i*] as well. Therefore, the model in Schöbel (2009) can be described as [$pm\frac{a}{i}dh$]. Two heuristics are proposed; one that fixes the order of trains and then solves DM with additional precedence constraints. Secondly, a heuristic that solves the DM problem without track capacities and then resolves the problem with headway constraints is developed.

In Schachtebeck and Schöbel (2010), taking the capacity of tracks into account and considering the order of trains and their headways is evolved. Priority constraints are added to the IP of the DM problem. The problem can be solved optimally. However, the additional constraints lead to longer computation times. Therefore, a preprocessing step was included to reduce the problem size similar to Schöbel (2007). After the preprocessing, the model performs significantly faster. Additional heuristics are also proposed. These heuristics decompose the problem by solving a DM problem with fixed priorities of trains in one step and the order of trains on a line in another step. The first two heuristics solve the subproblems in varying sequences. For the last two heuristics, the wait-depart decisions are fixed in a first step, too. The heuristics show a significantly shorter computation time, but the relative error of the solution grows with the size of the network.

The capacity of stations is taken into account in Dollevoet et al. (2015). The DM model with track capacity (Schachtebeck and Schöbel 2010) is supplemented with constraints that schedule the platform track assignment in stations. The model can be solved exactly, but for larger instances an iterative heuristic is developed. Firstly, the platform track assignment is fixed and based upon this wait-depart decisions and priorities of trains are determined. Afterwards, the platform track assignment is rescheduled for each station individually. This procedure can be repeated until no further improvement is possible. The platform track assignment alone can be solved in polynomial time. The passenger delays could be reduced, but the program reschedules many trains which might lead to further passenger inconvenience.

Dollevoet et al. (2012) consider the aspect of passenger rerouting, for the model formulation see Sect. 2.1. In other DM models, it is assumed that passengers, who miss a connection, have to wait a complete cycle time for the next train. In Dollevoet et al. (2012), a shortest path problem is included to look for alternative routes for passengers. The model can be solved exactly, but again the additional constraints extend solution times. The DM problem with rerouting is compared to the *classical* DM problem and a never-wait policy; the problem with rerouting outperforms the other.

In a follow-up paper from Dollevoet and Huisman (2014) heuristics for larger instances of DM with passenger rerouting (as described in Dollevoet et al. 2012) are evaluated. The penalty for a missed train connection is accommodated by a model that is used in Schöbel (2007) to reveal the assumption that passengers wait a complete cycle time. Additionally, an iterative heuristic is proposed that solves the model from Schöbel (2007) first and then computes new passenger routes. The proposed heuristics are tested against dispatching rules and the exact solution. Among the heuristics, the iterative heuristic performed best with a quite small gap to the optimal solution and in shorter computation time.

First attempts to determine the complexity of DM with passenger rerouting have been proposed in Dollevoet et al. (2012). Schmidt (2013) proves DM with passenger rerouting to be NP-hard. For one OD pair the problem is strongly NP-hard. A polynomial-time algorithm is developed that is able to find an optimal solution in certain cases. For general DM problems with passenger rerouting, in the sense that there is more than one OD pair, the calculation of lower bounds is proposed.

In König and Schön (2020), the capacity of trains is taken into account and spill effects are evaluated. The model further considers passenger rerouting as it is done in Dollevoet et al. (2012). Passenger streams are broken down into fractions leading to a MINLP. Three different linearizations (exact and approximate) are proposed. The approximation is based on McCormick envelopes that relax the problem. The exact linearizations are formulated first with SOS1 constraints (special ordered sets of type 1) and second with a logarithmic representation of integer variables. In a numerical study the three new proposed approaches are compared to the DM model from the literature. A considerable spill effect is measured as the DM model with train capacities outperforms the reference model neglecting train capacities in every scenario. For larger test instances, the exact formulations had problems to deliver results in a reasonable time while the McCormick approximation was able to obtain results.

A special case in this class is the model of Dollevoet et al. (2014), as it is a combination of a macroscopic DM model ([*pmade*]) and a microscopic rescheduling model ([*tmide*]). Both models are solved iteratively by first determining which connections to drop and which to maintain. For the achieved disposition timetable the microscopic model determines the feasibility for operating. The DM model is based on the model from Dollevoet et al. (2012) by only taking the scheduling constraints into account to derive the disposition timetable. Since the main objective of the approach is to minimize passenger delay, we assign this model to DM.

