Strategic interactions between minimization of train delays and passenger assignment in Microscopic Railway Delay Management

Francesco Corman

Abstract
Recent optimization models for railway traffic rescheduling mostly focused on incorporating an increasing detail of the infrastructure, with the goal of proving feasibility and quality from the point of view of the infrastructure managers, i.e. the train dispatchers actually managing traffic in real-time. Different approaches neglect precise train operations, but consider passenger flows to focus more explicitly on the quality of service perceived by the passengers, which is instead in the interest of the train operating companies.

This work deals with a microscopic railway delay management, determining precise train operations and including passenger flows. In particular, key point of research is the interaction between the problem (of the infrastructure manager) to reschedule trains and the problem (of the travellers) to find the optimal route in the network. In fact, changing passenger flows, respectively delaying trains and/or dropping passenger connections varies the setting under which the two decision makers respectively interact. This paper studies this interconnected problem as mediated by the information one decision maker has about the other.

State-of-the-art algorithms for the microscopic railway delay management problem are analysed from the point of view of passengers and train traffic. Some solutions correspond to Nash equilibria for different strategic settings of the game. Computational results based on a real-world Dutch railway network describe the trade-offs between the minimization of train delays and passenger travel times.

Keywords: Public transport, Transit Assignment, Delay Management, Rescheduling
1 Introduction and literature review

Optimization of railway service is a key factor to reduce congestion on multi-modal networks, especially in densely populated areas, and to provide an eco-friendly and sustainable way of transport. In order to attract new customers from other transport modes, European countries defined challenging targets in terms of Quality of Service (QoS) that the railway companies should provide to their customers (Caprara et al, 2006; EU, 2013; Pellegrini et al, 2014).

Many stakeholders interplay in the railway transport perspective: passengers who want to move from their origin to their destination, Train Operating Companies (TOC) which are selling transport services to passengers between station to station, the Infrastructure Manager (IM) who sells infrastructure capacity to the TOC along railway lines. The viewpoint of the different stakeholders is different and translates into different approaches to determine a best (or optimal) solution when managing railway traffic when delays occur. This paper addresses trade-off and strategic interactions among the objectives of the above-mentioned stakeholders, IM, and passengers (which we assume have the same interest of the TOC). While the IM objective relates to train delays, the passengers/TOC aims at minimizing the passenger travel time.

We show that those two slightly different objectives align sometimes, but not always and the interaction of them might result in interesting strategic interactions. The two objectives are very close to each other: in heavily used railway networks any small delay of a few trains easily propagates to the other trains, and rescheduling trains in real-time in order to minimize train delays, also directly addresses travel time of passengers.

This translates to a train rescheduling problem, designing in realtime a microscopically feasible schedule minimizing train delays (Cacchiani et al, 2014). The complexity of the microscopic train rescheduling problem (which includes determining typical control actions at level of single switches and signals, such as retiming, reordering, rerouting train traffic) is due to the limited available capacity of railway lines and stations, plus the constraints required to model the operating and safety rules, including the behavior of the fail-safe signalling system (Hansen and Pachl, 2014). Since the problem of computing a feasible solution has been shown to be NP-hard (Mascis and Pacciarelli, 2002), a lot of academic efforts tackled the complexity of this problem. Two main streams are to be found; one bases on combinatorial models exploiting graph theoretical structures, called alternative graph, i.e. a generalization of the disjunctive graph; and the other use more general purpose MILP formulations. Among the former category, we report here (Corman et al 2011; Corman et al, 2012; Corman et al, 201A; D’Ariano, 2008; D’Ariano et al, 2007; Samà et al, 2016); for the latter, we refer to (Carey & Crawford, 2007; Lamorgese and Mannino, 2015; Meng and Zhou, 2014; Pellegrini et al, 2014; Tornquist and Persson, 2007; Van Thielen et al, 2018). Most of those approaches neglect passengers, or
include them as static weights in objective functions, with some exceptions (priority classes: (Corman et al, 2011); Pareto Front of Connections kept: (Corman et al, 2012)).

The optimization of the QoS perceived by the passengers in customer-oriented train dispatching is introduced by (Schöbel, 2007). The passenger passenger dissatisfaction is often translated into some (weighted) travel time formula (for instance Espinosa-Aranda and García Rodenas, 2013). When passengers have a more active role, and can take decisions as reactions of the observed traffic, interesting optimization problem arise. In fact, changing train orders may result in extra delays for those passengers that miss a connection at some station, therefore the passengers/TOC would like to keep the connections that are more relevant to some passengers in order to minimize their passenger travel times.

This tradeoff has been already well recognized by the delay management approaches (to wait or not to wait for a delayed train?). Such delay management problem has been gradually investigated under increasing realism (Schachtebeck and Schöbel, 2010; Dollevoet et al, 2012; Dollevoet et al, 2015), including choices to reroute passengers between their origin and destination, which is a second crucial decisions for passenger flows. Approaches in this stream are typically based on macroscopic models, considering the arrival/departure events of trains at stations only, and neglecting further infrastructure capacity, or train separation due to the safety system. For this reason, there is a gap between the QoS promised by the solutions delivered and the one that can be achieved when implementing the solutions within the actual limitations and rules of practice.

