Assessing disruption management strategies in rail-bound urban public transport systems from a passenger perspective

Roelofsen, D.S. · Cats, O. · van Oort, N. · Hoogendoorn, S. P.

Abstract: This paper provides a framework for generating and assessing alternatives in case of disruptions in rail-bound urban public transport systems. The proposed framework considers the passenger perspective as well as the operator perspective, for the often-used measures of detouring and short-turning. An application of the framework demonstrates that currently used disruption management protocols often do not lead to the optimal solution from the passenger perspective. Furthermore, the optimal choice between alternatives from passenger perspective shows to be dependent on the passenger flows.

Keywords: disturbances; disruption management; passenger perspective; rail-bound urban public transport; smartcard data
1 Control of public transport systems

Service reliability of public transport systems is widely considered as one of the most important service aspects in the evaluation of its quality (Redman, Friman, Gärling & Hartig, 2013). Disruptions in operations highly affect the reliability of public transport systems by late, early or no arrival of vehicles, leading to prolonged waiting times and thus longer travel times (Van Oort, 2016). Furthermore, re-routing of vehicles (usually) causes longer trip times and possibly skipped stops, while cancellation of services causes crowded and thus less comfortable rides.

Previous research showed that current disruption management protocols are primarily focused on the operator perspective (Carrel, Mishalani, Wilson & Attanucci, 2013; Cats 2014). Babany (2015) considers the passenger perspective when restoring the schedule; that is when the cause of the disruption has been solved but operations are still disrupted. This phase is coined the third recovery phase in the bathtub model (Ghaemi et al. 2017). It remains however unknown how the passenger perspective can be considered in the first transition phase and the second disruption phase, when the cause of disruption has not been removed yet.

There has been research conducted regarding the passenger perspective of train operations plans and traffic control in heavy-rail environments. For example, Takeuchi et. al (2017) propose a framework which evaluates train operations plans from the passenger perspective (frequency, convenience and reliability), operational costs and energy consumption. D’Ariano and Corman (2017) formulate an optimization model which minimizes train delay and passenger travel times during real-time railway traffic control. However, rail-bound urban public transport systems differ from heavy rail operations, in for example the probability of disruptions to occur (disturbances in rail-bound urban public transport is more likely to occur than in heavy-rail (Cats et al., 2016b)), implemented measures by traffic controllers (rerouting vehicles are often used in rail-bound urban public transport, whereas this is not the case in heavy-rail), network density and redundancy (Jenelius and Cats 2015), and walking as an alternative for passengers.

Hitherto, no comprehensive framework is available which explicitly considers passenger impacts of disruptions and implemented measures, while also considering the operator perspective, in order to support making a well-informed decision on managing operations during the disruption period. The objective of this research is to provide a framework to be used in disrupted operations in rail-bound urban public transport system, which generates and assesses alternatives while accounting for both passenger and operator perspectives. Two commonly used measures by traffic controllers have been investigated, namely detouring and short-turning (see figure 1).
2 Framework for generating and assessing disruption management alternatives

The proposed framework consists of two parts: the generation of routing alternatives and the assessment of alternatives (see figure 2).
2.1 Generation of alternatives

For the generation model, the $k$-shortest path algorithm is used (Yen, 1971). This results in the set of all possible detours. Using the following series of filtering rules, a subset of detours to be further assessed is attained:

- A threshold value excludes detours exceeding a certain extra travel time compared to the original route.
- A threshold value excluding detours directly affecting a certain number of additional passengers compared to the alternative directly affecting the least number of passengers.
- Dominancy aspect, excluding detours which skip (at least) the same stops as another detour while yielding a longer travel time.
- Enforcing link capacity constraints. The effect of the increased frequency on the detour route is assessed. The dominancy aspect and the residual capacity of detour links are considered iteratively.

Besides detour alternatives, short-turn alternatives are also considered. These alternatives are generated by comparing the short-turn possibilities in the network. The set of possible short-turning locations for a given disruption is generated as a subset of short-turning locations which is given as an input.

2.2 Assessment of alternatives

**Passenger perspective**

After the alternatives are generated, they are assessed from both perspectives. To determine the consequences for the total generalized passenger time (TGTT), historical data of passenger flows (derived from smartcard data) is used as model input. Detailed information regarding these smartcard data is provided by Van Oort et. al (2016). The passenger impact of the different alternatives is determined using several assumptions regarding passenger route choice when encountering a disruption:

- Passengers for which their original boarding stop and/or alighting stop is not skipped by the alternative can continue using that stop.
- Passengers for which their original boarding stop and/or alighting stop is skipped are assumed to either walk directly to their destination, walk to the closest stop that is still served or wait until the disruption is over and service is resumed. The choice whether to walk or to wait is modelled by a probability distribution function conditional on the walking distance.