A different approach for modeling DM models is presented in Suhl et al. (2001) and Kliewer and Suhl (2011) (for the description of Kliewer and Suhl (2011) see

the class of [*pmanh*]). These models stem not from EAN. In Suhl et al. (2001) a model for scheduling arrival and departure times with a nonlinear objective function is developed. In the objective function different weights are assigned for the waiting place of the passengers (in the train or on the platform) and the additional waiting time to describe passenger inconvenience. The model is solved using SOS2 variables (special ordered sets of type 2) and the resulting MIP can be solved also for large instances. The model is evaluated for a few scenarios by deriving a solution solely on the basis of the optimization compared to an optimization considering waiting time rules for trains (see [*pmanh*]). It turns out that the optimization without waiting restrictions for trains performs better. In Suhl et al. (2001), two further approaches are evaluated independently which belong to different classes. As the macroscopic and deterministic attribute levels dominate, we assigned the overall work to this class. Suhl et al. (2001) introduce also several dispatching rules, e.g., AW, NW, RWT (see Sect. 3.1), as they have been used for the German railways. These heuristics belong to the class of [*pmanh*]. These dispatching rules are evaluated with different waiting times for RWT and it turns out that AW performs badly, especially for larger delays. RWT and NW perform quite similar, whereas NW is slightly better. RWT shows the best performance for waiting times of 2 or 3 min. Moreover, Suhl et al. (2001) propose a multi-agent system that is able to behave autonomously, classified as [*pmidh*]. The microscopic view is appropriate as agents represent microscopic items. The system consists of a passenger generator, a topology manager for the infrastructure of the network and an assistant for dispatchers. Everything is controlled by a simulation server. The German network served as test basis by applying NW and AW in the simulation. With the aid of the simulation, passenger information for dispatchers can be gained.

These software agents are further developed in Biederbick and Suhl (2007). The complete German network is implemented in the simulator. Passengers can be directed individually with a "passenger router". In a numerical study several dispatching strategies are tested with and without passenger rerouting. Passenger-related dispatching strategies show thereby a good performance.

Berger et al. (2011a) introduce a dynamic decision support system that includes updating delay information and respective new arrival and departure times as well as a simulation of passenger flows. The objective of minimizing passenger delay is reformulated in three different ways. The underlying EAN takes the passengers as multi-commodity flow into account. Solutions are obtained by an algorithm that uses the RWT of the German railways and updates information on passengers and timetables repeatedly. The disposition tool is able to demonstrate the effects of the dispatching decisions in the network in reasonable time.

A combination of optimization and simulation is proposed in Kanai et al. (2011) that minimizes passenger disutility. Several congestion formulas leading to different disutility functions in the objective are evaluated. Train traffic and passenger flows are simulated simultaneously. The optimization part determines if connections should be maintained by applying a tabu search heuristic. In a numerical study the different objectives and dispatching decisions are varied and it turns out that the interaction of simulation and optimization leads to decreased passenger disutility.

Rückert et al. (2017) introduce PANDA, a web-based decision support tool for dispatchers. PANDA reflects real-time information on passenger flows and evaluates the effects of wait-depart decisions in the network. The model formulation is proposed on the basis of an EAN that determines how stable a connection is and decides on the amount of affected passengers. A case study with the data of the German network shows that passengers benefit from PANDA's recommendation. In a second case study, the authors analyze the impact of an early rerouting which decreases the delay of passengers as well. Currently, PANDA is used in a study from DB on the German network as mentioned in Sect. 1.

Enhancements for PANDA are studied in Lemnian et al. (2016) by conducting a sensitivity analysis and expanding the scope of wait-depart decisions. In the sensitivity analysis the amount of passengers that is needed to change a dispatching decision of PANDA is analyzed. An IP formulation is given and experiments are performed revealing that decisions are either very stable or very unstable. The impact of joint subsequent waiting decisions is further evaluated in a conflict tree to take the propagation of decisions through the network into account. Experiments on this show no significant impact.

3.2.2 Macroscopic incomplete information models [**pmanh**]

The following class considers also a macroscopic view but possesses incomplete information. In theses works different heuristic methods are developed.

Kliewer and Suhl (2011) use a deterministic model, a simplified model from Suhl et al. (2001), in order to obtain a benchmark by computing ex-post optimal solutions and to derive a re-optimization policy for a large numerical study on dispatching rules. As the main investigation is on the rule-based methods, we classify the work of Kliewer and Suhl (2011) as [*pmanh*]. Kliewer and Suhl (2011) propose further dispatching rules on the basis of transferring passengers to different trains. In the numerical study these passenger-dependent strategies and dispatching strategies without considering information of passengers, (as mentioned above, e.g., RWT, AW, NW) are compared to the optimization with full information and a dynamic re-optimization policy. The passenger-related strategies outperform the other dispatching rules and even the re-optimization policy. The advantage of these rules is that they can be applied easily and much faster with less information (some need none at all).

Bauer and Schöbel (2014) developed dynamic heuristics by computing a solution for deterministic models repeatedly when new information is available. The model from Schachtebeck and Schöbel (2010) is modified and solved with and without track capacities. To yield a robust algorithm, a learning strategy that is able to cope with incomplete information is proposed. The solution is obtained by iteratively performing a re-optimization. In a numerical study the heuristics outperform simple dispatching rules and are able to compete with the solutions derived in deterministic settings.

