Switching Perspectives: Analysing train delays considering train transfer and cancellation

Johannes Bertram [*1] **Felix Böhm** [*2] **Nadja Buttke** [*3] **Fabian Morelli** [*4]

Abstract

Trains of the Deutsche Bahn AG are perceived to be notoriously delayed, but published delay statistics disregarding cancellations and transfer do not provide necessary information to estimate passenger's delays at their destination. We introduce an algorithm to analyse delay and cancellation data of long-distance trains from and to Frankfurt to estimate the final delay. The output of our algorithm is a dataset providing more accessible final delay information, also for individual connections. We show the necessity of considering the additional contribution of missed transfers and train cancellations to the mean train delays. We find that cancellations contribute about 6 minutes and, for low scheduled transfer times, missed transfers contribute between 30 and 60 minutes to the delay. For transfer times of 10 minutes or less, the additional delay due to missing connecting trains outweighs the benefit of lower waiting time and leads to later expected arrivals.

1. Introduction

Despite the well-known stereotype of German punctuality, the Deutsche Bahn AG (DB) is commonly known for its large delays and many cancellations. Passengers travelling with the DB typically start journeys across Germany with high uncertainty about their time of arrival. The DB publishes monthly statistics on train delays (DB, 2023), but they only consider the percentage of trains with delays larger than 5 or 15 minutes. Furthermore, even though train cancellations can lead to large delays for the passengers, they are disregarded in the DB statistics. This is especially critical as the DB cancels the final stops of trains with too large delays that are due to return in the same direction. The initial stops on the return journey are also cancelled to enable the train to resume its scheduled timetable. (DB, n.d.-b; Kriesel, 2019). While analyses on cancellations have been performed (Kriesel, 2019), train transfers, potentially causing delays when missed, are not taken into account. Thus, these analyses do not provide complete information about the delay at arrival for the passengers.

In this work, we examine the passenger delays arising from travelling with the long distance trains of the DB taking into account the three factors of arrival delays at the destination: Train delays, train cancellations and missed transfer trains. To make the computations feasible, we only consider transfers at Frankfurt central station. While Hamburg central station has more daily passengers, Frankfurt takes second place with 493,000 daily passengers and a more central location in Germany (FDP, 2020). In particular, our goal is to investigate the passenger delays for every pair of origin and destination train stations given that a transfer occurs in Frankfurt.

In the following we will introduce the dataset and our algorithm[1] and analyse the resulting passenger delays.

2. Methods

2.1. Data and Preprocessing

We use data from Bhatte and Preuß, 2024 including date, train station, scheduled arrival, scheduled departure, train ID, delay in minutes and cancellation information. The dataset contains data from 154 stations with all incoming and outgoing trains to and from Frankfurt main station from 20/01/2021 to 04/12/2023. We merged the two million data points by trains to receive about 166,000 incoming and 185,000 outgoing trains. Fourteen trains were deleted as their sequence of stops oscillated between different, far away regions in Germany or the scheduled driving time was not enough to reach the next stop.

[*]Equal contribution [1]Matrikelnummer 5461361, johannes.bertram@student.uni-tuebingen.de, MSc Machine Learning [2]Matrikelnummer 5443299, felix.boehm@student.uni-tuebingen.de, MSc Machine Learning [3]Matrikelnummer 6157302, nadja.buttke@student.uni-tuebingen.de, MSc Medical Informatics [4]Matrikelnummer 6634173, fabian.morelli@student.uni-tuebingen.de, MSc Machine Learning.

Project report for the "Data Literacy" course at the University of Tübingen, Winter 2023/24 (Module ML4201). Style template based on the ICML style files 2023.