Some recent approaches consider passenger flows integrated in microscopic train rescheduling. (Tomii et al, 2005) introduce a first microscopic railway traffic rescheduling model with the minimization of passenger dissatisfaction. (Sato et al, 2013) address the problem of minimizing passenger inconvenience (related to time spent) on simple railway lines, taking into account disruptions that might require adjustment of vehicle schedules. They report in general that trains can decrease passengers inconvenience by increasing delays. (Dollevoet et al, 2014) iterate a macroscopic delay management approach with a microscopic train rescheduling approach. A microscopic delay management model with passenger routing is introduced in (Corman et al, 2016), in which passengers are always assigned to a shortest path between their origin and destination. The problem is modelled as a MILP and solved by various heuristic, considering only the TOC view point. (Binder et al, 2017) directly focus on the problems of considering capacity limitations, albeit only in the vehicle (i.e. not in the infrastructure), as one critical aspect in the choice and actions of travelers. In conclusion, most of the microscopic train traffic management approaches of the literature still neglect the impact of train rescheduling decisions on the QoS to passenger, while most of the approaches focusing on the passenger perspective miss a detailed description of the rail infrastructure.

This work evaluates the strategies of passengers/TOC and IM by using microscopic optimization models, computing microscopically feasible real-time train rescheduling
actions for railway traffic under delays. The following research questions are investigated:

(a) Which is the impact of passenger travel time minimization on train delay minimization and vice versa?

(b) At what extent the rescheduling methods commonly adopted in practice can satisfactorily address the objectives of the stakeholders?

To address the above questions, we use the Microscopic Delay Management model proposed in (Corman et al, 2016). The peculiarity of the formulation proposed in this paper resides further in the combination of the microscopic level of detail in managing traffic flows inside the station areas. This level of detail is required to cope with train conflicts occurring at or around busy station areas and at interlocking areas, when dealing with delayed traffic. We use here generalizations of the heuristics proposed in (Corman et al 2016), which iteratively solve a train scheduling and routing problem integrated with a passenger assignment or routing problem (corresponding to the IM and passenger problems). The solutions delivered by these algorithms correspond to Nash equilibria of the strategic interaction between IM and passengers players.

We here remark that game theoretical (or algorithmic) studies in the field of railway transport scheduling are quite rare, with a few works tackling the offline timetabling point of view, where the strategic and political effect of decisions is much larger (Klabes, 2010). At that stage, actors are trains competing for capacity in the planning phase or in real time, and have utility by running with the best path, or the lowest delay. Actors can also not be trains per se, but train operators, while the economic appeal of the entire service intention is the goal of the game. For instance, combinatorial auctions have been used to competitively assign freight paths among train operators (Harrods, 2013) in a planning phase. Game theoretical approaches, mostly based on cooperative game theory, and transferable utility assumptions, to study this allocation can be found for instance in (Kuo and Miller Hooks, 2012). A non-transferable utility game theoretical study is proposed in the recent paper by (Fragnelli and Sanguineti, 2014) which examines the need for information exchange for cooperative timetabling among multiple train operators. Companies exchange preferences about their needs and get compensation, under limited information disclosure.

When actors are travelers competing for the limited capacity in a vehicle, or even the seats in a vehicle, the problem is typically known as assignment. This problem has been studied from a large range of different angles. (Bouman et al, 2017) report on minority games, focusing on crowding dynamics, in railway assignment under disturbances. They report on a study clarifying the impact that the availability of information has on crowding, as well as the impact of the optimization from the train operator, which can change the capacity by using different rolling stock. Both actions have direct impact on passengers’ satisfaction, here seen as the ultimate goal of the actors. (Binder et al, 2017) also studies the existence of fixed point, similar to a Nash equilibrium in a repeated game setting, for capacitated railway assignment under disturbed conditions. The importance of compliance to information is becoming
crucial, especially when the setting of a repeated game is not perfectly matching the reality of specific disturbances happening only once. We refer the interested reader to the applications of game theory in transport networks in (Zhang et al, 2010), as well as (Hollander and Prashker, 2006) and (Easley and Kleinberg, 2010). In conclusion, while many papers have some game theoretical approach to report on particular evaluations of utility, or outcomes of a shared decision process, no approach is currently known which considers the available capacity of the infrastructure, the infrastructure managers who has to assign it, and the passengers as actors competing/collaborating for a satisfactory transport performance.

In our setup, to move towards a Nash equilibrium, iterative traffic optimization models are considered, which alternate a move of IM, i.e. a microscopic train scheduling and routing model with the minimization of the total weighted train delays, and a move of the passengers, i.e. a passenger assignment model with minimization of the total passengers' travel time for a given train schedule. We show that the inherent extension to include infrastructure capacity, i.e. the microscopic detail included in the microscopic delay management, brings issues in existence of equilibrium point, even when no capacity of the vehicle is considered. To better understand this, and also to result in benchmarks, we thus consider also approaches which consider a limited or no form of interaction. Those are for instance the timetable solution, and the exact train rescheduling algorithm in (D’Ariano et al, 2007). Computational experiments are based on a real-world and large-scale Dutch railway network. We study the implications from the point of view of the algorithmic development, computation of a solution, as well as the practical implications.

