Delay Propagation and Process Management at Railway Stations

Rob M.P. Goverde and Ingo A. Hansen

Delft University of Technology Faculty of Civil Engineering and Geosciences Transportation Planning and Traffic Engineering Section Stevinweg 1, 2628 CN Delft, The Netherlands Phone: +31.15.2783178, fax: +31.15.2783179, e-mail: goverde@ct.tudelft.nl

July 2001

1 Introduction

Process operators at large railway stations have the difficult task to secure a fluent train traffic flow while minimizing deviations from established timetables. Variation in actual train departure times is inevitable due to many circumstances such as arrival delays and fluctuations in alighting and boarding time, even if some buffer time is contained in the dwell time. Moreover, a departure may be delayed by waiting for a feeder train to secure a connection and by conflicting train movements prohibiting an outbound train path. The predictability of train processes is even more degraded when in similar situations different control actions are pursued depending on for instance individual dispatchers.

In the Netherlands, passenger train services operate basically according to a cyclic timetable, repeating the same arrival and departure times each hour, with the exception of additional passenger trains in rush hours and freight trains that are scheduled in between the regular train services. It is hence anticipated that the traffic processes are mainly variations on a repetitious pattern. Analysis of historical realization data then yields operational insight that can be used to improve or support process management.

To gain accurate operations data a software tool, TNV-Prepare, has been developed that filters relevant train detection data from train describer records. This paper starts with a brief account of the collection and preparation of train detection data. Then for the particular case of station Eindhoven a detailed punctuality analysis is reported including the performance of dwell and transfer times (tightness or possible recovery time) and train waiting times to secure connections. Departure delays are predicted from arrival delays using regression analysis, whereas the remaining noise is attributed to human factors.

2 Train Detection Data

In railway systems the presence of trains is detected automatically by means of track circuits, axle counters, coils, or induction loops. In general, the occupation and clearance of a signal block or track section by a passing train is recorded and the data is saved a certain time for safety reasons. The location of the devices, however, varies and depends on the track layout and the design of the signalling system. In most cases the last measurement point before a station is situated some hundred meters or even more than a kilometre upstream of the platforms, whereas the first one after a station is located typically close to the departure signal. Moreover, the stop position of trains at a platform may vary if the length of trains is changing over time-of-day or day-of-week and the passenger access to the platform is not located at only one end. Therefore, the distances between the last (first) train detection devices before

(after) the station and the stop position of the different trains at the platform are to be determined in order to estimate the remaining deceleration (acceleration) time of the trains until (from) the stop.

Accurate data of Dutch railway operations are obtained from information of interlocking and signalling systems as received by train describer systems, which in the Netherlands are implemented as the so-called TNV-systems. A TNV-system is continuously logging in real-time the actual state of all relevant signal controls and monitoring information in a traffic control area, including the attached track sections, signals, points, and route relays. The resulting TNV-logfiles of a traffic control area consist of about 25 MB ASCII-format each day. The Dutch railway network is divided into 13 traffic control areas, which communicate through the TNV-systems of neighbouring traffic control areas.

