Network Reduction and Dynamic Forecasting of Passenger Flows for Disruption Management

Evelien van der Hurk, Leo Kroon, Gábor Maróti, Paul Bouman, Peter Vervest

Department of Decision and Information Sciences, Rotterdam School of Management
Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands,
e-mail: ehurk@rsm.nl
Innovation and Process Quality, Netherlands Railways
Laan van Puntenburg 100, Utrecht, The Netherlands

Abstract

Current policies in disruption management for railways mainly focus on solving the resulting rescheduling problems for crews and rolling stock. Maximizing passenger service level as objective will likely improve the solution for the passengers. Realized passenger flows emerge from a complex interaction between the timetable, the vehicle schedules and the passengers' behavior. The required detailed data on passenger journeys became available only recently as resulting from smart card fare ticketing systems. This paper studies modeling and forecasting passenger journeys based on this newly available detailed data for disruption management in public transport networks.

A two-step methodology consisting of Network Reduction and forecasting leads to time and origin destination specific forecasts of passenger journeys. Immediate insight into how passengers are affected by a disruption results from the Network Reduction. Real-time forecasts of passenger behavior are derived based on historic information from smart card data. Together this provides insight in how passengers are affected by a disruption, enables a forecast of demand for seats given a route choice decision, and allows to estimate passenger delay given any timetabling decision. Results are based on a 10 months’ real-life smart card data set of the Dutch passenger railway operator Netherlands Railways (NS).

Keywords

Passengers, Disruption Management, Forecasts, Econometrics

1 Introduction

In our current time pressured society, reliability of public transport is very important. Unfortunately disruptions occur regularly in these systems. Passengers expect quick solutions that minimize their inconvenience, while disruption management policies in practice are focused on the logistic aspects as crew and rolling stock rescheduling, which will not always lead to solutions that are best for passengers. A focus on passenger service requires detailed data on demand. Recently introduced smart card systems collect data on passenger jour-
neys. How to forecast demand and estimate the impact of a disruption on passengers using this detailed data is the focus of this paper.

Disruption management aims at adjusting the pre-set schedules to the unexpected circumstances while maintaining passenger service. Current disruption management practice largely focuses on the timetable as well as on vehicle and crew re-scheduling. Disruption management is carried out under serious time pressure, and must be able to deal with the uncertain duration of the disruption. The resulting real-time re-scheduling problems quickly become rather complex, as shown by Potthoff (2010) for crews and by Nielsen (2011) for railway rolling stock.

Passenger service level is the objective of recent research in this area. Nielsen (2011) takes into account passenger flows in case of rolling stock rescheduling and aims to minimize passenger delays. Dollevoet et al. (2012) include rerouting passengers in delay management. However, the required input is generally not available in real life.

Forecasting passenger demand is not straightforward despite the available data. Realized passenger flows emerge from a complex interaction between the timetable, the vehicle schedules and the passengers’ behavior. The timetable specifies the departure and arrival times of the services, while the vehicle schedules determine the available capacities. The passengers react to the disruption by choosing an alternative route, mostly by considering the anticipated travel time. The vehicles’ capacity may not be sufficient to accommodate the demand; in such case the passengers must compete for space. Passengers that are left behind have then to re-consider their travel options, and choose another route through the network. Consequently, the per-trip passenger demand on a disrupted day may substantially differ from that on a normal, undisrupted day.

How to forecast passenger demand for disruption management is the focus of this paper. Forecasts should be detailed enough to study the interaction between passengers, the timetable and the vehicle schedule. A second requirement is to minimize the dimension of the origin destination matrix as rescheduling problems grow in complexity with this dimension. Finally, forecasts should be easy to analyze in real-time to be useful in practice.

These requirements make the forecast of passenger flows for disruption management significantly different from the origin destination (OD) matrix forecasts in traffic, that forecast for a given set of origins and destinations information in given time intervals. For one because time cannot simply be aggregated in globally fixed time intervals as this is too coarse to assign passengers to trains. Furthermore, the focus in disruption management is on changes in demand and passengers affected by the disruption instead of overall demand. Finally, reducing the dimension of the OD matrix benefits both practical analysis and rescheduling models, and is possible thanks to the focus on changes in demand and the structure of the public transport network.

We propose a two-step methodology for forecasting demand. In the first step, a Network Reduction selects journeys affected by the disruption and groups journeys based on their detour routes. This significantly reduces the number of origin, destination, time combinations while maintaining enough detail for estimates of passengers per train, even in case of complex interactions. The second step links smart card data to specific journeys and paths through the network. Combining this with the Network Reduction, we derive dynamic forecasts of passengers affected by the disruption. Combining this with any behavioral as-
sumption of passengers and a logistic schedule, we can estimate the number of passengers per train and the delays of the passengers.

The forecast model is tested on a real-life smart card data set of Netherlands Railways (NS), the largest passenger railway operator in the Netherlands. The forecasts can be used both by computer-aided disruption management tools and by the dispatchers themselves to improve the service level during and after the disruption and to obtain insight into how passengers are affected by a disruption. Our future work on disruption management applications builds upon the detailed demand forecast. In particular, we are interested in using the forecast in capacity rescheduling and in providing travel advice to passengers with the objective of elevating the achieved service quality.

The remainder of this paper is organized as follows. In Section 2 we present the problem description. Section 3 contains a short literature overview of related research on forecasting, smart card data and disruption management. The technique for Network Reduction and forecasting is described in Section 4. Results based on smart card data of NS are presented in Section 5. We conclude this paper with a discussion and comments on future research in Section 6.

2 Problem description

In this section we focus on how passengers react to a disruption. We illustrate the concepts with an example drawn from the Dutch railway network. Thereafter we sketch our proposed two-step methodology.