Some of the works within this class propose online algorithms for different scenarios and determine the competitive ratio of the algorithms, a performance measure as introduced in Sect. 3.1. Anderegg et al. (2002) are the first who propose a bound

for the competitive ratio for the solution of a simplified DM problem with unknown delay. The central decision that has to be made, is how long a vehicle should wait at a station for another delayed vehicle with focus on minimizing passenger waiting time. An extended version of the paper can be found in Anderegg et al. (2009).

In Gatto et al. (2007) the DM problem on a single train line under incomplete information is compared to the Ski-Rental problem, a well-known problem from the literature. The authors prove that this DM problem can be solved with algorithms, belonging to the class of 2-competitive online algorithms. The exact value of the competitive ratio is determined to be the golden ratio (a value of ≈ 1.618).

In a follow-up paper, Gatto et al. (2008) consider the DM problem from Gatto et al. (2007) for a weakened adversary (usually the opponent is assumed to be omniscient). Further special cases are evaluated to close some gaps on the bounds for the competitive ratio.

Krumke et al. (2011) model the DM problem from Gatto et al. (2007) as a two-person zero-sum game and achieve an improved lower bound for the competitive ratio. The problem is further extended for the case of two possible delays for passengers and, therefore, a 3-competitive online algorithm is presented. Additionally, they propose a new objective that models the operator's total profit and find out that no deterministic algorithm can have a bounded competitiveness for this problem.

In Bender et al. (2013) the DM problem of a single train line from Gatto et al. (2007) is evaluated with other measures than the competitive analysis. The adversary is weakened, i.e., assuming that for the DM problem the delay at the following station is known (the algorithm can use *lookahead*). They measure the performance of the proposed online algorithm with weaker versions of the competitive ratio, namely comparative and average-case analysis where the expected cost of the online algorithm is determined with a probability function. Furthermore, a stylized stochastic program is developed that includes the number of delayed passengers as a discrete random variable. The decision to wait for a delayed feeder train is allowed to be taken only once. For a small example of three stations, the stochastic program outperforms the algorithm from average analysis and a balancing algorithm from literature. Due to the complex and time-consuming computation, the stochastic program was not part of the numerical study. But with the stochastic program, ideas for the class of [*pmash*] are proposed.

3.2.3 Macroscopic stochastic models [**pmash**]

The class of macroscopic stochastic models is rather scarce. All works in this class derive their solution heuristically.

Berger et al. (2011b) develop TOPSU-RDM, a simulation platform that evaluates different heuristics for the DM problem, drawing delays from underlying distribution functions. The platform combines the tasks of building a model, finding an appropriate solution algorithm and experimentally evaluating it. The implemented solution algorithms, called engines, contain several dispatching rules and a Monte Carlo tree search. The performance of the engines is evaluated

and the Monte Carlo search shows a rather poor performance as the solution time is restricted. Furthermore, a proof that making wait-depart decisions under incomplete information is PSPACE-hard is also given. The decisions depend on the global structure of the network, the schedule, the passenger routes and the imposed delays.

A stochastic dynamic program (SDP) incorporating delay distributions from statistical literature is developed in Schön and König (2018). Potential recourse actions for the decision process are determined on single train lines, considering effects on feeder and connecting trains. The objective function is modeled with a Bellman equation that minimizes passenger delays. A state space reduction speeding up solution times is applied for the solution. The SDP outperforms simple dispatching rules and a re-optimization strategy in a numerical study and yields results close to a full-information model.

3.2.4 Microscopic deterministic models [**pmidh**]

Finally we review DM models with a microscopic perspective on the network, consisting only of a small number of works. All models are deterministic and derive their solutions with heuristics [*pmidh*].

Corman et al. (2012) propose a *hybrid* approach that combines goals of DM and RTR. The bi-objective function minimizes train delays and missed passenger connections. Moreover, the model is built on the basis of an alternative graph with detailed infrastructure components as it is common for RTR. Therefore, we classify this work as $[\frac{p}{t}midh]$. To determine the Pareto front of non-dominated schedules, two heuristics are proposed and tested on data of the Dutch railways. The "compromise" solutions obtained by both heuristics seem promising for taking good dispatching decisions.

In Corman et al. (2017) DM and RTR are merged together yielding a "microscopic DM model". The microscopic perspective models the infrastructure while the passenger-centric objective aims to minimize the time spent in the system by passengers. Lower and upper bounds for the passenger flows of the resulting MIP are proposed. Several heuristics are designed that fix the train order or solve parts of the model iteratively. In a large numerical study with data from the Dutch railways the heuristics were able to solve also larger problem instances and reduce passenger waiting times.

In Xu et al. (2018) wait-depart decisions for last connections at a day's end are made. The model incorporates the passenger's choice behavior for transferring with the goal of maximizing the number of maintained connections and minimizing average waiting times. The constraints are formulated in a microscopic view to ensure feasibility of the disposition timetable. A genetic algorithm is developed and tested in a case study of Beijing's subway. With the aid of the algorithm, maintained connections of last trains could be increased.