rayon : gemaakt	EHV : 2000/03/	in 13-9-1997 serie 15 ?	: 1500	periode :	1997/09/13	t/m 1997/0	9/20					
Datum	TreinNr	Drg	Drg	23BT	27BT	33BT	47AT	24T [*]	107AT	108AT	121BT	193BT _
		Aankomst	Vertrek	BEZET	BEZET	BEZET	BEZET	BEZET	BEZET	BEZET	BEZET	BEZET
997-9-15	1519	07:24	07:29	07:22:15	07:22:27	07:22:33	07:22:42	07:22:50	07:29:45	07:29:53	07:30:16	07:30:22
997-9-15	1523	08:24	08:29	08:23:49	08:23:58	08:24:03	08:24:10	08:24:17	08:30:59	08:31:09	08:31:33	08:31:38
997-9-15	1527	09:24	09:29	09:21:51	09:21:59	09:22:05	09:22:13	09:22:21	09:33:36	09:33:43	09:33:59	09:34:03
997-9-15	1531	10:24	10:29	10:22:12	10:22:20	10:22:27	10:22:35	10:22:42	10:32:54	10:33:02	10:33:23	10:33:28
997-9-15	1535	11:24	11:29	11:23:25	11:23:34	11:23:40	11:23:48	11:23:55	11:29:57	11:30:10	11:30:37	11:30:42
997-9-15	1539	12:24	12:29	12:23:22	12:23:36	12:23:45	12:23:56	12:24:04	12:32:14	12:32:26	12:32:48	12:32:53
997-9-15	1543	13:24	13:29	13:21:05	13:21:13	13:21:18	13:21:24	13:21:31	13:30:02	13:30:09	13:30:28	13:30:32
997-9-15	1547	14:24	14:29	14:25:17	14:25:26	14:25:31	14:25:39	14:25:46	14:34:03	14:34:11	14:34:32	14:34:37
997-9-15	1551	15:24	15:29	15:31:14	15:31:22	15:31:28	15:31:35	15:31:42	15:37:02	15:37:09	15:37:27	15:37:31
997-9-15	1555	16:24	16:29	16:23:11	16:23:19	16:23:25	16:23:32	16:23:39	16:30:57	16:31:05	16:31:26	16:31:31
997-9-15	1559	17:24	17:29	17:31:58	17:32:08	17:32:14	17:32:23	17:32:30	17:37:20	17:37:35	17:38:00	17:38:05
997-9-15	1563	18:24	18:29	18:26:28	18:26:49	18:26:57	18:27:06	18:27:14	18:34:23	18:34:34	18:34:55	18:35:00

Figure 1. TNV-Prepare table of section occupancies before and after platform

Delft University of Technology (DUT) has recently developed the tool TNV-Prepare (Goverde & Hansen, 2000) that converts TNV-logfiles into tables per train line and route suitable for data analysis. Within TNV-Prepare, the rail infrastructure and signals are implemented as a set of coupled and connected objects. TNV-Prepare filters the files on the relevant objects, automatically tracks the (standard and non-standard) train routes from the data, recovers the signalling and interlocking events corresponding to individual train movements along the route, and checks the consistency of the results. TNV-Prepare gives reliable and compact tables of successive events along a train route, including successive section occupations and clearances, proceed and stop signals, and point switches. The user can choose to view/export a subset of events such as for instance the successive section entrances, see Figure 1.

The actual train length and speed trajectory of each train entering or departing a station can be estimated on the basis of the occupation and clearance times and the scheduled train characteristics. Estimation errors due to round-off errors (passage times are given in seconds) and deceleration variations during the approach are filtered by means of a least squares method and comparison of the calculated speed with the design speed, e.g. at signals and

turnouts. The occupation time of the platform track section itself includes dwell time and lasts until the clearance time of the train at the departure signal. The precise standstill time at the platform and the start of acceleration at departure, in general, are not recorded automatically, except when the train is equipped by an on-board processor and the data is transmitted to the trackside control system. The remaining deceleration and acceleration time of the train from and to the signals, however, can be estimated on the basis of the known standard deceleration and acceleration rates per type of train. This way, the arrival and departure times of each train at the platform tracks are determined with a precision in the order of a second (Goverde, 2000). The estimation procedure has been implemented in Matlab (Figure 2).



Figure 2. Estimated speed trajectory and arrival/departure delay using TNV-Prepare output

3 Station Eindhoven

TNV-Prepare has been applied to Eindhoven, an important transfer station with 6 platforms in the south of the Netherlands, where passenger trains from 4 main directions meet: Den Bosch, Tilburg, Venlo, and Roermond, see Figure 3.



Figure 3. Eindhoven station and surrounding route directions

Each hour 16 passenger trains arrive and depart in Eindhoven: 4 trains from and to each direction. These trains correspond to 9 train lines, including 3 intercity (IC) lines, 1 international (INT) line, 2 interregional (IR) lines, and 3 local (AR) lines. Eindhoven is the

origin/destination station of 5 lines (6 trains per hour) and 4 lines pass through (in both directions) (10 trains per hour), see Table 1. In rush hours, an extra passenger IR line of 2 trains per hour is scheduled from/to the direction Den Bosch. Also additional freight trains pass Eindhoven on 4 through tracks. Trains from (to) Den Bosch and Tilburg merge (emerge) at Boxtel from which a double-track route leads to Eindhoven. Between the long-distance lines 6 cross-platform transfers are scheduled, see Table 2 in Section 8.