2.1 The disruption: line-based and network perspectives

The Dutch railway network is composed of lines; a line provides direct connections between certain stations, and is operated with a given frequency, usually once every 30 minutes. Figure 1a shows the so-called ‘800’ line connecting Amsterdam (Asd) to Maastricht (Mt) via Den Bosch (Ht) and Eindhoven (Ehv).

Suppose now that the rail segment between Den Bosch and Eindhoven becomes unavailable. From a vehicle scheduling point of view, the natural solution is to let the train run on the remaining infrastructure: between Amsterdam and Den Bosch as well as between Eindhoven and Maastricht, as shown by the grey bending arrows in Figure 1a.

Passengers, however, still need to travel from one side of the disrupted segment to another. Fortunately, there are other train lines, as well. In particular, the ‘1900’ line passes through Eindhoven and Tilburg (Tb), and the ‘16000’ line passes through Den Bosch and Tilburg (see Figure 1a). These two additional lines happen to provide the geographically shortest possible route from Den Bosch to Eindhoven and back. Therefore, in practice, the operator assumes all passengers will use this detour.

The network actually contains many more lines than the three ones mentioned so far. A passenger who wants to travel from Amsterdam to Eindhoven may also choose to travel to Rotterdam (Rtd) and to transfer to a direct train to Eindhoven. This alternative is geographi-
Operator reaction to disruption | Network perspective on disruption

(a) Disruption and rerouting advice to passengers | (b) Early warning leads to more and shorter reroutes

Figure 1: A disruption on the 800 line between Eindhoven and Den Bosch requires passengers to reroute. The line based perspective (top) results in different detours of passengers than the network perspective (bottom), which better spreads additional demand of passengers over the network and reduces their delay in comparison to the line based perspective.

We can distinguish two perspectives based on how passenger behavior during a disruption is looked at.

**Line-based perspective:** The passengers of the disrupted line (i.e., the ‘800’) travel on the most plausible detour via Tilburg by taking the ‘1900’ and ‘16000’ lines. That is, the passengers are bound to their intended line as much as possible.

**Network perspective:** The passengers are free to select any path in the entire network based on the information that is available at that time.

These two perspectives play an important role in how the operator assesses the disrupted situation. Current railway practice tends to lie rather close to the line-based perspective. The operator may want to allocate additional resources to the ‘1900’ and ‘16000’ lines in order to prevent overcrowded trains. Even more importantly, the operator’s perspective is reflected by the travel advice it communicates to the passengers.

Passengers are free to choose their own route in time, provided that they stay on the shortest geographical route, within the setting of Netherlands Railways. An operator can allow passengers to enter a wider part of the network, for instance in case of an disruption. An OD-dependent detour equal to the network perspective in the example can reduce the delay of passengers and can help spread the additional demand more evenly over the network, both leading to an increase of passenger service level.

In this research we assume that both the passengers and the operator have the network perspective. The network perspective is never worse (and mostly better) for an individual
passenger than the line-based perspective. Therefore, assuming that the train capacities are sufficient, the network perspective leads to a passenger flow with a superior service quality. It is the operator’s responsibility that the passenger flows are facilitated by providing the necessary train capacity, on one hand, and by spreading carefully selected information and travel advice to the passengers, on the other hand.

This paper aims at designing methods for accurately forecasting the passenger demand, taking the passenger’s free will into account. Our future research studies use this forecast. We are particularly interested in the problem of informing the passengers. The travel advice can advertise easily overlooked yet optimal routes (such as the aforementioned route Amsterdam - Rotterdam - Eindhoven); also, it may convince a few passengers to take an equally good (or slightly worse) alternative route in order to decrease the crowdedness of a train whose capacity cannot be extended. This future application gives a further reason for the forecasting model to consider multiple routes for each passenger.

2.2 The proposed two-step methodology

The example in Section 2.1 indicates that not every part of the geographical network is equally relevant. The disrupted segment clearly is important, so are the segments Den Bosch - Tilburg and Tilburg - Eindhoven as they are likely to be part of many alternative routes.

On the other hand, the segments to the South of Eindhoven are irrelevant. Indeed, passengers between Maastricht (Mt), Roermond (Rm) and Eindhoven (Ehv) do not enter the disrupted segment at all. Passengers from any of these three stations to any other station must pass through Eindhoven. That is, if a passenger’s journey starts at these three Southern stations, the origin can be substituted by Eindhoven. A similar substitution can be applied for the destination stations. Finally, suppose that, after these substitutions, a few passengers have the same origin and destination; then these passengers can be treated as a homogeneous group of passengers, there is no need to distinguish them individually.

These observations motivate our two-step approach.

- The first step of our methodology is Network Reduction which amounts to identifying, based on the geographical network and on the timetable, the subnetwork consisting of all relevant stations and line segments. Informally speaking, the subnetwork is an area around the disrupted infrastructure within which the passengers’ path selection decisions do matter.

- The second step of our methodology processes a database of historical travel information and computes a demand forecast for the disrupted day. We assume that the disruption does not deter passengers, they do not choose another mode of transport. The disruption may, however, alter the passengers’ route choice.

The forecasting model’s input contains the set of historical travel data. The output is defined over triplets origin, destination and start time (shortly ODT triplets). The entries of an ODT triplet refer to stations and start times in the reduced network. For each ODT triplet, the output specifies the number of passengers, obtained by aggregating passengers with the same detour options. The output also specifies the ODT’s route on an undisrupted day and
their intended route on the disrupted day. Note that the time dimension of the ODT triplet is necessary due to the large fluctuation of the passenger figures during a day.

We note that our choice to allow one single undisrupted route and one single disrupted route to an ODT triplet was made for the sake of a proof-of-concept implementation. Our method can be extended easily to handle cases where a passenger can select one of multiple routes based on an utility model.