Given these assumptions, passenger travel times associated with each alternative are estimated. The passenger perspective is reflected by the (generalized) total passenger
travel time of an alternative, which depends on the total number of passengers affected and the route of the alternative (and corresponding skipped stops).

The origin stop \( i \) and the destination stop \( j \) can be located at three sections of the line; upstream of the disruption \( (i, j \in S_{l,p,q}) \), at the skipped section of the line due to the disruption \( (i, j \in S_{l,p,p}) \) or downstream of the disruption \( (i, j \in S_{l,p,v}) \). Depending on the location of both the origin \( i \) and destination \( j \), the passenger travel time for a route \( p \) between origin and destination \( t_{i,j,p}^l \) can be determined. Fig. illustrates the definitions of the abovementioned sets used to denote the location of the stops with respect to the location of the disruption.

Fig. 3 Illustration of used terminology for stops in the relation of an original route (solid line) and a detour (dotted line).

The travel time between an origin \( i \) and a destination \( j \) during a disruption is dependent on several factors. First, it depends on the location of origin \( i \) in respect to the applied alternative. Depending on the location of the origin stop, passengers have to wait at the origin stop or walk to a stop served. For a route \( p \), the origin stop can be located upstream of the disruption \( (i \in S_{l,p,q}) \), downstream of the disruption \( (i \in S_{l,p,v}) \) or can be skipped \( (i \in S_{l,p,p}) \).

Second, passenger travel time depends on the location of destination \( j \). Just as for the origin stop, for a route \( p \) the destination stop can be located upstream, downstream of the disruptions, or skipped. However, destination stops cannot be located prior to the origin stop; for instance, it is not possible for a destination stop to be located upstream of the disruption if the origin is located downstream of the disruption.

Third, passenger travel time can be dependent on the selected mitigation measure. Whether short-turning or detouring is considered, can affect the passenger travel time elements. For instance, for detouring it is possible to board upstream of the disruption and alight downstream of the disruption, whereas for short-turning some distance should be crossed by foot.

Finally, passenger travel time depends on the previously discussed passenger path choice. Depending on the path choice, the passenger travel time consists for instance of directly walking from the origin to the destination, walking to the last stop upstream
and boarding a vehicle there, or walking to the first stop served downstream of the disruption \((S_{l,p,r_1})\) and boarding a vehicle there.

The following trip components are considered: walking time, waiting time, transfers and in-vehicle time. All of which in relation to the undisrupted case. The different passenger trip components are perceived differently by passengers. Therefore, each element is assigned a corresponding weight, in order to capture the difference in perception. Transfers are penalized by a fixed penalty, since ceteris paribus direct trips are preferred over trips with a transfer. Due to the high frequency that is typical for rail-bound urban public transport, it is assumed that passengers arrive randomly at the stop.

An example of the generalized travel time \((Gt_t)\) between on origin downstream of the disruption and a destination upstream of the disruption without a detour is then:

\[
G_{t,i,j,p} = \beta \cdot t_{i,p}^{wait} + \beta^{int} \cdot t_{i,j,q}^{int} + \beta \cdot t_{j,q}^{walk} \cdot t_{r_t,1,p}^{wait} + \beta \cdot t_{r_t,r_1,p}^{wait} + \beta \cdot t_{r_t}^{int} + \beta \cdot \#_{transfers} \tag{1}
\]

The different weight factors for the different trip elements are denoted by \(\beta\). A transfer occurs when there are multiple legs in a trip. This is for instance the case when a passenger boards a vehicle between \(i\) and \(|q|\), and then walks from \(|q|\) to \(j\). An additional transfer would be made if he walks from \(|q|\) to \(r_1\) and boards a second vehicle there.