		Train	Forwards (+)		Back (-)				
Туре	Line	/Hr	Platform	Α	D	Platform	Α	D	Origin-Destination
IC	800	1	1	57	59	5	35	37	Haarlem-Maastricht
IC	900	1	2	27	-	5	-	07	Haarlem-Eindhoven
IC	1500	1	1	24	29	6	05	09	Den Haag-Heerlen
INT	1800	1/2	4	-	33	4	01	-	Eindhoven-Koln
IR	1900	1	2	54	59	6	35	39	Rotterdam-Venlo
IR	2700	1/2	4	-	33	4	01	-	Eindhoven-Venlo
IR	3500	2	1	07	-	5	-	27	Utrecht-Eindhoven
AR	5200	2	2	15	17	6	14	16	Tilburg West-Deurne
AR	6400	2	3	-	05	3	27	-	Eindhoven-Weert
AR	9600	2	3	20	-	3	-	12	Utrecht-Eindhoven

 Table 1. Timetable and platform allocation in Eindhoven (1997/1998)

4 Arrival Delays

Arrival delays are usually the result of a late departure at the preceding station, too large running time, or route conflicts between trains. The running time of a train between two particular stations (slightly) varies for each trip depending on a wide range of factors from within the railway system and from exogenous sources. Internal sources are technical failures (of signals, switches, tracks, power supply and distribution, superstructure, and rolling stock), operations personnel behaviour (driver, conductor, dispatcher), and passenger flows (fluctuating alighting and boarding times depending on the amount of involved passengers). Exogenous sources are for example weather conditions. These sources of random variations are difficult to forecast and constitute a fundamental part of the practice of railway operations. Therefore there will always be a certain amount of trips that exceed a scheduled process time leading to primary delays, although scheduled running times and dwell times usually contain some margin or slack time to compensate for small variations. Also hinder by other trains via the signalling system such as a slow train upstream a single track or a conflicting train movement at a junction or crossing, influence the train running time (the so-called secondary delays). Departure delays even increase the probability of mutual hinder of trains as the delayed trains deviate from the scheduled train paths.

Analysis of realized train running times is therefore a crucial step in punctuality management, regardless of the used method in the planning process for the running time calculations. The variation in running times should be small in a well-designed railway system. Simulation is usually used to see the effect of interactions between trains in case of random primary delays of certain trains. Buffer times between conflicting train movements decrease the possibility of secondary delays. Analysis of realization data reveals tight headway times or the actual amount of buffer times at infrastructure bottlenecks.

Statistics of arrival delays at a station may help to identify unstable timetable designs and give directions to further analysis of train interactions. Figure 4 and 5 show some statistics of arrival delays at station Eindhoven (Goverde *et al.*, 2001). The general view is that intercity (IC) trains perform worse than interregional (IR) and local (AR) trains. The mean arrival delay

per IC line varies between 30 and 90 seconds (Figure 5). Special attention should be given to IC 900 and IR 3500, which clearly have inferior performance. In these figures e.g. IC800+ denotes IC train line 800 in forward direction. In this paper we will not analyse the arrivals in more detail, as we are here mainly concerned with delay propagation in Eindhoven.



Figure 4. Late arrivals and arrivals less than 3 minutes late in Eindhoven (Sept. 1997)



Average arrival delay [s] Standard Deviation [s]

Figure 5. Mean and standard deviation of arrival delays at Eindhoven (Sept. 1997)

5 Departure Delays

The major causes of a (large) departure delay are a late arrival, a prolonged dwell time due to boarding/alighting passengers or logistic reasons, waiting to secure connections, and a delayed outbound route setting due to conflicting train movements. Recall that a train is not allowed to depart early, i.e., before its scheduled departure time, and hence practically always departs late. Therefore a time window has to be detailed in which we assume a train to depart on time. A reasonable window is within one minute after the scheduled departure time.

Percentage arrivals<3 min Percentage departures<3 min



Figure 6. Share of arrival and departure delays smaller than 3 min in Eindhoven (Sept. 1997)

Figure 6 shows the share of arrival delays and departure delays smaller than 3 minutes, whereas Figure 7 gives the mean arrival and departure delays for the through and turning train lines (Goverde *et al.*, 2001). These figures clearly show that punctuality and mean delay at departure is far worse than at arrival, with the exception of the intercity line 900 that is terminating in Eindhoven. The mean departure delay varies from 90 to 150 seconds per train line. The considerable standard deviations of all train lines indicate that the timetable and dispatching in Eindhoven is unstable. Possible causes are too tight dwell times or transfer times and conflicting train movements. These possible sources of departure delay are investigated in the next sections.