[1]Publicly accessible code: https://github.com/17ex/DataLiteracyPublic

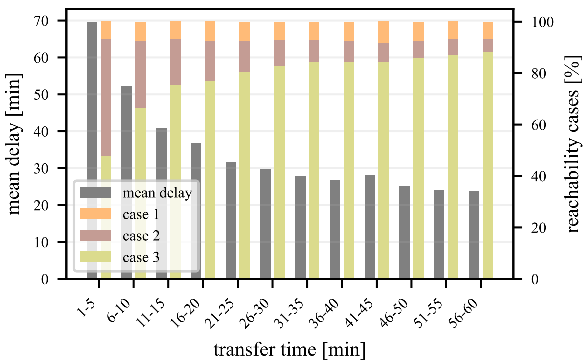

Figure 1. Mean delay at destination ( ) in minutes and reachability percentages of case 1: the first train is cancelled ( ), case 2: the connecting train is cancelled or missed ( ) and case 3: both trains are reachable ( ), plotted over scheduled transfer time in minutes on all origin-destination pairs

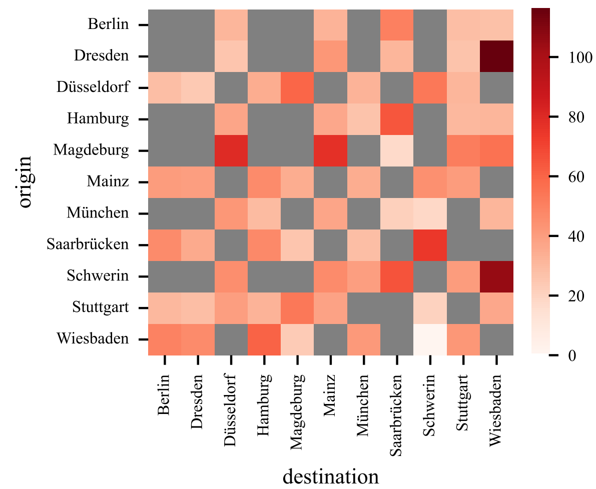

Figure 2. Mean delays in minutes for origin (y-axis) and destination (x-axis) pairs. The greyed-out pairs were excluded in the analysis, as it would not make sense to take a connection via Frankfurt for them.

2.2. Algorithm

To estimate delays on routes with a scheduled transfer in Frankfurt, we consider all pairs of origin and destination stations where a transfer in Frankfurt would be sensible. For example, when travelling from Stuttgart to Munich, a stop in Frankfurt is too large of a detour. Specifically, we only consider origins and destinations such that 1.5 times the distance between them is smaller than the sum of the distances between origin and Frankfurt and destination and Frankfurt. This results in keeping 79% of the 23562 possible routes.

To decide if a connecting train is reachable, the time of departure from Frankfurt is needed. However, the data only provides information about scheduled departure and arrival delay, not departure delay. Given the potential for a train to wait at the station for a delayed incoming train - a feeder train - from which passengers are transferring, or due to technical issues, the delay upon arrival does not provide an adequate estimate of the actual departure time. The research landscape using DB data is concerned with delay management (DM), the question whether, to minimize delays, a connecting train should wait for delayed feeder trains (Dollevoet et al., 2015). The approaches in DM are only evaluated on small subsets of German train stations (König, 2020; Liebchen et al., 2007). Typically, DM tries to minimize passenger delays (König, 2020; Schöbel, 2009) whereas the DB states that they minimize train delays while also considering passengers (DB, n.d.-a): Disponents decide on waiting for feeder trains case by case according to unpublished guidelines (Scheffer, 2023). While the DB conducted research on DM (Bissantz et al., 2005) and on estimating the number of passengers doing a particular train transfer (Liebchen et al., 2007), the exact decisions about waiting times remain unpublished. Thus, no clear assumptions about the waiting time of trains can be made. Therefore, we estimate the time gain on the route to the next stop and utilize the delay in the next stop to estimate the actual departure time in Frankfurt. To account for the uncertainty of this approximation, we use five different gain estimation methods including a lower and and upper bound.

No-wait Policy: All trains leave as scheduled.

Zero Gain: The scheduled driving time is subtracted from the actual arrival at the next station to obtain the actual departure

Average Gain: For every station, the average driving time from Frankfurt to the station is estimated and subtracted from the actual arrival at the next station to obtain the actual departure.

Average Positive Gain: For every station, the average of driving times faster than the scheduled time from Frankfurt to the station is estimated and subtracted from the actual arrival at the next station to obtain the actual departure.

Maximal Gain: The minimal driving time being 73% of the scheduled time (Dambeck, 2019) is subtracted from the actual arrival at the next station to obtain the actual departure.