Finally, all discussed works are summarized for each attribute class in Table 2, sorted in the sequence as discussed above. As all of the mentioned works relate to small disturbances [*m*], a column for the attribute of delay cause was omitted.

Table 2 Summary of discussed literature on DM neglecting the attribute level [m]

Publication	Attribute			
	Focus	Perspective	Input	Solution
Schöbel (2001)	p	a	d	e
Gatto et al. (2004)	p	a	d	e
Gatto et al. (2005)	p	a	d	e
Schöbel (2007)	p	a	d	e
Heilporn et al. (2008)	p	a	d	e
Ginkel and Schöbel (2007)	p, t	a	d	e
Schöbel (2009)	p	a, i	d	h
Schachtebeck and Schöbel (2010)	p	a	d	e, h
Dollevoet et al. (2012)	p	a	d	e
Schmidt (2013)	p	a	d	e
Dollevoet and Huisman (2014)	p	a	d	h
Dollevoet et al. (2014)	p, t	a	d	e
Dollevoet et al. (2015)	p	a	d	e, h
König and Schön (2020)	p	a	d	e, h
Suhl et al. (2001)	p	a, i	d, n	e, h
Biederbick and Suhl (2007)	p	a, i	d	h
Berger et al. (2011a)	p	a	d	h
Kanai et al. (2011)	p	a	d	h
Rückert et al. (2017)	p	a	d	h
Lemnian et al. (2016)	p	a	d	h
Kliewer and Suhl (2011)	p	a	d, n	e, h
Bauer and Schöbel (2014)	p	a	n	h
Anderegg et al. (2002)	p	a	n	h
Anderegg et al. (2009)	p	a	n	h
Gatto et al. (2007)	p	a	n	h
Gatto et al. (2008)	p	a	n	h
Krumke et al. (2011)	p	a	n	h
Bender et al. (2013)	p	a	n, s	h
Berger et al. (2011b)	p	a	s	h
Schön and König (2018)	p	a	s	h
Corman et al. (2012)	p, t	i	d	h
Corman et al. (2017)	p	i	d	h
Xu et al. (2018)	p	i	d	h

Applications on real-world data in the reviewed literature of Sect. 3.2 are presented in Table 3 for every country in alphabetical order. The applications are distinguished between countries the data set is taken from. The majority of the works are applied to the German network, which is a massive network requiring a lot of operational services every day, followed by the Dutch network. Some works from

Table 3 Applications on real-world data

Country	Publication
Belgium	Heilporn et al. (2008)
China	Xu et al. (2018)
Germany	Bauer and Schöbel (2014), Berger et al. (2011a, b), Biederbick and Suhl (2007), Ginkel and Schöbel (2007), Kliewer and Suhl (2011), König and Schön (2020), Lemnian et al. (2016), Rückert et al. (2017), Schachtebeck and Schöbel (2010), Schöbel (2001, 2007, 2009), Schön and König (2018), Suhl et al. (2001)
Greece	Bender et al. (2013)
Japan	Kanai et al. (2011)
Netherlands	Corman et al. (2012, 2017), Dollevoet and Huisman (2014), Dollevoet et al. (2012, 2014, 2015)

the class of [*pmade*] and [*pmanh*] are missing as these works contain theoretical considerations only. Other networks in the world that are not mentioned here might be also worth an investigation, providing further opportunities for research.

4 Conclusion and future research

We have reviewed the literature for railway DM problems. The area of DM models is embedded on the operational level of planning problems among related short-term problems. Influences between DM and the other problems such as, e.g., disruption management, RTR etc. are depicted. For the short-term problems, a new taxonomy scheme is developed that classifies the literature on the basis of five different attributes and their attribute levels. The taxonomy is applied to classify literature assigned to DM. With this classification scheme areas with scarce or even no works can be detected easily.

As seen in Sect. 3.2, all models have minor disturbances in common and nearly all of them focus on passengers only. There exist only three exceptions taking a train perspective additionally into account (Ginkel and Schöbel (2007), Corman et al. (2012), Dollevoet et al. (2014)). Furthermore, a macroscopic view is included in most of the works considered. Only six works build the model on a microscopic view, three of them together with a macroscopic view. This reflects the key criteria as explained in Sect. 2.1. Moreover, models with deterministic input represent the largest part of the existing literature (about two-thirds of all papers). Exact and heuristic solution methods therein are represented half-and-half. The other third consists of models with incomplete information, usually solved with heuristics. Stochastic

models are rather rare. In Berger et al. (2011b) a simulation platform using stochastic distribution functions is presented; Bender et al. (2013) briefly sketch a stochastic program and Schön and König (2018) model an SDP for a single train line. When considering the stochastic nature of delays, the question arises why not more stochastic approaches exist. One may argue that taking stochasticity into account may lead to problems that are harder to solve or cannot be solved fast enough for real-world applications. On the other hand, including delays after a known distribution results in models that are closer to the real world. As a deterministic setting seems to be too optimistic while a setting where nothing about the future is known might be too pessimistic, stochastic models could be a compromise that are worth future investigation.