The Network Reduction step is not strictly necessary for the forecast. However, our motivation to embark on it are three-fold. First, the reduced network is substantially smaller that the entire network. More dramatic is the effect of grouping passengers by their substituted origin and destination. The Dutch railway network features 160,000 possible origin-destination (OD) pairs (not even accounting for the time aspect), while 80% of the journeys is realized between 15% of the OD pairs. The Network Reduction step helps in decreasing the computation time of subsequent steps. These subsequent steps include the demand forecasting as well as our future applications.

Second, the high segmentation of the historical data means that a significant part of the network load comes from ODT triplets that have just a few observed passengers in the historical data. By combining ODT triplets based on their substituted origin and destination, we expect a statistically more reliable forecast.

Third, we aim at involving practitioners in the evaluation of our methods. It is impractical to consider hundreds of thousands of forecasts. The Network Reduction’s more compact output has better prospects to give insight into the spatial and temporal character of the demand.

### 2.3 Data sources

We forecast the number of passengers per ODT triplet, where the stations and time periods result from the reduced network. The specific set of ODT triplets depends on the Network Reduction, which changes based on a specific disruption. A journey given origin, destination and start time, can always be linked to an ODT triplet.

Until recently data for forecasting ODT triplets was not available. Currently periodic field counts of passengers, a passengers per train forecast model, and a two-annual survey are used for forecasting passenger flows. The field counts and forecasts per train do not contain OD information. The survey is not detailed and frequently enough to capture changes over time.

However in most recently introduced smart card ticketing systems the required data is stored. Such systems are currently available in amongst others Tokyo, Seoul, Singapore, the London Underground and the Netherlands. These systems require a check in and a check out that register the origin, destination, and start and end time of each journey. Often this data is available with some delay, for the Netherlands this delay is between 15 minutes and a day. The ticketing system hence creates historical data that is an ideal source for forecasting. In this paper we use a real life smart card data set of Netherlands Railways for forecasting.
3 Literature Review

In this section we give an overview of related research to short term forecasts of passenger flows for disruption management. We start with literature on demand forecasting for transport, followed by a short overview of research on the application of smart card data and concluding with some key references for disruption management.

3.1 Forecasting OD flows and demand for transport

In this section we focus on research on demand forecasting for public transport. Before smart card data existed, these models usually relied on panel data or aggregated data, and focused on long term predictions of demand and elasticities.

Some research focuses on long term predictions of demand, like Gaudry (1975) who uses aggregated data on the demand for public transport in a specific urban area, to forecast future demand based on the price of public transport, the price of alternatives to public transport and the demographics of the population, like the income distribution of its inhabitants. Wardman (2006) specifically focuses on forecasting railway demand in the long run, using similar variables as the model of Gaudry (1975).

Others focus on the elasticity of demand. Hsiao and Hansen (2011), for instance, focus on predicting demand for air passengers, combining in their model predictions for demand generation and assignment. They base their model on panel data. They focus on predicting the sensitivity of demand to both time and price of the journey. Rolle (1997) also forecasts the elasticity of demand to price, specifically for railway demand, taking into account the different kind of services offered. Batley et al. (2011) focuses explicitly on the long term effects of lateness of trains on railway demand. He compares a market level model with an individual level model based on panel data, and comes to the conclusion that the effects on the aggregated level in terms of elasticity is less severe than the individual level models based on panel data suggest.

From the area of complexity research, González et al. (2008) show that human mobility patterns are very well predictable, based on a US experiment using cell phone location data. The AURORA\(^1\) project that focuses within NS on predicting numbers of passengers per train, is very successful in their approach based on several data sources, that since recently also include smart card data.

We are interested in short term predictions of passenger flows as opposed to the long term predictions and focus on elasticities that is prevalent in most of the above literature, and in the specific focus on OD flow prediction. The success of long term predictions for similar kinds of flows, and the different approaches from both macro and micro models, are an indication of the viability of this approach.

\(^1\)AURORA is the name of a series of projects focused on long term predictions of the number of passengers per train.
3.2 Research on Smart card data

Literature on smart card data and applications of smart card data mostly stem from the start of the current millennium. Blythe (2004) was one of the first to give a functional overview of the introduction of smart card systems for ticketing in public transport. The recent paper by Pelletier et al. (2011) provides a review of literature focused on analyzing smart card data. They divide the literature into the categories strategic level, tactical level, operational level and commercialization. We use a different categorization of those articles on smart card data that are related to this paper, and divide them into the sections passenger behavior and route choice and OD. We refer the interested reader to Pelletier et al. (2011) for a broader review of smart card data.

Analysis of passenger behavior

One of the uses of smart card data is for analyzing passenger travel behavior. Bagchi and White (2005) are one of the first to use smart card data. They use smart card data of a UK bus company, where they analyze the passenger population of that company. Similar is the work of Park and Kim (2008) who analyze the usage and mode choice of passengers based on smart card data of the Seoul public transport system.

Morency et al. (2007) and Agard et al. (2006) extend these analyses by using data mining techniques to deduce travel behavior for different groups of passengers. They show that clustering techniques can reveal patterns of groups of passengers that are valuable for the analysis of transport usage for a PTO. They focus on the analysis of past performance and behavior of passengers. Also, in their data sets the destination of the journey is not included.

Zureiqat (2008) focuses on the prediction of passenger travel behavior for revenue management. He focuses on the prediction of product type choice together with the number of journeys per passenger. He develops a model for forecasting the effect of product and price changes of a PTO.

Route choice and ODs

Second to the analysis of passenger travel behavior, few articles exist on estimating route choice and ODs from smart card data. These estimations are in general focused on in hindsight constructing route choices and ODs of passengers, as the destination is not registered in every source of smart card data. Seaborn et al. (2009) analyze smart card data of the London system, investigating transfer times between different legs and modes of transportation. Kusakabe et al. (2010) focus on linking check ins and outs to trains for the Tokyo system. In this system they focus on one line having different services. Based on the check ins and check outs, they deduce the route choice as the (combination of) services the passenger has chosen for his journey. The algorithm they propose is very similar to the algorithm used in practice by NS for linking journeys to routes.