The *No-wait policy* is a lower bound for the departure of the train, as trains do not leave early. While the *No-wait policy* is a heuristic that does not include the data, the other methods use the delay in the next stop to approximate the delay in Frankfurt. The delay in Frankfurt is equal to the delay in the next stop, adjusted for the time gained or lost on the way to the next stop. With *Zero gain* we make the assumption that no time is gained or lost on the way. *Maximal Gain* assumes the biggest possible gain of 27% of the driving time (Dambeck, 2019) and therefore is an upper bound on the departure of the train in Frankfurt. *Average Gain* and *Average Positive Gain* are based on data of trains that are passing through Frankfurt. For these trains we can calculate the delay leaving Frankfurt by subtracting the wait time from the arrival delay. With the departure delay we can calculate the average time gained or lost for all the stations directly connected to Frankfurt. This time gain is then used to estimate the departure time for all trains.

In every case, if the estimated actual departure would be earlier than the scheduled departure, the scheduled departure is used instead. Additionally, if the estimated actual departure would be earlier than the actual arrival in Frankfurt, we assume the train leaves immediately, that is, one minute after its arrival.

Our estimated departure time of the trains now allows us to check if connecting trains are reachable. Hence, we can apply Algorithm 1 to calculate the delays for every origin and destination pair. Our algorithm treats three cases: First, the incoming feeder train is cancelled, then the reachable train with earliest arrival in Frankfurt is chosen as the alternative and the second case is checked for the alternative connection. Second, the outgoing connecting train is cancelled or unreachable due to a delay of the feeder train. Again, the reachable train with earliest arrival at the destination is chosen as the alternative. Third, both the incoming and outgoing train are reachable as planned. The total delay for a connection is the sum of the scheduled delay that possibly arises from taking another train and the train delay of the train to the destination.

Along with general assumption such as passengers taking the fastest connection, several other assumptions need to be made. First, a transfer time of 1 minute is assumed to always be sufficient. Second, set the maximum planned transfer time to 1 hour and search for alternatives in case of a missed transfer in the next 4 hours after arrival of the incoming train in Frankfurt. If no alternative is found, we set the delay at the destination to 4 hours minus the transfer time. This occurs for 3.2% of all connections. Additionally, for both incoming and outgoing trains, if a train is cancelled either in Frankfurt (for incoming trains) or at the destination (for outgoing trains), we make the assumption that the passenger does not board this train. Instead, we presume they arrive via the next train that is not cancelled. If a train is missed or cancelled, the set of trains considered as alternatives are all trains that are reachable. This can lead to negative delays for connections as the original connection was cancelled and a faster alternative connection was found. These connections are ignored as they do not comply with our assumption that passengers take the fastest connection.

Algorithm 1 Delay calculation

Input: in, out: Incoming and outgoing train datasets
$orig, dest$: Origin and destination stations
Output: $delays$: List of calculated delays
$in \leftarrow$ filter($orig$ in $in.origins$)
$out \leftarrow$ filter($dest$ in $out.destinations$)
$connections \leftarrow$ merge(in, out, on=$Date$)
for each $connection$ in $connections$ **do**
  $arrival \leftarrow connection.planArrivalDest$
  if incoming train cancelled **then**
    Find next train from $orig$ not cancelled
    update $connection$
  end if
  if outgoing train cancelled or transfer missed **then**
    Find next train to $dest$ not cancelled and reachable
    update $connection$
  end if
  $newArrival \leftarrow connection.planArrivalDest$
  $planDelay \leftarrow newArrival - arrival$
  $trainDelay \leftarrow connection.delayDest$
  $delays$.append($trainDelay + planDelay$)
end for

The result of the algorithm is a new dataset for further analysis containing all origin-destination pairs. For every possible connection between the origin and the destination, the date, scheduled switching time, delay at the destination, and a marker indicating the reachability case (1, 2, or 3) is saved.