The passenger perspective could be improved by learning more about passenger patterns, as e.g., in Ortega et al. (2018). Currently, DB also uses passenger patterns to represent different passenger groups. They developed in their research department the "persona concept" to better understand individual needs of their customers (Deutsche Bahn 2015). At the moment it is used for product development but it might also be helpful for a passenger-oriented dispatching. A further possibility is to integrate the passenger directly into dispatching decisions with an automated feedback system as it is proposed in Stelzer et al. (2016).

A different strategy to model short-term problems with an even stronger focus on the passenger is studied in Lijesen (2014). They anticipate the decisions of passengers how to reach the destination. In Keyhani et al. (2017) the latest point in time, when the journey of a passenger should start to reach his destination in time with a probability of nearly 100%, is determined. The included delay distributions originate from historical delay data of DB. The decision-making process from a passenger's point of view is also used in Schmidt et al. (2017) (see Sect. 2.2). They show how a passenger should decide for the continuation of his trip when a disruption of unknown dimension has occurred. This might be worth further investigation as all of these works take stochasticity for the delay into account and are "close to the customer".

Nearly all models in the literature on DM assume that passengers will always reach (even if delayed) their destination. In reality, passengers might abort their journey (be it on their own decision or due to external circumstances). In König and Schön (2020), a first model that focuses on spilling passengers due to overloaded trains is presented. However, further reasons for aborted trips should be analyzed, e.g., if no alternative connection is possible anymore. In last train scheduling, a special emphasize is put on how to dispatch the last train of the day. So far, only literature on metro systems exists, such as, e.g., Kang et al. (2015) with focus on timetable rescheduling and Xu et al. (2018) (see the class of [*pmidh*] in Sect. 3.2). Last train scheduling might be also interesting for railway providers of long-distance or regional trains, especially when railway companies have to pay for a hotel if the passenger has missed the last connection of the day (a common practice in Germany, see Deutsche Bahn 2019).

An overview of applications on real-world data in the reviewed literature is provided in Table 3. Numerical studies are done for some countries more often (e.g., Germany and the Netherlands) than for others. For many countries, no studies on

DM exist at all and a first step towards passenger-oriented solutions at the occurrence of delays might be worth looking at. Furthermore, the infrastructure of the investigated railway networks differs in size and shape. While, e.g., the network in Germany is massive and rather unstructured, the network in France has the shape of a star, concentrated on Paris (SNCF 2019) and Japan's Shinkansen runs on lines from North to South (Japan Rail Pass 2019). The evaluated literature always focuses on one country but it might be also interesting to compare the performance of the same DM model or dispatching rule on networks of different countries.

In Fig. 3 railway problems on the operational level and their interconnections are shown. Several links between these problems already exist, but literature for combined approaches is rare, e.g., Veelenturf et al. (2017) consider aspects from DM, RTR and rolling stock rescheduling in case of disruptions. The delay for passengers and the costs for the rescheduling of timetable and rolling stock are minimized equally. Passenger delays could be reduced by adjusting timetables without increasing rescheduling costs for rolling stock. Focusing on the delay for passengers or restoring a valid timetable solely might be falling short of an "optimal" solution for handling disturbances in the railway system. For railway providers it seems desirable to run holistic models that are able to serve passengers' and operators' needs.

The interconnection between the different problems on the operational level offers additional potential for further research. This seems to be not only possible for railways but also for other industries. A first work, motivated by railway DM is proposed in Santos et al. (2017). The authors introduce the Airline DM problem considering priority decisions and capacity restrictions for an airport; it is based (among others) on the model from Schachtebeck and Schöbel (2010).

References

Agrawal S, Wang Z, Ye Y (2014) A dynamic near-optimal algorithm for online linear programming. Oper Res 62(4):876–890. https://doi.org/10.1287/opre.2014.1289

Anderegg L, Penna P, Widmayer P (2002) Online train disposition: to wait or not to wait? Electron Notes Theor Comput Sci 66(6):32–41

Anderegg L, Penna P, Widmayer P (2009) Online train disposition: to wait or not to wait? In: Ahuja R, Möhring R, Zaroliagis C (eds) Robust and online large-scale optimization, Lecture Notes in Computer Science, vol 5868. Springer, Berlin, pp 387–398. https://doi.org/10.1007/978-3-642-05465-5_17

Bauer R, Schöbel A (2014) Rules of thumb: practical online-strategies for delay management. Public Transp 6(1):85–105. https://doi.org/10.1007/s12469-013-0082-8