We also mention here the work of Zhao (2004) and Gordillo (2006) because they focus on OD matrix estimation, although their problem setting is different from the one studied in this paper. In their paper, they do not aim to forecast the OD flows, but they focus on deducing these from historic data, as the end locations of trips are not registered. From the
estimated OD matrix, secondly the routes of the passengers are inferred. The article focuses on the London bus system in which, different from the London subway system, just a single check in is required.

Summary smart card data research

Current literature on the analysis of smart card data shows that it is a valuable source of information to deduce and qualify passenger travel behavior. However, to our knowledge, none of the research has focused on forecasting passenger behavior or OD flows from the smart card data.

3.3 Disruption Management

Caprara et al. (2007) give an excellent overview on Operations Research (OR) in passenger railways. They present many operations research models used for line planning, rolling stock circulation, crew planning and shunting, among others.

The focus on passengers in disruption management is rather recent, of which good examples can be found in Dollevoet et al. (2010), Nielsen (2011) and Veelenturf (2010). In their research on disruption management they take a passenger oriented approach. Minimizing passenger delay is the objective of their research. They include the reaction of passengers to a disruption. However, they assume initial OD flows given, while in practice this is in general not true. Therefore the results of the current paper would be valuable input for their research.

4 Methodology

This section presents the two-step methodology for studying the effect of large scale disruptions on passenger behavior. The methodology results in forecasts of passengers affected by a disruption, given their origin, destination and departure time. This information is required to study the interaction between the passengers, the vehicle schedules and the timetables. These forecasts will be such that given any assumption on passenger behavior and the given adjusted timetable and vehicle schedules, the additional demand per train can be calculated and shortages in capacity can be signaled. This is the essential information required for passenger oriented disruption management.

4.1 Processing smart card data

Smart card data contains the origin, destination, start and arrival time of each passenger journey. To focus on passengers affected by a disruption, we need also information on the path of a passengers besides the origin and destination. Therefore, we link journeys to specific routes. This processed data will serve as input for the forecasting model.

A journey, $j$, contains a departure time $t_d$, an arrival time $t_a$, a departure station $s_d$ and an arrival station $s_a$. To make the journey, the passenger travels by train. The train schedule
can be represented as a directed graph $G(V, E)$ with $V$ the set of vertices representing a station and time, and $E$ the set of arcs representing either trips of a train connecting two stations, or waiting arcs connecting two events in same station over time.

In the timetable graph $G(V, E)$ we search for journeys connecting a specific origin and destination over time. As passengers do not always travel along the shortest path, we find multiple paths by the following procedure. To find paths for a specific OD, $s_1, s_2$, we take all outgoing arcs from the origin station $s_1$, then compute the shortest path from the end points of all these arcs to the destination $s_2$ while forbidding all arcs ending or starting in $s_1$. This results in a collection of paths $p(V_p, E_p) \in \mathcal{P}$ in the network connecting $s_1$ to $s_2$. Some of these paths are dominated by other paths. A path $p$ is dominated by another path $p'$ when $p'$ departs no earlier than $p$, arrives no later and has no more transfers than $p$. We reduce the set of paths to a nondominated set by these criteria, denoted as $\mathcal{P}'$. Note that each OD has its own collection of paths $\mathcal{P}'$ associated to it.

Next we link the smart card data to a path in the network. For each journey $j$ we take the collection of paths $\mathcal{P}'$ belonging to the OD of the journey $j$. Next we select all paths in $\mathcal{P}'$ that fit within the departure and arrival time of $j$. For 70 percent of all journeys, just a single path remains. For 98 percent of all journeys, less than three paths remain based on which we need to select one path $p$.

A path is selected according to the following three rules. The first rule is based on the fact that most passengers check in and check out close to departure and arrival time, where the check out time $t(a)$ is very close to the arrival time of the train, and hence the path. Our first rule checks whether there are any paths $p$ that have a small margin to $t_d$ and $t_a$. When multiple paths remain, we make use of the second rule to prefer paths with less transfers, as passengers are known to generally prefer routes with less transfers. When neither of the criteria has lead to a single remaining path, we choose one from the remaining set based on minimal cost.

We process 10 months of smart card data based on this process and a planned timetable. We use a planned timetable as this allows for comparison of passengers per path over time, aiming to overcome the effects of disturbances in the past on the forecast. This process results in a number of passengers per path $p$ per day. These are the route based time series we will use for forecasting and for the Network Reduction described in Section 4.2.

4.2 Network Reduction

Disruptions are likely to affect a larger part of the network. The example in Section 2.1 showed that although a larger part of the network is affected, there are just a few key points at which passengers can decide to take a different path. Whether passengers pass such a point depends on their origins, and the set of such points along their paths depends on their destinations. There can be several ODs that have the same set of key decision points and the same rerouting options.

Our objective is to find an alternative OD matrix that can provide estimates of passengers per train given a timetable and behavioral information of the passenger, but that is also small in size. We propose to aggregate all journeys with similar rerouting options into one origin-destination pair, OD'. By selecting only ODs with a path through the disrupted area, we
explicitly focus on the changes in the network. By aggregating ODs with the same rerouting options we reduce the size of the OD matrix but are still able to study the interaction between passengers and trains. Finally, the OD’s enable time aggregation.

We refer to this technique as Network Reduction (NR), because the new OD’s together with the set of rerouting options span a smaller network in comparison to the full public transport network, while containing all edges where we expect a change in demand due to the disruption. Consequently the focus is on the changes in demand in the network. Section 4.2 presents a formal description of the NR.