3. Results

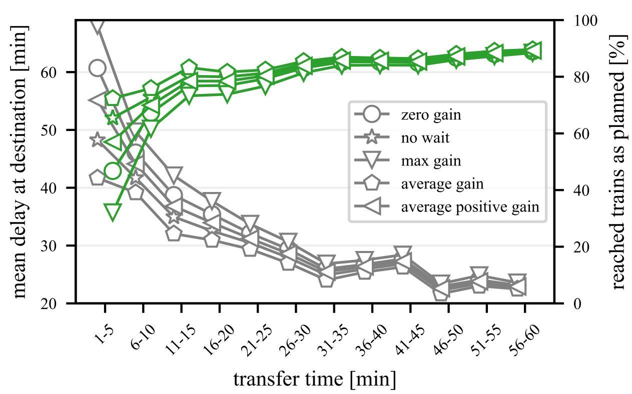

Figure 3. Comparison of mean delay at destination (■) in minutes and train reachability percentage (case 3) (■) of the different gain estimations plotted over scheduled transfer time in 5 minutes on all origin-destination pairs

First, we evaluated the five different gain methods from Section 2. Due to the high time complexity, we used a subset of 24 stations consisting of the 16 state capitals and all cities with a category 1 train stations in Germany (DB, 2024). As Figure 3 shows, the resulting delays for different gains are, as expected, bounded by the worst case (*No-wait Policy*) and the best case (*Maximal Gain*). For low scheduled transfer times, the methods vary the most as the number of missed transfers varies greatly with small changes in estimated departure. For larger transfer times, most transfers are possible and the methods produce similar results. Both the worst case and best case scenarios are unrealistic and an intermediate method should be chosen. Due to the qualitatively similar behaviour of all methods, any choice would be sensible. To interpret our results as expected delays, we choose the *Average Gain* for further analyses.

Using the *Average Gain*, Figure 1 shows the results on the full dataset. The expected delay with a scheduled transfer time of 1 to 5 minutes is about 70 minutes, largely because of missing the connecting train. With increasing transfer time, the delay decreases. Up to a transfer time of 15 minutes, the expected delay increases faster than the transfer time increases. Thus, a passenger is expected to arrive earlier if the transfer time is higher. For larger transfer times, most trains are reached and this effect does not persist. Accordingly, the number of reachable transfers increases from one third for low transfer times up to about 90% for a transfer time of 1 hour. This increase is due to the higher number of reachable transfers, while the number of cancellations (case 1) remains stable. The case 2 percentage reaches similar final values as case 1 indicating that the remaining impossible transfers are due to cancellations and not missed transfers.

As the expected delays and reachability percentages converge for increasing transfer times, it is possible to decompose the expected delay into its three causes: Missing the connecting train, train cancellations and arrival delays of trains. In Figure 4, the mean arrival delay of ca. 9.5 minutes for all trains irrespective of any transfers is plotted as a comparison. For all connections, this delay will be part of the final expected delay. The difference between the mean arrival delay and the limit of the expected delays is the delay caused by trains cancellations as for high transfer times nearly all uncancelled trains can be reached. The delay caused by cancellations and train delays amounts to a total of about 15.2 minutes. The additional delay present, especially for low transfer times arises, because of the missed transfers. It reaches a value between 30 minutes and 60 minutes for low transfer times while converging to 0 for increasing transfer time.

Our produced dataset also enables analyses of specific routes as displayed in Figure 2. For example, passengers travelling from Schwerin to Wiesbaden can expect high delays. However, the opposite direction produces lower delays. Apart from this, a symmetry in the plot is notable, i.e. the expected delay in both directions between two cities is similar.

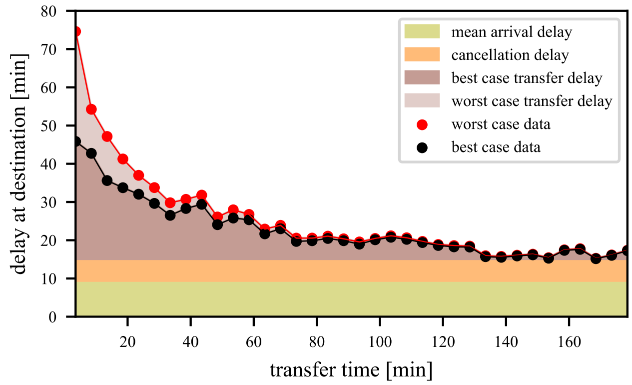

Figure 4. Cumulative decomposition of the mean delay at the destination into mean train delays, delays caused by cancellations and delays caused by missing trains in the best case and worst case

4. Discussion

We introduced an algorithm to analyse train delay data that takes into account train delays, cancellations and missed transfers for an expected arrival delay at the passenger's destination. This allows an analysis of estimated delays for specific connections and a look at transfers at different train stations. A limitation of our method is the high noise of train delays. Estimated delays have to be interpreted with great caution because individual train delays are noisy and depend on many factors such as weather, strikes, time of day and more. To deal with this, we used a large dataset dating back three years. The computational complexity of our algorithm forced us to search for connecting trains in limited time frames and to focus on Frankfurt as our sole transfer station. Thus, the generalisability to other connections across Germany is not given. However, we do not expect qualitative changes of the results for other large train stations.