Bender M, Büttner S, Krumke S (2013) Online delay management on a single train line: beyond competitive analysis. Public Transp 5(3):243–266. https://doi.org/10.1007/s12469-013-0070-z

Berger A, Blaar C, Gebhardt A, Müller-Hannemann M, Schnee M (2011a) Passenger flow-oriented train disposition. In: Demetrescu C, Halldórsson MM (eds) Algorithms—ESA 2011. Lecture Notes in Computer Science, vol 6942. Springer, Berlin, pp 227–238. https://doi.org/10.1007/978-3-642-23719-5_20

Berger A, Hoffmann R, Lorenz U, Stiller S (2011b) Online railway delay management: hardness, simulation and computation. Simulation 87(7):61–629. https://doi.org/10.1177/0037549710373571

Biederbick C, Suhl L (2007) Decision support tools for customer-oriented dispatching. In: Algorithmic methods for railway optimization. Springer, pp 171–183

Binder S, Maknoon Y, Bierlaire M (2017) The multi-objective railway timetable rescheduling problem. Transp Res C Emerg Technol 78:78–94

Cacchiani V, Toth P (2012) Nominal and robust train timetabling problems. Eur J Oper Res 219(3):727–737. https://doi.org/10.1016/j.ejor.2011.11.003

Cacchiani V, Caprara A, Galli L, Kroon L, Maróti G, Toth P (2012) Railway rolling stock planning: robustness against large disruptions. Transp Sci 46(2):217–232

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 Methodol 63:15–37

Caimi G, Fuchsberger M, Laumanns M, Lüthi M (2012) A model predictive control approach for discrete-time rescheduling in complex central railway station areas. Comput Oper Res 39(11):2578–2593

Caprara A, Toth P, Vigo D, Fischetti M (1998) Modeling and solving the crew rostering problem. Oper Res 46(6):820–830

Cicerone S, Di Stefano G, Schachtebeck M, Schöbel A (2012) Multi-stage recovery robustness for optimization problems: a new concept for planning under disturbances. Inf Sci 190:107–126

Corman F, D'Ariano A, Pacciarelli D, Pranzo M (2012) Bi-objective conflict detection and resolution in railway traffic management. Transp Res C Emerg Technol 20(1):79–94

Corman F, D'Ariano A, Marra AD, Pacciarelli D, Samà M (2017) Integrating train scheduling and delay management in real-time railway traffic control. Transp Res E Logist Transp Rev 105:213–239

Deutsche Bahn (2015) Im Bahn-Labor der innovativen Ideen. https://inside.bahn.de/innovationslab-interview/ (Accessed 29 Oct 2018)

Deutsche Bahn (2017) PANDA: Ein Assistent, der Ihre Anschlüsse verbessert. https://inside.bahn.de/anschlusszug-panda (Accessed 28 Oct 2018)

Deutsche Bahn (2019) Ihre Rechte als Fahrgast im nationalen Eisenbahnverkehr. https://www.bahn.de/p/view/service/auskunft/fahrgastrechte/nationale_regelungen.shtml (Accessed 10 Sept 2019)

Die Welt (2017) Lutz ist neuer Bahnchef—und das ist sein Problem. https://www.welt.de/wirtschaft/article163058133/Lutz-ist-neuer-Bahnchef-und-das-ist-sein-Problem.html (Accessed 28 Oct 2018)

Die Welt (2018) Zahl der Zugausfälle bei der Deutschen Bahn 2017 deutlich gestiegen. https://www.welt.de/newsticker/news1/article176137413/Verbraucher-Zahl-der-Zugausfaelle-bei-der-Deutschen-Bahn-2017-deutlich-gestiegen.html (Accessed 28 Oct 2018)

Dollevoet T, Huisman D (2014) Fast heuristics for delay management with passenger rerouting. Public Transp 6(1–2):67–84. https://doi.org/10.1007/s12469-013-0076-6

Dollevoet T, Huisman D, Schmidt M, Schöbel A (2012) Delay management with rerouting of passengers. Transp Sci 46(1):74–89. https://doi.org/10.1287/trsc.1110.0375

Dollevoet T, Corman F, D'Ariano A, Huisman D (2014) An iterative optimization framework for delay management and train scheduling. Flex Serv Manuf J 26(4):490–515

Dollevoet T, Huisman D, Kroon L, Schmidt M, Schöbel A (2015) Delay management including capacities of stations. Transp Sci 49(2):185–203. https://doi.org/10.1287/trsc.2013.0506

Dollevoet T, Huisman D, Schmidt M, Schöbel A (2018) Delay propagation and delay management in transportation networks. In: Handbook of optimization in the railway industry. Springer, pp 285–317

Fang W, Yang S, Yao X (2015) A survey on problem models and solution approaches to rescheduling in railway networks. IEEE Trans Intell Transp Syst 16(6):2997–3016