**Formal description of Network Reduction**

We take the timetable graph $G(V, E)$ as described in Section 4.1 with $V$ the set of vertices (station, time) and $E$ the set of arcs (trips by train or wait arcs). We consider all journeys to be all possible combinations of two stations $s_i, s_j$, with $i \neq j$. For each $s_i, s_j$ we have a set of non dominated paths $P'$ where each path $p \in P'$ contains an ordered list of vertices connected by arcs, $p : (V_p, E_p)$ as discussed in Section 4.1. Note that for each $s_i$ there exists a subset $V_{s_i} \subset V$ representing station $s_i$ over time.

The Network Reduction selects those journeys affected by a disruption, then finds their first and last deviation from the original path. These stations are selected as new origin and destination, the OD'. The intuition behind this is that as long as passengers are on their planned path, the capacity that is available will be sufficient, and the trains will run according to plan. Hence the passenger will not incur any further delay or experience inconvenience on these parts of the path. Delay and insufficient capacity can therefore only be encountered on the detour or affected trains. The idea is to cluster passengers that have the same set of detour options.

The detour options depend on the origin and destination of the journey. Several journeys can have the same rerouting options. The idea is to assign new origins and destinations to each journey, such that based on this OD' we still find the same rerouting options for the journey, while at the same time having only one OD' per uniquely defined set of rerouting options. We do this by selecting the first point of deviation from the path in case of a disruption and last point of return to the path as the origin and destination of the OD'.

We show that the path assigned to an OD' is the same as the path that is assigned based on the original origin destination information of that journey. Moreover, we show that there is no other OD pair contained in the planned path of the OD' that spans a smaller area of this path while providing the same information.

We choose a weighting function $f(x)$ that assigns costs to the arcs of a path $p$. We take this function such that the costs defined by $f(x)$ are a linear combination of the weights of the arcs. Weights of arcs are always positive and represent the time difference between the start and end of the arc. The weight function does not distinguish between arcs connecting two nodes of the same stations and arcs connecting two nodes of different stations.

We consider two graphs. The first is the planned timetable graph, $G$, the second graph $G'$ belongs to the disrupted network. A disruption means that a set of arcs connecting two stations $s_i, s_j$ are taken out for as long as the disruption lasts. Passengers whose route contained such an arc in their path $p \in P'$ now need to replan their journey in this second
graph $G'$, that we assume to represent the adjusted timetable given the disruption. These journeys are considered to be affected by a disruption.

The OD' give the same path assignment Take an affected journey $j_i$. The associated path $p \in P'$ thus contains at least one disrupted arc $e_{\text{disr}}$ in the set of arcs connecting departure station $s_d$ with arrival station $s_a$. Consequently the passenger needs to replan the route based on the graph $G'$ and the cost function $f(x)$. We denote $E'_{u,v}$ to be the set of arcs that are in $p'$ but not in $p$ based on a station-wise comparison, independent of time. There must be at least one arc that connects the vertices that were connected by $e_{\text{disr}}$ that was deleted from the path. If the set is empty, there is an alternative direct connection for the disrupted arcs. The first origin station and last arrival station of the disrupted arcs in the path are taken as new clustered origin and destination. However, these situations in practice are rare.

In case the set $E'_{u,v}$ is nonempty, arcs are ordered over time. The first vertex $u$ defined as the departure station of the first arc is where the path $p'$ started to divert from the path $p$, meaning that it ended in another node $w$ which station was not on the original path $p$. The last arc in $E'_{u,v}$ returns path $p''$ to $p$ in a vertex $v$, meaning that from $v$ and onwards all the stations represented by vertices in the path are also in the original path $p$. The stations belonging to vertices $u$ and $v$ are the candidates for OD'.

The detour path for OD' should be the same detour path as for the original journey $j_i$. Let us take $t_s$ as the arrival time in node $u$ given the original path $p$. Now we calculate the shortest path from $u$ to $v$ starting at time $t_s$ denoted by $l$. This has to be the same path as the path in $p'$ connecting $u$ and $v$, $t'_u,v$, because if $l < t'_u,v$ then surely we could improve on $t'$ by replacing $t'_u,v$ by $l$, which contradicts that this is the shortest path from $s_d$ to $s_a$ in the graph $G'$. Similarly, if $l < t'_u,v$ then this contradicts that $l$ is the shortest path in the graph $G'$. It is possible that the cost of $l = t'_u,v$ while the paths are not the same. In this case we assume the passenger to be indifferent between the two and to find $l$ equally good as $t'_u,v$.

So we showed that given the graph $G'$ and the function $f(x)$ we find the same detour path for OD' as the original journey. Moreover, given OD' we extract the change in route, and hence the change in demand per train.

Selecting ODs to uniquely represent a detour We now show that OD', as selected above, is the only OD to represent the whole change in path due to the disruption.

In OD' origin station $o'$ and destination station $d'$ need to be stations included in the path $p$. If $o'$ is not in $p$, then a passenger will have to use an arc $e_{o'}$ not in $p$ to arrive in station $o'$. As $e_{o'}$ is not in $p$, this arc is part of the detour route, and hence should be in the path belonging to OD'. However, arc $e_{o'}$ cannot be in the path of OD' as this arc arrives at the start point of the path, $o'$. Take path $p'$ to be the path of OD'. As $p'$ is a detour path of $p$ there are at least two stations of $p$ in $p'$, the start and end of the detour. As $p'$ starts in $o'$ which is not in $p$, there is at least one arc $e_{o'}$ that connects path $p'$ to a station in $p$, and therefore cannot have been part of the detour path belonging to $p$. Therefore, $o'$ needs to be a station that is in $p$. A similar argumentation holds for $d'$.
Assume we choose a different station for the origin than $o'$. Say station $o''$ is in path $p$ after station $o'$. There are two cases. One, the detour path $p''$ for $o''$, $d$, contains the detour path $p'$ belonging to $o$, $d$. As this detour path started in $o'$, there must be at least one arc connecting $o''$ to $o'$ which would not be in $p''$ as $o''$ is after $o'$ and then the costs of this path would be more than $p'$. In the second case, the path $p''$ does not fully contain $p'$. If $p''$ is contained in $p'$, this would contradict that the first point of the change in path is in $o'$ or that $p''$ contains all arcs that are not in $p$. Otherwise, $p''$ has to be more expensive than the change presented in $p'$ because $p'$ represents the least cost path in $G'$.