Section 2 presents the problem, which is analyzed from a game theoretical perspective in section 4. Section 5 reports on some larger scale numerical experiments; Section 5 concludes the paper.

2 Problem Description and Model Formulation

We address the problem of computing in real time a microscopically feasible disposition schedule for some disturbed train traffic, and a passenger route advice for each passenger so that passengers have the least possible discomfort. In general, discomfort can be related to the time spent in various conditions. However, according to (Binder et al, 2017), in this paper we directly adopt the passenger travel time as a surrogate for the passenger discomfort.

We refer to Figure 1 for a graphical summary aiming at considering all aspects of the model. Passengers start from an origin station at given departure time (which is the time they enter the station, and not the time they board their first train) and want to reach a destination station as soon as possible.
We discretize the departure time of passengers, and refer to a group of passengers departing at the same origin $o$ at the same time $w$ and having the same destination $d$ as a demand $odw$. We call passenger generation the definition of all demands. Passengers have available some information which makes them determine a route advice. This is a strategy in terms of travelling along the network, involving a series of trains connect the departs stations and the destination station, through a series of feasible connections. Typically, route advices are now available to passengers by online applications and websites, or by personal construction of a plan when looking at printed timetables at stations.

A passenger connection is the transfer of some passengers from a feeder train to a connected train at an intermediate station along the demand route (either a route advice or a realized path). The train rescheduling decisions determine the possibility to activate a passenger connection, in order to reach the destination. In the given train schedule, the arrival time of the feeder train must be sufficiently in advance compared to the departure time of the connected train.

In presence of traffic disturbances, train rescheduling decisions are necessary in order to recover feasibility of operations, and keep fluidity of running. The plan of arrivals/departures described by the timetable is adjusted during operations, while respecting the constraints on the limited capacity of the railway infrastructure.

The latter constraints limit the possibility of rescheduling train movements, since the railway safety regulations must be respected while considering the signalling system, the speed of each train and its relative position with respect to the other trains in the network. According to the commonly used fixed block signaling system (Hansen and Pachl, 2014), the railway network is partitioned in block sections separated by signals. When a block section is occupied by a train, the signaling system forces other incoming trains to stop before the signal, while the signal aspect at the entrance of the previous block section forces incoming train to reduce their speed. A train can run at high speed only if the two block sections ahead are empty, while it can run at reduced speed if only the block section ahead is empty. Moreover, each block section can host at most one train at a time. This focus on limited infrastructure capacity is a key innovation from the state of the art of delay management.

A train cannot depart from a station before its scheduled departure time and its departure from a station can be constrained to be sufficiently larger than the arrival of another (feeder) train so that the passengers can move from the latter to the former. Moreover, a train is late when its arrival time at a station is larger than the scheduled
arrival time. The delay of a train can be caused either by external disturbances related to that specific train or by the propagation of delays from other trains due to the operational constraints of the railway system. Specifically, *initial delays* are due to disturbances that can be recovered only to a certain extent by exploiting running time supplements of the timetable. *Consecutive delays* are determined by rescheduling decisions in response to initial delays, and are the result of a train delay propagation due to some conflicting situations. A *conflict* occurs whenever two trains require the same block section at the same time, due to some (relatively small) delay. The conflict is typically solved by adjusting departure times from stations and passing times along the network (re-timing), or by specifying a passing order for the trains at the block section (re-ordering). We do not consider rerouting actions at the current stage, and leave this extension for future research.

The problem of controlling railway traffic with microscopic detail corresponds to a job shop scheduling problem with blocking no-swap constraints (Mascis and Pacciarelli 2002). We use the general alternative graph model for this purpose, which refers to a graph described by nodes $N$ (associated to trains entering block sections), fixed arcs $F$ (representing physical constraints of the movement), alternative arcs $S$ (representing ordering decisions and separation) and connection arcs $C$ (representing a minimum headway between a feeder and a connected train). Blocking no-swap constraints model the so-called fixed-block railway regulation that a train on a given block section cannot move forward if the block section ahead is not available or if it is occupied by another train. Detailed information on the modelling of the occupation time of block sections can be found, e.g. in (Hansen and Pachl, 2014).

The passenger routing problem (or passenger assignment problem) studies the distribution of passengers onto the railway network. We consider a time discrete model for passenger arrivals at each station. Hence, we assume to know the number of passengers willing to reach the same destination $d$ (out of a set $D$), from the same origin $o$ (out of a set $O$) starting their journey at the same time $w$ for a discrete set of arrival times $W$. We also assume that each train has infinite capacity and each passenger aims at reaching his/her destination in the minimum time. In this paper, we refer to a group of passengers going from $o$ to $d$ and arriving in $o$ at time $w$ as a triple $odw$, hereinafter denoted as *demand*, and let $ODW$ be the set of all demands $odw$. Note that, once the train schedule is fixed, each demand $odw$ moves in the network independently from the other triples, a minimum multicommodity minimum cost flow problem on a particular graph $(N, F \cup S \cup C)$ (comprising fixed arcs $F$, orders $S$ and connections $C$ between services connecting Nodes $N$ relating to events), can well represent their realized path, i.e., the choice of a particular routing for a given $odw$. 
3 Game theoretical analysis

Solution procedure

We next present the solution scheme investigated in this paper, which determines some heuristic approaches to solve the studied problem. The scheme is based on a game-theoretical setting, in which the decisions of the two players are successive and reactive on each other's decision. In general, a game is a set of decision by a set of players. In our case, multiple players exist, namely a IM player, and a multitude of passengers players. Differently from the majority of the solution frameworks, in our case their interests align, but only partially.