Gatto M, Glaus B, Jacob R, Peeters L, Widmayer P (2004) Railway delay management: exploring its algorithmic complexity. In: Scandinavian workshop on algorithm theory. Springer, pp 199–211

Gatto M, Jacob R, Peeters L, Schöbel A (2005) The computational complexity of delay management. In: Kratsch D (ed) Graph-theoretic concepts in computer science, Lecture Notes in Computer Science, vol 3787. Springer, Berlin, pp 227–238. https://doi.org/10.1007/11604686_20

Gatto M, Jacob R, Peeters L, Widmayer P (2007) Online delay management on a single train line. In: Geraets F, Kroon L, Schöbel A, Wagner D, Zaroliagis C (eds) Algorithmic methods for railway optimization, Lecture Notes in Computer Science, vol 4359. Springer, Berlin, pp 306–320. https://doi.org/10.1007/978-3-540-74247-0_17

Gatto M, Nunkesser M, Schöbel A, Widmayer P (2008) Delay management with restricted adversaries and ski rental. Tech. rep., ARRIVAL Project 1–8

Ghaemi N, Cats O, Goverde RM (2017) Railway disruption management challenges and possible solution directions. Public Transp 9(1–2):343–364

Ginkel A, Schöbel A (2007) To wait or not to wait? the bicriteria delay management problem in public transportation. Transp Sci 41(4):527–538. https://doi.org/10.1287/trsc.1070.0212

Goerigk M, Schachtebeck M, Schöbel A (2013) Evaluating line concepts using travel times and robustness. Public Transp 5(3):267–284

Goerigk M, Schmidt M, Schöbel A, Knoth M, Müller-Hannemann M (2014) The price of strict and light robustness in timetable information. Transp Sci 48(2):225–242. https://doi.org/10.1287/trsc.2013.0470

Heilporn G, De Giovanni L, Labbé M (2008) Optimization models for the single delay management problem in public transportation. Eur J Oper Res 189(3):762–774. https://doi.org/10.1016/j.ejor.2006.10.065

Jaillet P, Wagner MR (2010) Online optimization—an introduction. INFORMS TutORials in Operations Research, chap 6, pp 142–152. https://doi.org/10.1287/educ.1100.0072

Japan Rail Pass (2019) Shinkansen—the Japanese bullet trains. https://www.jrailpass.com/shinkansen-bullet-trains (Accessed 10 Sep 2019)

Kanai S, Shiina K, Harada S, Tomii N (2011) An optimal delay management algorithm from passengers' viewpoints considering the whole railway network. J Rail Transp Plan Manag 1(1):25–37

Kang L, Wu J, Sun H, Zhu X, Wang B (2015) A practical model for last train rescheduling with train delay in urban railway transit networks. Omega 50:29–42

Kecman P, Corman F, D'Ariano A, Goverde RM (2013) Rescheduling models for railway traffic management in large-scale networks. Public Transp 5(1–2):95–123

Keyhani MH, Schnee M, Weihe K (2017) Arrive in time by train with high probability. Transp Sci 51(4):1122–1137. https://doi.org/10.1287/trsc.2017.0758

Kliewer N, Suhl L (2011) A note on the online nature of the railway delay management problem. Networks 57(1):28–37. https://doi.org/10.1002/net.20381

König E, Schön C (2020) Railway delay management with passenger rerouting considering train capacity constraints. Working paper revised and resubmitted. Eur J Oper Res, pp 1–31

Kroon L, Maróti G, Nielsen L (2014) Rescheduling of railway rolling stock with dynamic passenger flows. Transp Sci 49(2):165–184

Krumke S, Thielen C, Zeck C (2011) Extensions to online delay management on a single train line: new bounds for delay minimization and profit maximization. Math Methods Oper Res 74(1):53–75. https://doi.org/10.1007/s00186-011-0349-2

Lan Y, Gao H, Ball M, Karaesmen I (2008) Revenue management with limited demand information. Manag Sci 54(9):1594–1609. https://doi.org/10.1287/mnsc.1080.0859

Lemnian M, Müller-Hannemann M, Rückert R (2016) Sensitivity analysis and coupled decisions in passenger flow-based train dispatching. OASIcs-OpenAccess Series in Informatics, Schloss Dagstuhl-Leibniz-Zentrum für Informatik 54:1–15. https://doi.org/10.4230/OASIcs.ATMOS.2016.2

Liebchen C, Schachtebeck M, Schöbel A, Stiller S, Prigge A (2010) Computing delay resistant railway timetables. Comput Oper Res 37(5):857–868. https://doi.org/10.1016/j.cor.2009.03.022

Lijesen MG (2014) Optimal traveler responses to stochastic delays in public transport. Transp Sci 48(2):256–264. https://doi.org/10.1287/trsc.2013.0465

Louwerse I, Huisman D (2014) Adjusting a railway timetable in case of partial or complete blockades. Eur J Oper Res 235(3):583–593