The different alternative routes are to be ranked from a passenger perspective based on their total generalized passenger travel time. To calculate the total generalized passenger travel time for a route \(p\), \(G_{t,p}^{tot}\) during a time interval \(\tau\), being element of a set of time intervals \(T\), the travel time between each \(i\) and each \(j\) is multiplied with the corresponding demand for that time interval \(\pi_{i,j,\tau}\), and summed over all \(i,j \in S_t\).

\[
G_{t,p}^{tot} = \sum_{i \in S_t} \sum_{j \in S_t} G_{t,i,j,p} \cdot \pi_{i,j,\tau} \forall \tau \in T, p \in P_t \tag{2}
\]

The number of passengers travelling from \(i\) to \(j\) varies over time. The travel time between \(i\) to \(j\) is assumed fixed within the analysis period. The total generalized passenger travel time over the whole set of time intervals \(T\) is then:

\[
G_{t,p}^{tot} = \sum_{\tau \in T} \sum_{i \in S_t} \sum_{j \in S_t} G_{t,i,j,p} \cdot \pi_{i,j,\tau} \forall p \in P_t \tag{3}
\]

An important aspect in the evaluation of passengers of public transport is the issue of crowding. In-vehicle crowding has not been taken into account in this study, because only effects on the disrupted line are taken into account. Effects on other lines are excluded, and taking in-vehicle boarding into account on the disrupted line only could lead to alternatives being underestimated which do not cause in-vehicle crowding on the disrupted line, but do so on other lines; in other words, in order to consider in-
vehicle crowding explicitly, a network-wide assessment is necessary. However, denied boarding on the disrupted line are considered. Denied boarding are perceived very badly by passengers. This is taken into account by assigning an extra penalty to the waiting time of denied boarding passengers. So, until arrival of the first vehicle the regular waiting time weight factor is applied, and if denied boarding occurs the waiting time from the arrival of the first vehicle until arrival of the second vehicle, the weight factor is multiplied with the denied boarding factor (Cats et al. 2016a).

Passengers on-board during disruption

The previously discussed travel times during disrupted operations are valid for passengers starting their journey during the disruption. Passengers who are already on-board the vehicle when the disruption occurs are also affected by it, for instance due to their stop not being served anymore, or a longer in-vehicle travel time due to a detour. Since the trip elements prior to the start of the disruption are not affected by the possible alternatives, these are not considered while analysing the alternatives from the passenger perspective. In other words, for passengers already on-board the vehicle, the waiting time at the origin is neglected, as well as the in-vehicle time up to the start of the disruption. From the start of the disruption the different trip element travel times are calculated in the same manner as has been discussed before, since these can be affected by the chosen alternative.

Resource perspective

Disruptions do not only affect the demand side of the network (passengers) by leading to different travel times and routes than expected, but can also affect the supply side (resources) of the network. As discussed previously, resources are planned minutely, defining their activity and location for all times. The goal in the planning process is, given a certain timetable, to minimize the number of resources used.

The resource perspective of the different alternatives is given here without intervention in the schedules, i.e. rescheduling personnel and vehicles. The possibilities to reschedule by switching driver shift, changing vehicles, or the usage of spare drivers/vehicles are so extensive and dependent on the actual situations, that providing a generic rescheduling approach for resources for the different alternatives is considered out of the scope. Therefore, the consequences from a resource perspective of the different alternatives without rescheduling are assessed here.

For the alternatives incorporating a detour, the resource schedule and timetable is likely to be affected. The effect of the detour on the timetable has been discussed when assessing the alternatives from a passenger perspective.

The resource schedule is significantly affected if the delay caused by a detour propagates onto the next scheduled activity, which is the case if it arrives delayed at the destination terminal and the delay cannot be compensated for by any buffer times.
Buffer times can be incorporated in the schedule when the time scheduled for an activity is longer than the actual time needed. If it arrives earlier than planned at the destination terminal, i.e. the detour was effectively a shortcut, the planned activities can follow through as planned for the remainder of the duty. The same accounts for a detour arriving at the destination terminal at the same time as has been planned.

A delay arises on a disrupted line if the route trip time due to a detour is longer than the route trip time of the original line. The amount of delay of a route $p$ on a line $l$ $t_{l,p}^d$ is the difference between the scheduled route trip time and the route trip time of the detour. Or formally:

$$t_{l,p}^d = t_{l,p}^r - t_{l,p}^r$$

(4)

Note that the delay can have a negative value as well, which indicates that the detour is actually a shortcut. Only delays caused by the different routes $p$ are considered.

With the amount of delay $t_{l,p}^d$, the delayed arrival time at the destination terminal of the resource can be determined. In order to assess the effective delay for the subsequent activity it needs to be compensated with the buffer time. For any delay of an activity $k \in K$ (set of all activities scheduled), the delay of the subsequent activity $k+1$ can be determined by subtracting the original delay with the buffer time $t_k^b$.