The decision of a player is the set of all actions that a player will take while dealing with the problem at hand. In our case, the possible actions of a passenger are to wait at their origin station, board some trains, possibly alight somewhere, wait and board a new train, and possibly many times, finally alight at the destination station. The objective of each passenger is always to minimize their travel time, i.e. the total time spent in those actions. The possible actions of an IM player are to change times and orders of trains. Those can be changed to a fixed value, or can be changed to a functional form, i.e. the times and orders of trains would be such that a specific function is met. In our setup, IM and passengers have a complete information of their actions, and also have a complete information of the other players choices.

Formally (see Stankova, 2009 for a longer discussion), we call \( D \) the set of all possible decisions of player \( i \), where \( i \) can be for instance \{IM, passenger\_1, passenger\_2, ... passenger\_n\}. The set of taken decisions is \( u \). The combination \( u = \{u_{IM}, u_1, u_2, ..., u_n\} \) is a decision profile, which we associate to a solution to the problem. We call \( J_i = J_i(u) \) the objective function of the player \( i \). Given a solution profile, it is possible to compute the objective function of all players involved. The objective function is individual and depends on the decision profile chosen by the same player, and by the other players. We assume a smaller objective function is preferred by each players (minimization).

Nash equilibrium is a solution \( u^* = \{u_{IM}^*, u_1^*, u_2^*, ..., u_n^*\} \) such that no unilateral deviation from it would result in an improvement of objective function for any player; formally, \( J_i(u_{IM}^*, u_1^*, ..., u_i^*, u_{i+1}^*, ..., u_n^*) \leq J_i(u_{IM}, u_1, u_2, ..., u_{i-1}, u_i, u_{i+1}, ..., u_n) \) for any \( u_i \neq u_i^* \).

A Stackelberg game identifies explicitly two roles, a leader and some followers. It is natural to identify the leader in our case as the IM (L=IM), and the followers as the passengers \( F = \{1, 2, ..., n\} \).

In a multi follower Stackelberg setting, the leader announces the decision \( u_{IM} \) to the followers, which can then react choosing their own decisions, i.e. \( u_i^* = \text{argmin} J_i \) \( (u_{IM}, u_1, u_2, ..., u_{i-1}, u_i^*, u_{i+1}, ..., u_n) \) or \( \text{argmin} J_{F_i}(u_{IM}, u_i^*) \) in the case that a the objective function related to a decision is independent from the decisions of other followers (which we actually assume in our setting). The decision \( u_i^* \) could be expressed as a reaction function \( l() \) applied to the choice of the leader: \( u_i^* = l(u_{IM}) \).
In a multi-follower inverse Stackelberg setting, the leader follows a strategy. A strategy \( \gamma_{IM} \) is a function able to determine a decision \( D_{IM} \) based on the set of all decisions from the followers. In general, this function might return a vector of values (i.e., a multi-dimensional decision, for instance involving multiple train times, and train orders). The followers then minimize their objective function, conditional to this knowledge. In other terms, \( u_i^* = \arg\min J_i(\gamma_{IM}(u_1, u_2, \ldots, u_N), u_1, u_2, \ldots, u_n) \) for each follower \( i \), given some decisions taken by the other followers. If the leader would desire some decision \( u^*_F = \{u_1^*, u_2^*, \ldots, u_n^*\} \) from the followers, a dominant strategy \( \gamma_{IM}^* \) is one that satisfies:

\[
\arg\min J_i(\gamma_{IM}^*(u_1, u_2, \ldots, u_N), u_1, u_2, \ldots, u_n) = u^*_i \quad \text{for any arbitrary decisions from the followers} \quad u_1, u_2, \ldots, u_n, \quad \text{for any follower} \quad i.
\]

We assumed in the derivation of those formulas that the IM is the leader and the passengers are the followers, for the reason that the IM is a single role, with a clear business objective; in a sense, the IM has a direct advantage of improving the utility of the passengers. If the IM do take decisions that are unnecessarily against the interest of the passengers and train operating company moving them, complaints and investigations might arise (in fact, in the EU legal framework, in the strictest application, the utility of the IM does not relate directly to the passengers’ utility, but via the intermediation of the TOC objectives). We also assume that the passengers are many and do not have direct decision power upon the IM, and that moreover the passengers do not influence directly each others’ decision, but do so only via a decision of the IM affecting both. This relates to the independence of travel and dwell times on the amount of passengers onboard, as well as the infinite capacity of the vehicles. Having finite capacity of the vehicle a direct competition between travelers would arise, which would complicate further the scheme.