5. Contribution

Felix Böhm was responsible mainly for data acquisition and preprocessing while also helping with the algorithm. Fabian Morelli implemented the algorithm and helped with the writing. Nadja Buttke was responsible for plotting the results. Johannes Bertram did the writing and helped with preprocessing and plotting. All team members discussed the algorithm and results together.

References

Bhatte, T., & Preuß, P. (2024). Bahn-analysen.de [https://www.bahn-analysen.de [Accessed: 22/01/2024]].

Bissantz, N., Guttler, S., Jacobs, J., Kurby, S., Schaer, T., Schöbel, A., & Scholl, S. (2005). Diskon - disposition und konfliktlösungsmanagement fur die beste bahn. *Eisenbahntechnische Rundschau (ETR)*, *45*(12), 809–821.

Dambeck, H. (2019). Ice versus tgv warum deutsche schnellzuege deutlich langsamer sind [https://www.spiegel.de/reise/deutschland/ice-versus-tgv-warum-deutsche-schnellzuege-deutlich-langsamer-sind-a-1259209.html [Accessed: 22/01/2024]].

DB. (n.d.-a). Anschlusszug wartet nicht [https://www.bahn.de/service/ueber-uns/inside-bahn/hintergrund-technik/anschlusszug-wartet-nicht [Accessed: 24/01/2024]].

DB. (2023). Erläuterung pünktlichkeitswerte dezember 2023 [https://www.deutschebahn.com/de/konzern/konzernprofil/zahlen_fakten/puenktlichkeitswerte-6878476 [Accessed: 22/01/2024]].

DB. (n.d.-b). Infografik: Warum züge manchmal ausfallen [https://www.bahn.de/service/ueber-uns/inside-bahn/hintergrund-technik/infografik-warum-zuege-manchmal-ausfallen [Accessed: 22/01/2024]].

DB. (2024). Stationspreisliste 2024 [https://www.deutschebahn.com/resource/blob/10549430/8b4014c743df66a357abf2c652860d8b/Anlage-1-Geplantes-Entgelt-pro-Station-alphabetisch-sortiert-getrennt-nach-Bahnhof-und-Bahnsteig_R1-data.pdf [Accessed: 29/01/2024]].

Dollevoet, T., Huisman, D., Kroon, L., Schmidt, M., & Schöbel, A. (2015). Delay Management Including Capacities of Stations. *Transportation Science*, *49*(2), 185–203. https://doi.org/10.1287/trsc.2013.0506

FDP. (2020). Verlässlichkeit des schienenverkehrs an knotenbahnhöfen [https://dserver.bundestag.de/btd/19/194/1919475.pdf [Accessed: 24/01/2024]].

König, E. (2020). A review on railway delay management. *Public Transport*, *12*(2), 335–361. https://doi.org/10.1007/s12469-020-00233-1

Kriesel, D. (2019). Bahnmining [https://www.dkriesel.com/blog/2019/1229_video_und_folien_meines_36c3-vortrags_bahnmining [Accessed: 22/01/2024]].

Liebchen, C., Schachtebeck, M., Schöbel, A., Stiller, S., & Prigge, A. (2007). Computing delay resistant railway timetables. *Computers & Operations Research*, *37*(5), 857–868. https://doi.org/10.1016/j.cor.2009.03.022

Scheffer, K. (2023). Aktualisierungsschreiben der bekanntgabe 12 zum handbuch 42002 - zusammenarbeit mit eisenbahnverkehrsunternehmen [https://fahrweg.dbnetze.com/resource/blob/10109672/3318e6ce87e8ef5defebf3d9a2754ea2/rw2024_420-02-data.pdf [Accessed: 24/01/2024]].

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