We choose again a station different than $o'$, $\hat{o}$ that is on $p$ but before $o'$. The detour path $\hat{p}$ from $\hat{o}, d$ has to contain the detour path $p'$. If there would be a different path $\hat{p}$ better than $p'$, this would contradict that $o'$ is the first location of a change in path. As path $\hat{p}$ thus has to contain the detour path $p'$, it also contains at least one additional arc, $e_{\hat{o}}$ to connect $\hat{o}$ to $o'$ that is not part of the detour path. Therefore $\hat{o}, d$ does not only represent the change in path.

We showed that $o'$ is the only station that can help to uniquely find the change in path due to the disruption. A similar reasoning holds for $d'$.

Adding behavior and capacity constraints The above strongly depends on the shortest path. However, in case of capacity shortages this shortest path not only depends on the adjusted timetable $G'$ but also on the path choice of other passengers. To find any path, we need a timetable graph $G$, a cost function $f(x)$, an origin station $o$, a destination station $d$ and a departure time $t_o$. We denote the latter three as the tuple $j_{o,d,t}$, a journey $j$ from station $o$ to station $d$ departing at time $t$. We now extend the above presented method to find those $j_{o,d,t}$ that even in case of capacity constraints will find the best path. We describe the procedure for the origin candidate alone, as the search for a destination candidate is similar to this procedure.

We introduce two behavioral assumptions for the passengers. The first is that a passenger will never make a detour longer than a time $t_m$. This is a common assumption in disruption management with passenger flows acknowledging that there is a limit to the willingness of a passenger to incur delay for a given journey before the passenger cancels the trip. Secondly, we assume that there is a limit to the delay a passenger is willing to incur, $t_n$, in comparison to the shortest detour path out of precaution. This model assumes the passenger has some knowledge or feeling about his minimal delay, and assumes the willingness to make a longer detour to increase the chance of a seat or avoid competition for capacity, not just related to the maximum delay but also the price that is to be paid for such a possibly more convenient path.

Note that $j_{o,d,t}$ does not include a specific path, but just contains the first point in space and time that a passenger will start to reroute. Hence our question reformulates to: would there be a station before the found OD where a passenger, given the two above described behavioral constraints, would be willing to start rerouting. If not, we assume the passenger would never start rerouting before OD, and hence we can take OD to be the shorter representation of the original path and OD. If so, we can replace the origin of OD by the new and earlier origin station of the new rerouting path.

The methodology to select an OD is thus extended by an iterative procedure. In each step all arcs leading to the current candidate for the clustered origin are taken out. Next a new
shortest path is calculated given this adjusted graph. The new path is compared and evaluated against the best detour path and the original path given thresholds $t_m$ and $t_n$. As long as the comparison falls within the thresholds, a new and earlier origin is selected and the procedure continues. Finally, the departure time of the tuple $j_{o,d,t}$ is derived from the arrival time of the original path $p$ at station $o$.

By this procedure we gain that in case of capacity constrains the reaction of passengers can be calculated, though losing the property that the path of $OD'$ represents the detour path alone.

**Network Reduction** The procedure for Network Reduction reduces the set of OD pairs to a smaller set, $OD'$, that still represents all information needed to calculate the reaction of passengers to the disruption. It focuses on the change in demand in the network. It not only reduces the dimensions of the OD-matrix, but also enables a reduction of the network by taking all arcs in the detour paths and all stations that are part of the $OD'$ set. Thereby it serves all three criteria important for demand forecasting for disruption management.

### 4.3 Forecasting the OD matrix

The processing of smart card data leads to a time series of passengers per path $p$ per day. Based on the Network Reduction, we select those paths that contain at least one of the disrupted arcs. Next, we aggregate paths on origin, destination, and time, that have the same rerouting options. This leads to a new and reduced set of time series, which we use for forecasting.

We compare a set of forecasting models that are widely used in practice. Exponential Smoothing and a forecast based on the average are used to give an indication of the predictability of these series. Secondly, we use an autoregressive model for forecasting. The AR(1) model is defined as:

$$y_t = c + \beta y_{t-1} + \varepsilon$$  \hspace{1cm} (1)

With $y_t$ the number of passengers for $OD'T$ on a specific day $t$. The forecasts per weekday are estimated in separate time series, as the correlation between weekdays is known to be higher than the correlation between consecutive days. Furthermore, there is a trend in demand over the year due to holidays. Therefore a second model, an AR(1) model with trend. As the latter two may differ in forecasts due to aggregation, we compare the outcome of the forecasts based on network reduction with forecasts based on the single series.

### 5 Results

Our dynamic forecasting model provides detailed time dependent forecasts of the number of passengers that are affected by a disruption. Through Network Reduction we greatly reduce the set of journeys to a key set of journeys, which simplifies our analysis and in many cases improves our forecasts. Based on these forecasts we can calculate the diversion of passengers given any behavioral rule and any new timetable.
In this section we present results of our dynamic forecasting model. As results depend on a disruption, we use three examples chosen such that estimates of passengers are likely to be important to improve service level in case of a disruption, because of expected capacity shortages and the possibility for passengers to reroute.