In our setup, different various utility functions and reaction functions between IM and passengers can be considered. The objective function of the passengers and their decision is discussed first. Every passenger decides their route advice based on the expected travel time. In other terms, they would minimize their total travel time \( \sum_{odw} n_{odw}(T_{odw} - \Pi_{odw}) \), based on some information and route advice available. This problem has been already investigated much in the literature as a transit assignment problem. The total travel time is defined as the amount of passengers \( n_{odw} \) for each \( odw \), multiplied by the corresponding travel time. This latter is expressed as the arrival time at the destination, \( T_{odw} \), minus the generation of the group of passengers at the origin station, \( \Pi_{odw} \). A solution to the passenger assignment problem is the sequence of trains to board and change, for each passengers; and aggregating over all passengers, the amount of passengers assigned per train, between any two stations and sections. Related to this, one can compute the amount of passenger disembarking the train at each stop; the amount of passengers changing the train at each stop; the amount of passengers arriving at their final destination at each stop. Those figures could be useful in determining the strategy of the IM.

The problem of the IM relates to a train rescheduling problem, which has been defined in the literature as online updating of train paths, by train re-timing, re-ordering decisions with the minimization of a certain function of traffic performance,
related to delays. In the literature, no consensus is found on which exact delay expression can be used (see for instance, Cacchiani et al. 2014; Samà et al., 2015) when only the train-related operations are to be considered; and enlarging the study to include also passenger decisions would not make the problem easier. We consider objective functions which are based on some form of a linear combination of a lot of factors $f_e z_e$, which refer to some events $e$, related to a train arriving at a station, or passing at relevant points in the network; $f_e$ is the amount of passengers related to event $e$; and $z_e$ is the delay associated to event $e$ compared to the published timetable.

We consider different ways to combine together those $f_e z_e$, as follows. A solution that has all $f_e = 1$ practically reflect the case where the IM does not directly care of the amount of passengers on board the trains, nor on their travel time, but only on the delay of each train, further measured at all events $e$. If the event $e$ is only one, and in particular, the one associated to the maximum consecutive delay, the objective function is then analogous to the model of (D’Ariano et al., 2007). Otherwise, a total delay, or average delay form can be determined. As those decisions of the IM do not take into account the passengers, one can think of them as a Stackelberg Game, where the IM published their strategy, which depends on the delays of trains, and passengers take their decisions accordingly. Two cases we consider in more detail are the solution indeed minimizing the maximum consecutive delay, i.e. the solution of D’Ariano et al., 2007; and the solution keeping the timetable order, i.e. a solution for which the train order are as in the timetable, and the train times are a direct consequence of them.

**Timetable.** With this strategy, the train schedule is simply obtained by keeping the same train sequence of the timetable and delaying each train by the minimum amount needed to achieve feasibility. Passengers then follow the path which would bring them as fast as possible to their respective destinations. This approach simulates the common practice of railway management in which IM keeps the order of trains prescribed by the timetable, while passengers react individually to delays by choosing the most convenient route in real-time, based on the information available to them.

**Train Rescheduling only.** With this strategy, the train schedule is simply obtained by keeping the same train sequence of the timetable and delaying each train by the minimum amount needed to achieve feasibility. Passengers then follow the path which would bring them as fast as possible to their respective destinations. This approach simulates the solution which would be taken if the optimization approach introduced in (D’Ariano et al., 2007), possibly extended to deal with routing flexibility, would be implemented; this approach neglects passenger travel time or passenger delay. Passengers would react individually to delays by choosing the most convenient route in real-time, based on the information available to them.

Some other approaches considering passenger flow and interaction to some extent are now reported. Some of them have been defined in (Corman et al., 2016):

**Nash 1 (Integral of flows onboard).** This Nash equilibrium achieving a compromise between the passengers and IM objectives. In this game the IM strategy is the minimization of a weighted train total delay, the weight of each train being equal to the number of passengers onboard the train. Thus, the passengers are
considered as a flow onboard the trains, as the integral of passenger boarding and alighting along the travel. The starting solution is the one in which trains are rescheduled minimizing their delay, regardless of passengers. The TOC strategy is based on route advices which minimize the average passenger travel time, i.e. passenger react individually to the rescheduling actions by choosing the most convenient route in real-time.

**Nash 2 (Gradient of disembarking flow).** This Nash equilibrium, obtained with a slightly different game. Also in this game the IM strategy consists is the minimization of a weighted train total delay, but the weight of each train is equal to the number of passengers disembarking the train. Thus, the passengers are considered as the variation in flows entering and leaving the train along the travel. The starting solution is the one in which expected passengers flows are computed according to the timetable plan, i.e., regardless of delays. The passengers strategy is as for Nash 1. The final solution corresponds numerically to the heuristic H3 in (Corman et al, 2016).

**Nash 3 (Optimizing connections).** This Nash equilibrium, is obtained further as a variant of Nash 2. This heuristic differs in the selection of the active connections. This is based on the identification of promising connections to be enforced, i.e. of the lost connections in which the arrival time of passengers is slightly later than the departure time of the connected train. Connections are iteratively enforced as far as passenger traveltime improves. In a sense, this mixes up objective function of passengers and IM. The IM and TOC strategies are the same of Nash2. The starting solution is the one in which expected passengers flows are computed according to the timetable plan, i.e., regardless of delays. The final solution corresponds numerically to the heuristic H4 in (Corman et al, 2016).