Lusby RM, Larsen J, Ehrgott M, Ryan D (2011) Railway track allocation: models and methods. OR Spectrum 33(4):843–883. https://doi.org/10.1007/s00291-009-0189-0

Lusby RM, Larsen J, Bull S (2018) A survey on robustness in railway planning. Eur J Oper Res 266(1):1–15. https://doi.org/10.1016/j.ejor.2017.07.044

Malucelli F, Tresoldi E (2019) Delay and disruption management in local public transportation via real-time vehicle and crew re-scheduling: a case study. Public Transp 11(1):1–25. https://doi.org/10.1007/s12469-019-00196-y

Müller-Hannemann M, Rückert R (2017) Dynamic event-activity networks in public transportation. Datenbank-Spektrum 17(2):131–137. https://doi.org/10.1007/s13222-017-0252-y

Ortega FA, Pozo MA, Puerto J (2018) On-line timetable rescheduling in a transit line. Transp Sci 52(5):1106–1121. https://doi.org/10.1287/trsc.2017.0807

Papadimitriou CH (2003) Computational complexity. In: Encyclopedia of computer science. Wiley, Chichester, pp 260–265

Rückert R, Lemnian M, Blendinger C, Rechner S, Müller-Hannemann M (2017) Panda: a software tool for improved train dispatching with focus on passenger flows. Public Transp 9(1–2):307–324

Santos BF, Wormer MM, Achola TA, Curran R (2017) Airline delay management problem with airport capacity constraints and priority decisions. J Air Transp Manag 63:34–44

Schachtebeck M, Schöbel A (2010) To wait or not to wait—and who goes first? Delay management with priority decisions. Transp Sci 44(3):307–321. https://doi.org/10.1287/trsc.1100.0318

Schmidt M (2013) Simultaneous optimization of delay management decisions and passenger routes. Public Transp 5(1):125–147. https://doi.org/10.1007/s12469-013-0069-5

Schmidt M, Kroon L, Schöbel A, Bouman P (2017) The travelers route choice problem under uncertainty: dominance relations between strategies. Oper Res 65(1):184–199. https://doi.org/10.1287/opre.2016.1564

Schöbel A (2001) A model for the delay management problem based on mixed-integer-programming. Electron Notes Theor Comput Sci 50(1):1–10. https://doi.org/10.1016/S1571-0661(04)00160-4

Schöbel A (2007) Integer programming approaches for solving the delay management problem. In: Geraets F, Kroon L, Schöbel A, Wagner D, Zaroliagis C (eds) Algorithmic methods for railway optimization, Lecture Notes in Computer Science, vol 4359. Springer, Berlin, pp 145–170. https://doi.org/10.1007/978-3-540-74247-0_7

Schöbel A (2009) Capacity constraints in delay management. Public Transp 1(2):135–154. https://doi.org/10.1007/s12469-009-0010-0

Schön C, König E (2018) A stochastic dynamic programming approach for delay management of a single train line. Eur J Oper Res 271(2):501–518. https://doi.org/10.1016/j.ejor.2018.05.031

SNCF (2019) Sncf réseau. https://www.sncf.com/en/network-expertise/rail-network/sncf-reseau (Accessed 10 Sep 2019)

Stelzer A (2016) Automatisierte Konfliktbewertung und -lösung für die Anschlussdisposition im (Schienen-) Personenverkehr. Schriftenreihe des Instituts für Verkehr, Technische Universität Darmstadt

Stelzer A, Englert F, Hörold S, Mayas C (2016) Improving service quality in public transportation systems using automated customer feedback. Transp Res E Logist Transp Rev 89:259–271

Suhl L, Biederbick C, Kliewer N (2001) Design of customer-oriented dispatching support for railways. In: Voß S, Daduna JR (eds) Computer-aided scheduling of public transport. Springer, Berlin, pp 365–386. https://doi.org/10.1007/978-3-642-56423-9_20

Veelenturf L, Kidd M, Cacchiani V, Kroon L, Toth P (2016a) A railway timetable rescheduling approach for handling large-scale disruptions. Transp Sci 50(3):841–862. https://doi.org/10.1287/trsc.2015.0618

Veelenturf LP, Potthoff D, Huisman D, Kroon LG, Maróti G, Wagelmans APM (2016b) A quasi-robust optimization approach for crew rescheduling. Transp Sci 50(1):204–215. https://doi.org/10.1287/trsc.2014.0545

Veelenturf LP, Kroon LG, Maróti G (2017) Passenger oriented railway disruption management by adapting timetables and rolling stock schedules. Transp Res C Emerg Technol 80:133–147

Verhaegh T, Huisman D, Fioole PJ, Vera JC (2017) A heuristic for real-time crew rescheduling during small disruptions. Public Transp 9(1–2):325–342

Xu W, Zhao P, Ning L (2018) Last train delay management in urban rail transit network: bi-objective MIP model and genetic algorithm. KSCE J Civ Eng 22(4):1436–1445