Formally, this can be denoted as follows:

$$t_{k+1}^d = t_k^d - t_k^b \forall k \in K$$

(5)

Subsequent planned activities can be of any kind, such as the return trip for both driver as well as vehicle, a shift on another line for just the driver or vehicle, a break for the driver, end-of-shift for the driver, scheduled maintenance for a vehicle, etc. Please note that a driver is not necessarily coupled to a vehicle, and they should thus be seen as two separate resources.

The consequences of the delay can range in severity based on the subsequent activity scheduled. For instance, if it is the last shift, the consequence is a later end-time of the shift. The delay is cleared and other than a longer shift time there are no consequences. The longer shift time can have consequences, such as violation of legal regulations, overtime, or the later end-time of the shift shortening the time in-between shifts. Delays can also propagate onto next shifts, which can among other vary in terms of passenger demand.

In order to take into account that the effect of a delay can range in severity based on the activity, delays are weighted. Weights are assigned a low value if the severity of the delayed activity is relatively low, such as end of duty time or a shift on a line with low passenger demand, and will be assigned a high value if the severity is high, such
as a line with high passenger demand. Formally, this is denoted as follows, with \( \beta_k \) indicating the weight factor and \( Gt^d_k \) indicating the weighted delay of an activity \( k \):

\[
Gt^d_k = t^d_k \cdot \beta_k \quad \forall k \in K
\]  

(6)

The total of weighted delays \( Gt^d \) for all activities \( k \) is then represented by:

\[
Gt^d = \sum_{k \in K} t^d_k \cdot \beta_k
\]  

(7)

Since the logistics of vehicles and drivers might be affected fundamentally by short turning (vehicles will not reach the final terminal), we used expert judgment of operators and (strategic) planners for this aspect.

Summarizing, the generated alternatives are assessed from the passenger perspective as well as the resource perspective, using the indicators presented in table 1.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Indicators in assessing generated alternatives.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Detour</td>
</tr>
<tr>
<td>Passenger perspective</td>
<td>Total generalized passenger time (TGTT) [minutes]</td>
</tr>
<tr>
<td>Operator perspective</td>
<td>Delay at destination terminal [minutes]</td>
</tr>
</tbody>
</table>

The methodology is implemented in the stochastic event-based simulator Simio. This is perceived to be suited for mimicking the disrupted operation and passenger behavior prior to, during and after the disruption. Each passenger is generated as a separate entity, based on random (from Poisson distribution) arrival for a given arrival rate of passengers per hour. Arrival rates are retrieved from historical data. Events in this context range from arrival of passengers at a stop, the arrival of a vehicle at a stop or the start of a disruption. In this manner, the disruption operation is mimicked as well as the passenger behaviour prior to, during, and after the disruption.

3 Case studies and results

The developed framework is applied by means of discrete event-based simulation to four (hypothetical) disruption locations in the HTM urban rail network in The Hague, as well as one historical disruption.

3.1 Rail-bound urban public transport in The Hague

The city of The Hague is the political centre of the Netherlands and with approximately 515,000 inhabitants it is ranked as the third city in the Netherlands according to the population, with a population over 1,000,000 in its agglomeration. The rail-bound urban public transport system is operated by HTM, transporting an average daily of
over 250,000, using 115 buses and 219 trams and light-rail vehicles. With a total network length of 336 kilometre, rail-bound disruptions for which disruption management is applied, occur daily.

3.2 Case study disruption locations

In determining the case study disruption locations, two main criteria have been used:

- No reasonable alternatives for passenger path choice are available since historical passenger flow data is used. Considering this limitation, a disruption duration of one hour is assumed.
- On the other hand, in order to demonstrate the frameworks’ face validity, different non-obvious alternatives should be available for the rerouting of vehicles.

The case studies differed in disruption locations in the network, along with the set of alternatives available and passenger flow patterns. To examine the effect of different passenger demand levels, for each of the four locations, two different passenger demand levels were taken into account (morning-peak and rest-of-day). To give an idea of the generated and assessed choice-sets, Figure 3 shows one of the disruption locations (marked by X), as well as the original route (red), different detouring alternatives and short-turning alternatives.

![Fig. 4 Disruption location and alternative choice-set for disruption location A.](image-url)
3.3 Case study results

For all four locations, 7 to 9 disruption management alternatives are generated by the model. These alternatives have been assessed from the passenger perspective as well as from the resource perspective, for both morning-peak- and rest-of-day passenger demand levels. Additionally, the assessed alternatives have been compared to the current disruption management protocols. Table 2 presents the difference between the detour alternative resulting in the lowest TGTT and the short-turn alternative resulting in the lowest TGTT, for all locations. A positive value indicates that the TGTT is lower when applying short turning. It also shows the difference in extra TGTT incurred by the alternative with the lowest TGTT, as compared to the current protocol.