The game-theoretical approaches based on the proposed solution scheme alternate a train rescheduling phase, optimizing train orders and times for given passenger flows and network capacity, to the passenger assignment phase, in which passengers’ travel time is minimized and computed for the given train schedule. The procedure iterates until convergence (i.e., until the current overall solution of train rescheduling and passenger assignment is equivalent to the one found in the previous iteration), that would correspond to an equilibrium. We assume here that such equilibrium can be found, even though in general this might not be the case, see next Subsection. In case no equilibrium is found within a maximum amount of iterations, the procedure is stopped and the best solution is reported.

Note that this is a static problem, i.e. all actors are not expected to learn from the past, or take different actions over time, if they would be again the same situation. This means, that compliance of actors is not considered, for instance based on past performance (as suggested instead in Binder et al, 2017); that the conditions under which actors take their decisions do not change over time; and finally, that there is no dynamic learning process as expressed for instance in the dynamic games of (Stankova, 2009).
Existence of Equilibrium point

We now discuss the possibility of having equilibrium in pure strategies for heuristics of the form comparable to H1 H2 H3, under some restrictive conditions. In general, depending on the way the infrastructure manager takes decisions based on the observed choices of the passengers, some equilibrium point might exist, or not. We show in this section, that some reaction types might result in situations that do not converge to a situation of equilibrium, and the game might keep oscillating. This is due to the discrete nature of the scheduling problem. Similar findings, though only in an empirical result of an algorithm and without a properly defined counter example, have been reported in (Binder et al, 2017). More in general, the existence of equilibrium point for continuous, convex, or non-continuous, non-convex domains, and possibilities to make games have an equilibrium point in general, is discussed in (Rosenthal, 1973).

We refer to Figure 2, Figure 3 and Table 1, where we describe a delayed situation of train traffic running in a network. 4 trains are running from left to right (labeled A B C D), across 5 stations (labeled P Q R S V).

Train A starts at P and ends at V; train B starts at P and ends in R; train C starts Q and passes V before reaching S, train D runs from R to V. Train A and B share a part of the infrastructure, and so do train B and C. We will see that those influences for the available capacity are the key mechanisms to inhibit reaching a fixed point. We assume train A is delayed by an unspecified amount \( Y \), and as a consequence of this delay it departs at the same time as train B. Both trains arrive at the same time at the bottleneck section, around time 4.

Figure 2 (middle and bottom) are time distance paths. Time (vertical, increasing downwards) goes from 0 to 20, while distance (horizontal, matching the infrastructure above) spans the 4 stations. Trains running on parallel stretches of the network (such as C and D, for instance) are reported anyway on the same horizontal location, and conflicts, if occurring, are highlighted, to keep the figure visible. We also report in dotted line the planned time distance path, i.e. assuming no infrastructure conflict. We report as vertical white boxes time spent at a station where people can change train and continue. In particular, there is a transfer between train B and D at station R; and a transfer time between train A and C at station V.

In Figure 3, we report the passenger loading for 4 situations, i.e. the same time-distance path as in Figure 2, but only focusing on where passengers are moving, symbolized by the thick lines. The top plots in Figure 3 represent two different passenger assignments on the top plot of Figure 2, and similarly for the bottom plots. The left (respectively right) plot of Figure 3 represent two possible solutions of passenger assignment for each time-distance path (Train rescheduling solution).

Table 1 represent, graphically for the same 4 situations as in Figure 3, the (total and consecutive) delays of trains, and the amount of passengers on the train just before the stop. As for the delay, each train has only one stop, apart from train C, so 5 rows are reported. We also consider some initial delay \( Y \) for Train A, which is actually
non-influent to the reasoning. As for the passengers, we assume a (negligible) small amount $\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D$, respectively for each train, plus a single passenger (we would name it X) who is the focus of our study. One could think of this passenger as a larger passenger group, and proportionally scale the $\epsilon$ conversely.

Please note that the order between train B and C is not relevant, as train C cannot come before train B in any case. Also assume that no other train conflict with no other train, i.e. capacity at stations P, R, V and S is enough to allow two trains to simultaneously stop. Two possible solutions are possible at the bottleneck. If train A is first, train B is held back and faces some delay, which we represent in the time distance path as a parallelogram (blue, stretching between times 4 and 6) depicting the area in time space where a train is blocking the infrastructure. For the sake of simplicity and clarity, we do not report formally blocking times, nor section boundaries, but all the reasoning applies. This delay results in a delay for train B, which is further propagated as train C goes after train B (another parallelogram, this time in red color, between time 8 and 14). Overall, passenger X travels on train A till station V where a transfer is possible to train C arriving at station P at time 19. This solution is reported in Figure 2 (middle), and Figure 3 (top row).