Results of the Network Reduction are presented in Section 5.1. Forecasts are derived based on the smart card data. Section 5.2 states the results of the forecasting. We find that we can somewhat improve forecasts based on the aggregation. Finally Section 5.3 shows how these forecasts can be used to estimate the need for capacity and the delay of passengers given behavioral assumptions of passengers.

The presented results are based on a 10 months’ sample of real life smart card data from Netherlands Railways, the largest passenger railway operator in the Netherlands. This sample contains over a third of all journeys made by rail in that period.

5.1 Network Reduction

Network Reduction selects journeys affected by the disruption and aggregates them based on available reroutes as described in Section 4.2. A general measure of success is hard to define, as the result depends on the specific disruption and the network. To get a feeling of how well it works, we have selected three scenarios of a disruption. These examples indicate that Network Reduction leads to a significant reduction of the size of the OD matrix of at least fifty percent. We discuss these results in this section.

The three scenarios are specifically chosen to be on busy and central parts of the network where we would expect a high number of journeys to be affected by a disruption. All scenarios consider a complete blocking of the track for some time. All scenarios are chosen such that passengers can divert to another route to get to their destination, which causes the need for forecasts of passenger flows for capacity rescheduling and provides the opportunity to improve service by providing travel advice to passengers, as replanning their route is in their benefit.

Table 1 contains the results for the Network Reduction in the three scenarios, stating the total number of affected journeys and the size of the key set of journeys. Secondly, the different numbers of stations in these two respective sets are given. Finally, the reduction is presented as percentage of the size of the full set. A percentage of nearly 5 percent indicates a reduction by a factor of 20. We have limited the set of total journeys to those journeys that on average in the past year had at least 3 journeys a day. These 6500 journeys are responsible for at least 95 percent of all journeys in the network. Though the algorithm is perfectly capable of taking on all possible 160,000 journeys, given the practical application the given comparison seems more fair. Moreover, we are now able to show that Network Reduction on its own reduces the number of journeys in a way that cannot directly be derived from data analysis.

The first scenario is that of a blockage between Utrecht and Breukelen, blocking the connection Utrecht-Amsterdam, which is chosen as it is one of the busiest routes in the network. The highest number of stations and journeys are affected by the disruption of this scenario. Network Reduction reduces the 821 journeys to 190, less than a fourth of the original set. The second scenario blocks the tracks between Gouda and Rotterdam Alexander, blocking
the route to the east of Rotterdam. Thanks to the introduction of a high speed line, there are numerous ways for rerouting, and the best ones strongly depend on the origin and destination. This is probably why the reduction factor is least for this scenario, but still reduces the number of journeys by a factor 2. The third scenario considers a blockage between Den Bosch and Eindhoven. The key journeys are just 5% of the total number of affected journeys, a reduction factor of over 20. Though less journeys are affected, a reduction from 276 to 12 shows that Network Reduction can greatly simplify the study of changing and rerouting passengers.

<table>
<thead>
<tr>
<th>Disruption Setting</th>
<th>OD</th>
<th>Stations</th>
<th>Reduction Percentage OD : Station</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ut, Bkl Network Reduction</td>
<td>190</td>
<td>52</td>
<td>23%; 35%;</td>
</tr>
<tr>
<td>Full Network</td>
<td>821</td>
<td>146</td>
<td></td>
</tr>
<tr>
<td>Gd, Rtd Network Reduction</td>
<td>123</td>
<td>41</td>
<td>39%; 47%</td>
</tr>
<tr>
<td>Full Network</td>
<td>318</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Ehv, Ht Network Reduction</td>
<td>12</td>
<td>7</td>
<td>4.3%; 4.6%</td>
</tr>
<tr>
<td>Full Network</td>
<td>276</td>
<td>151</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: For three disruption scenarios considering the blockage of part of a track, the number of affected journeys and stations in the full network is compared to the key set resulting from Network Reduction. Network Reduction reduces the set size by at least 50 percent, sometimes even by a factor 20.

5.2 Forecasting passenger journeys

We forecast the number of passengers that plan to make a specific journey based on the smart card data. Though the smart card data itself contains information on time, origin and destination of the journey, we need the route to link a journey to a specific key-journey and a departure time from the key-origin. This section is about forecasting the number of passengers on a specific day that plan to go from key-origin to key-destination before they have reacted to the disruption. This forecast is the basis for further analysis of passenger behavior, such as the spread of passengers over the network as described in the example in Section 5.3.

For each key journey and specific departure time forecasts are derived by time series analysis. The time series consist of the number of passengers on a specific day that belong to such an OD’T. The linked smart card data to paths in the graph (Section 4.1) together with the Network Reduction (Section 4.2) are used to form these time series. For each OD we deduce the key-journey from the Network Reduction. The specific path of the journey gives the arrival time at the key-origin station, which is the departure time in the key-journey. Given a specific start time of the journey, it is therefore possible to take into account the expected location of a (group of) passengers at that point in time, and assign them to the appropriate key-journey according to their currently expected location.

The results in Section 5.1 show that Network Reduction reduces the number of journeys significantly. We study whether forecasts quality changes through the Network Reduction by estimating several types of forecasting models for the transformed time series, and compare them to original non-aggregated time series. We compare the forecast performance on the
Table 2: Different performance measures of forecast quality. In most cases the aggregation leads to as good or better results as the separate forecast models. Values represent the average over all estimated models per scenario. The AR models, excluding and including a trend over time (AR(1)T) outperform the average and exponential smoothing (ES) model. As the ES model and mean-forecast by definition result in the same values for single and aggregated series, we have only reported one value.