<table>
<thead>
<tr>
<th></th>
<th>Difference detouring vs. short-turning</th>
<th>Potential savings extra TGTT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Morning-peak</td>
<td>Rest-of-day</td>
</tr>
<tr>
<td>Location A</td>
<td>- 6%</td>
<td>+ 28%</td>
</tr>
<tr>
<td>Location B</td>
<td>+ 15%</td>
<td>- 2%</td>
</tr>
<tr>
<td>Location C</td>
<td>- 29%</td>
<td>- 40%</td>
</tr>
<tr>
<td>Location D</td>
<td>- 80%</td>
<td>- 64%</td>
</tr>
</tbody>
</table>

The results show that different locations, implying different passenger demand levels and patterns and re-route alternatives, affect the outcome in terms of the disruption management strategy resulting in the lowest TGTT. Furthermore, it shows that the proposed strategy outperforms the current disruption management protocols resulting with a significantly lower TGTT with the exception of one location and demand level where no improvement is made. It should be noted that these potential savings in extra TGTT only occur on the disrupted line, and second-order effects of for instance resource delay on subsequent activities are not captured (e.g. induced crowding). Further analysis of the current disruption management protocols showed these are mainly driven by the operator perspective, only minimizing resource delay. All current disruption management protocols showed to be part of the Pareto-front between the TGTT on the one hand and the delay at the destination terminal on the other. This means that no other alternative is better on both the passenger perspective and the operator perspective at the same time than the current disruption management protocol.
4 Conclusions and discussion

In this research we developed a framework to generate alternatives in case of disruptions in rail-bound urban public transport, and furthermore assess these alternatives from the passenger perspective as well as the resource perspective. The framework is applied to several cases. Based on these case studies, practical recommendations are derived, as well as an outlook to further studies.

4.1 Practical recommendations

The case studies showed that, from a passenger perspective, only in one out of eight cases (consisting of 4 locations, 2 passenger demand levels per case), the current disruption management protocol yielded the least impedance for passengers on the disrupted line.

Based on the characteristics of the different locations and the outcomes in terms of passenger impact on the disrupted line, three variables have been identified which are of key importance when considering detouring or short-turning in the management of disruptions: (i) the ratio between passengers benefiting from detouring versus passengers benefiting from short-turning; (ii) the distance between the two short-turning stops, and; (iii) the detour length. Based on the values of these variables, the following decision tree can provide the favourable alternative from a passenger perspective (figure 4).

![Decision-tree indicating favourable alternative from a passenger perspective (ST = short-turning).](image)

The passenger groups most likely to benefit from different disruption management strategies are illustrated in figure 4. Passengers favoured by detouring are those originating from stops in group 1 and destined for stops in group 4. Passengers favoured by short-turning are destined for stops in group 2 or originating from stops...
in group 3. For passengers originating from stops in group 2 or passengers destined for stops in group 3, it depends on the walking distance between their stop and the last / first stop served upstream / downstream of the disruption. The longer the distance, the more favourable short-turning is. The favourable measure depends on the composition of these passenger demand groups.

Fig. 6 Different stops in relation to detouring and short-turning, and the favourable alternative depending on the OD-relation (blue = detour, red = short-turn, black = depending on distance).

To provide some context, a high ratio of passengers between detouring and short-turning turned out to be 2.5 passengers favoured by detouring versus 1 passenger favoured by short-turning. A low short-turning length was a walking distance of 5 minutes between the two short-turning stops, whereas a high length was 12 minutes. For the detour length, a low detour length was 0 minutes while the high detour length induced 11.5 additional minutes.

4.2 Outlook

In its current form, the framework is applicable for a conditional set of locations as well as disruption durations. The main recommendation is extending the provided methodology by incorporating a passenger route choice model to allow for the redistribution of passenger demand by embedding a dynamic transit assignment model in the disruption management framework (see Cats 2016). This will enable conducting a network-wide assessment of the different alternatives, making the method suitable for analysing any disruption location.

Acknowledgements: This research is performed in cooperation with HTM and Delft University of Technology, Department of Transport & Planning. The authors thank HTM, the urban public transport operator of The Hague, the Netherlands, for their valuable cooperation and data provision. This research was performed as part of the TRANS-FORM (Smart transfers through unravelling urban form and travel flow dynamics) project funded by NWO grant
agreement 438.15.404/298 as part of JPI Urban Europe ERA-NET CoFound Smart Cities and Communities initiative.

References