**Fig. 2** Non convergence to an equilibrium point. (top): infrastructure and services; (middle): first case; (bottom) second case.
Otherwise, the opposite order is possible, namely that train B goes first and train A is delayed. This solution is reported in Figure 2 (bottom), and Figure 3 (bottom row). In this case, train C is not delayed anymore by a delayed train B, and thus arrives at its planned time, i.e. 16, which is one time unit before train D. The summary of the delays of trains at the respective stations is reported in Table 2, top layer (both left or right): For train C we reported both delay at station (V) and at station (S).

Let us look at how the passengers would react to those possible choices, starting from the former one, i.e. A goes first. This solution is the one minimizing the maximum consecutive delay, the average consecutive delay, and their weighted versions, given the passenger flows. We can also assume that the delay Y is smaller than 2.5, then this same solution also minimizes the maximum total delay, and also the average total delay. All trains have some passengers, which are assumed to be \( \varepsilon_A, \varepsilon_B, \varepsilon_C, \varepsilon_D \) per train; those can be supposed for instance to be very small, and in any case they are not influenced by the different train rescheduling actions. In fact, if the origin destination of each train correspond to those of the passengers onboard; there are no two trains, which are directly competing on the same OD pair. Moreover, due to infinite capacity, different passenger do not directly influence each other. The only relevant passenger is X. Confronted to the solution on Figure 2 (middle), the best choice for passenger X would be to board train B and then change for train D, for an arrival time of 17 time units (Figure 3 top right). Instead, boarding train A would mean needing to change to the delayed train C, for an arrival time of 19 (Figure 3 top left).

**Fig. 3 Flows and the two possible solutions**

In the latter choice of the IM player, i.e. train B goes first, the passenger would instead board train A and then change to train C, for an arrival time of 16. The opposite
choice of the passenger, taking train B and then changing to train D, would arrive only at 17, so one time unit later.

We remark here that the three Nash setups would not differ: the amount of people onboard the train, or disembarking, (respectively used by Nash 1 and Nash 2) are the same (In Table 1, the third column of each table). Finally, Nash 3 would focus on connections, which actually do not change between those two solutions, thus no further action or difference would arise from Nash 3.

In fact, the optimal reaction of each player to the best move of the other one would result in not having a Nash equilibrium in pure strategies. This is summarized in the payoff matrix reported in Table 2, summarizing for the two possible actions of the Passengers and the IM, the utility outcomes for the passenger and the IM. Given the schedule of the train, Passenger X can in fact use a different route which results in a shorter travel time. Namely, passenger X can board train B, suffer the delay from train A, and change to train D at station R, which would arrive at 17, i.e. two time units before the (delayed) train C. Thus a different passenger assignment is considered, where Passenger X takes train B and D.

But, given this updated passenger assignment, it is better to invert the order of trains at the bottleneck. Then again, given this schedule, passenger X then changes route, taking train A, and C in succession, which is the starting solution reported in Figure 2 (middle) and discussed here above.

<table>
<thead>
<tr>
<th>Train (stat)</th>
<th>Total/Cons</th>
<th>Flow/grad</th>
<th>Train (stat)</th>
<th>Total/Cons</th>
<th>Flow/grad</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Y</td>
<td>ε_A+1</td>
<td>A</td>
<td>Y</td>
<td>ε_A</td>
</tr>
<tr>
<td>B</td>
<td>2.5</td>
<td>ε_B</td>
<td>B</td>
<td>2.5</td>
<td>ε_B+1</td>
</tr>
<tr>
<td>C (V)</td>
<td>3</td>
<td>ε_C</td>
<td>C (V)</td>
<td>3</td>
<td>ε_C</td>
</tr>
<tr>
<td>C (S)</td>
<td>3</td>
<td>ε_C+1</td>
<td>C (S)</td>
<td>3</td>
<td>ε_C</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>ε_D</td>
<td>D</td>
<td>0</td>
<td>ε_D+1</td>
</tr>
</tbody>
</table>

Table 1. Delays and passenger flows corresponding to the solutions of Fig 3.

<table>
<thead>
<tr>
<th>Train (stat)</th>
<th>Total/Cons</th>
<th>Flow/grad</th>
<th>Train (stat)</th>
<th>Total/Cons</th>
<th>Flow/grad</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Y 3.5</td>
<td>ε_A+1</td>
<td>A</td>
<td>Y 3.5</td>
<td>ε_A</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>ε_B</td>
<td>B</td>
<td>0</td>
<td>ε_B+1</td>
</tr>
<tr>
<td>C (V)</td>
<td>0</td>
<td>ε_C</td>
<td>C (V)</td>
<td>0</td>
<td>ε_C</td>
</tr>
<tr>
<td>C (S)</td>
<td>0</td>
<td>ε_C+1</td>
<td>C (S)</td>
<td>0</td>
<td>ε_C</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>ε_D</td>
<td>D</td>
<td>0</td>
<td>ε_D+1</td>
</tr>
</tbody>
</table>

Table 2. Payoff matrix
3 Experiments and results

This section reports on the results of a campaign of experiments on the performance evaluation of the 5 solution schemes described in Section 2 (Nash 1, 2, 3, Timetable, Train Rescheduling Only). Extensive computational experiments on a set of disturbed traffic situations demonstrate the potential of the proposed approach for increasing the passenger satisfaction, the quality of railway service and the overall traffic performance. All experiments are executed on a normal computer, equipped with a processor Intel i5 CPU at 2.70 GHz, 8 GB memory. The commercial solver CPLEX 12.4 is used to solve the MILP formulations.