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Disruption Scenario 1 Utrecht - Breukelen</th>
<th>Disruption Scenario 3 Eindhoven - Den Bosch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Relative Root Mean Squared Error (NR, Full)</td>
<td>Relative Root Mean Squared Error (NR, Full)</td>
</tr>
<tr>
<td>AR(1)</td>
<td>2.85, 2.86</td>
<td>1.76, 1.79</td>
</tr>
<tr>
<td>AR(1)T</td>
<td>2.87, 2.88</td>
<td>1.79, 1.80</td>
</tr>
<tr>
<td>Mean</td>
<td>2.87, 2.87</td>
<td>1.82, 1.80</td>
</tr>
<tr>
<td>ES $\alpha$ 0.7</td>
<td>3.51, 1.25</td>
<td>2.20, 0.97</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Absolute Relative Error (NR, Full)</th>
<th>(Root) Mean Squared Error (NR, Full)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>1.19, 1.20</td>
<td>6.67, 8.17</td>
</tr>
<tr>
<td>AR(1)T</td>
<td>1.02, 1.01</td>
<td>8.47, 8.83</td>
</tr>
<tr>
<td>Mean</td>
<td>1.22, 1.25</td>
<td>14.78, 14.86</td>
</tr>
<tr>
<td>ES $\alpha$ 0.7</td>
<td>1.25, 1.25</td>
<td>14.86, 1.93</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Root Mean Squared Error (NR, Full)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>1.40, 1.50</td>
</tr>
<tr>
<td>AR(1)T</td>
<td>1.51, 1.53</td>
</tr>
<tr>
<td>Mean</td>
<td>1.77</td>
</tr>
<tr>
<td>ES $\alpha$ 0.7</td>
<td>1.93</td>
</tr>
</tbody>
</table>

The difference between the two approaches is that for the Network Reduction approach we first aggregate the separate series and use this aggregate to estimate the forecast model, while for the full-network we generate forecasts for each series separately, and then aggregate the results based on the Network Reduction. As the Network Reduction and the number of affected passengers depends on the disruption, we use two of the same disruption scenarios as presented in Section 5.1.

We compare the models based on the error, using the Relative Root Mean Squared Error, the Absolute Relative Error and the (Root) Mean Squared Error. We find that the results for both approaches are close, but differences are mostly in favor of the forecasts based on the Network Reduction. Note that the Network Reduction significantly reduces the number of models that need to be forecasted, as each journey has several starting times, and therefore several time series linked to it. By reducing the set of journeys using key journeys, we reduce the set of time series even further. Given the time pressure in disruption management situation, this provides a huge benefit.

Results of the forecasting methods in Table 2 indicate that there is only a minor difference in performance between the forecasts based on the network reduction and the forecasts based on the full ODT time series. This comparison is not made for the Exponential Smoothing and Mean model because these forecasts are by definition the same for the Network Reduction and the full set of time series. Though several values of $\alpha$ were tested, the performance was similar, hence the results of a single $\alpha$ are included.
The forecasting quality is sufficient for the purpose of estimating demand per train, given the values for the error in Table 2. The statistically more advanced technique of the AR(1) models outperform the standard Exponential Smoothing and average model, indicating that more sophisticated models may be able to further improve these forecasts.

5.3 Estimating change in demand for capacity and passenger

Figure 2: A disruption on the 800 line between Eindhoven and Den Bosch requires passengers to reroute. Given a sudden disruption from 7:00 - 11:00h passengers reroute not just on the shortest route around the disrupted area - but following their new shortest path some spread across a larger part of the network. Thin black lines represent the tracks in the Dutch railway network. Thicker gray lines represent passengers on a detour path. A thicker line represents more passengers.

Based on the forecasts we can assign passengers to a specific detour route. To this end we assume a path choice mechanism of the passenger. We choose to use a shortest path algorithm, though it is possible here to use any type of path choice algorithm. Knowing the origin, destination and departure time, with the help of the network reduction, we can calculate detour routes. The usage of network reduction significantly reduces the number of shortest path to be calculated. For a two hour disruption between Den Bosch and Eindhoven, the number of original ODT’s is over 3000, while the number of clustered ODT’s is just around 300, about a tenth of the original set. As shortest path computations are expensive, this significantly speeds up the calculation. The detour path, the disruption and the length of the disruption together enables us to estimate the additional demand per train.

Moreover, given the exact ODT’s, we can calculate the delay per passengers. These calculation take into account additional transfer time due to the detour, giving a measure of passenger delay dependent on the actual journey. We find that the experienced delay strongly
depends on the origin and destination of the passenger. Though almost half of ODT’s experiences 30 minutes delay or less, over 20 percent of ODT’s experience over an hour delay. The delay can even differ dependent on the departure time of the passenger. This indicates that not only the path choice, but also the affect of a disruption is strongly dependent on the journey of a passenger. Hence ODT information is important to estimate passenger service levels, and these models can provide the essential information as input for disruption management models to maximize this service.

6 Conclusion

Forecasts of passenger flows are required for disruption management policies focussed on maximizing passenger service level. To this end, quick, comprehensible forecasts are needed than provide sufficient information for estimating passenger service quality, forecasting additional demand per train and allow to incorporate passenger behavior such as their preferred rerouting option.

In this paper we presented a two step methodology for forecasting these passenger flows. First, network reduction selects those journeys that are affected by a disruption. Furthermore, by focusing on the key-decision points, the number of ODTs is significantly reduced while these forecast provide sufficient information to estimate additional passengers per train and evaluate service quality. Secondly, this network reduction is used together with smart card data to forecast the number of passengers per OD’. A set of common forecasting models in practice showed that we can successfully forecast this number of passengers, and that the network reduction leads to no worse forecasts than the forecasts based on the full set of ODs.

Future research will focus on using these forecasts as input for disruption management models focussed on maximizing passenger service level through providing personalized route advice. Moreover, more advanced statistical models may be used for deriving better forecasts of the passenger flows.

References