The same instances of Corman et al 2016 are used, which refer (on average) to 22 OD pairs, 150 ODW passenger groups, 45 minutes of traffic prediction, order of magnitude of 4000 nodes N and 40000 ordering decision/arcs S, 156 trains, an average initial delay of the trains 120 seconds, and average travel time of passengers of 1830 seconds.

We report in Figure 4 the tradeoff between the interest of the two decision makers. More precisely, we refer to a bidimensional surface, the x-axis being the average passenger travel time, (i.e. the objective of the passengers and TOC), and the y-axis being the average consecutive train delay, (i.e. the objective of IM). For both axes, we report the percentage gap with regard to a benchmark, which is the Nash 3, online advice solution, which corresponds to the best performing heuristic H4 in (Corman et al 2016). Each solution scheme results in a point.

![Figure 4](image-url)  
**Figure 4.** Tradeoff between passengers travel time (passengers) and consecutive delay (infrastructure manager) objectives.
From the plot, many interesting conclusions are available. The timetable solution scheme performs very bad in terms of train delay, being about one order of magnitude greater than the other approaches. The performance in terms of passenger travel time is also quite poor. Nash 1 and Nash 2 perform in a very similar manner. We also remark how both approaches deliver always a train delays below the benchmark, for all advice configurations. Nash 3 performs remarkably different, instead. We comment briefly on the amount of iterations required for convergence. 3.3 for Nash1 and Nash 3 slightly higher at 3.5 for Nash2. Timetable and train rescheduling need a single iteration.

We report in Figure 5 a graphic representation of the iterative behavior of the solution, in the same plane defined by the x axis as train delay and y axis as passenger travel time, as Figure 4, for the Nash approaches In Figure 5, the x and y axis have been scaled to report all final solutions for the 80 cases (i.e. the Equilibrium) at the reference point (100, 100). For every case, the intermediate solutions have been reported as a line, connecting the values of train delays and passenger travel time at the iterations. Moreover, the first iterations are in light blue, and the color becomes darker as the iteration increase. In other words, there are thus 80 sequences of line, which start as light blue somewhere in the plane, connect all intermediate points related to the iterations of the solution scheme, become darker at every iteration, and ultimately reach the point (100,100) corresponding to the final solution. The three plots in Figure 6 report respectively the Nash 1, Nash 2 and Nash 3 solution.

From a graphic analysis of the Figure 5, some trends are evident. Again, the variation in train delays is relatively much larger than the variation in passenger travel time, the former being easily within a 50% variation, while the latter rarely exhibiting a variation larger than 5%.

The first Iteration of Nash 1 is generally smaller in size than the first iteration of Nash 2 and Nash 3 (the light blue part of the figure is smaller). In fact, Nash 1 achieves a rather limited spread of solution values compared to the final one, rarely getting out of a range of [98-102] for passenger travel time and [50-110] for train delay. Nash 2 has generally a larger spread of the first iteration, which is mostly associated to much larger train delays and slightly larger passenger travel times. The iterations afterwards are generally quite limited. Nash 3 has a remarkably different behavior, with no point achieving a passenger travel time smaller than the final solution (i.e. the semi-plane left of the (100,100) point is empty). The first step, as well as the following ones, is generally much larger than the two other solution schemes. In general, the first iteration reduces train delay and passenger travel time, while the successive iterations reduce passenger travel time but increase train delay, by optimizing some connection to be kept.

We here comment that a larger spread in the space of the two objective function is an indicator of some instability and a strong trade-off between the two objectives, while a compact spread is associated to a robust process where the interaction between the decision makers is more limited. The latter case is also associated to cases where the reduction of delays and passenger travel time is more limited in general.
4 Conclusions

This paper integrates train rescheduling and delay management approaches into a series of mathematical models to control railway traffic in real-time (Train Rescheduling, TR) with the objective of minimizing passenger travel time (directly tackled by Delay Management, DM). This paper addresses the trade-off and the strategic interaction between the objectives of the above-mentioned stakeholders, in different game-theoretical settings.

Based on this model and different approaches to specify the interaction between the players, we discuss general existence of Nash equilibria in pure strategies, and compute Nash equilibria of the resulting game. Computational experiments on a real-life network with a large set of OD pairs show that the two objectives are indeed competitive, though one objective has much more variation than the other; and we also show how the different setup of updates and reaction to a move of the other players influences quality and convergence of the overall procedure.

Several possible directions are open for future research, starting from the obtained results. One could study other ways to reach the Nash equilibria, and possibly determine all of them, if they would exist. Furthermore, it would be interesting to compare the Nash equilibria with the optimal solution of the individual objectives of the IM and the TOC. The transit assignment problem with capacity on vehicle is already considered leading to strategic interaction among travelers to board the limited capacity. In theory, including capacity of vehicles in the MDM model might require considering more sophisticated models for the passengers’ behavior.

References