Figure 7. Mean arrival and departure delay at Eindhoven (Sept. 1997)

6 Dwell Times

Four train lines pass through Eindhoven in both directions resulting in dwell times for 8 train directions. We analyse the dwell times of late arriving trains only and hence eliminate large dwell times caused by early arriving trains. From Figure 4 we see that this covers most trains. The scheduled and mean dwell times at Eindhoven are shown in Figure 8. In general, the mean dwell times at Eindhoven last about 40 seconds larger than scheduled, although one expects the opposite if buffer time is available. It can be concluded that the scheduled dwell times are very tight and there is not sufficient scheduled buffer time available. Possibly, some train drivers do not prepare in time for departure and the trains leave often 30 seconds or more behind schedule, even if no more boarding of passengers takes place. The stops at Eindhoven for through trains are thus structurally unstable.



Figure 8. Dwell times of late trains in Eindhoven (Sept. 1997)

7 Transfer Times

Transfer connections are one of the major sources of secondary delays. Especially in the highly urbanized Netherlands, where intensive train services are co-ordinated in an integrated periodic timetable offering good connections between any pair of stations with (cross-platform) transfer opportunities between main lines when direct connections are not available. In Eindhoven, 6 cross-platform transfers are scheduled between train lines from all 4 directions (recall Figure 3).

We analyse the transfer times for late feeder trains, since for these trains the connections might be cancelled. According to the Dutch railway operations regulations (WRT) e.g. IC trains may in general depart up to 2 minutes late if a feeder IC train is delayed. Only a fraction of the observed transfers is cancelled due to an extremely large arrival delay of the feeder train. The mean transfer times are slightly larger than scheduled, see Figure 9. The delay increase at Eindhoven of the 6 train lines involved in interconnections may hence (partly) be explained by waiting time to secure the connection, see also the next section.



Figure 9. Transfer times of late feeder trains in Eindhoven (Sept. 1997)

8 Regression Analysis

The train interconnections at Eindhoven may explain (some of) the variation in a departure delay, which can be explored using regression analysis. Table 2 shows the transfers along with the transfer times and the maximal waiting times according the WRT.

Train line	Dwell time [s]	Feeder train line	Transfer time [s]	Maximal waiting (WRT) [s]	Regression model of departure delay [s]	Explained variation [%]	Residual standard error [s]
IC800+	120	IR1900+	300	180	$38 + 0.91 A_{800+}$	92	36
IR1900+	300	IC800+	120	300	$41 + 0.76 A_{800+}$	86	31
IC800-	120	IR1900-	120	120	$73 + 0.82 \max(A_{800}, A_{1900})$	86	45
IR1900-	240	IC800-	240	120	$82 + 0.71 \max(A_{800}, A_{1900})$	94	32
IC1500+	300	IC900	120	180	$27 + 0.86 \max(A_{900}, A_{1500+})$	94	30
IC900	-	IC1500-	120	120	$64 + 0.53 A_{1500}$	42	51

 Table 2. Transfers and departure delay regression models

The regression models that best fit the data are given in the 6th column of Table 2. The last two columns give the associated percentage of departure delay variation that is explained by the model (multiple R^2 times 100%) and the standard error of the residual (remaining error between data and model). Figure 10 shows an example of a regression model.

The departure of IC 800 trains to Maastricht does not depend on the arrival of the feeder IR 1900 trains from Rotterdam. Recall however that the IR 1900 and IC 800 trains approach Eindhoven in this order over the same double-track route from Boxtel. In all observations this order is respected, and hence a large IR 1900 arrival delay also resulted in a large IC 800 delay. The departure time of the IC 1900 trains to Venlo also depends only on the arrival delay of the feeder IC 800 trains from Haarlem. Here, the large dwell time eliminates the arrival delays of the IC 1900 trains but the tight transfer time propagates the IC 800 arrival delays to the IR 1900 trains. In only 2 cases (out of 102) the IR 1900 trains did not wait for late IC 800 trains. Both regression models perform very well, i.e., the models explain most of the

variation in the departure delay and the residual errors are small. It can be concluded that in this bilateral transfer connection the IC 800 trains clearly dominate.



Figure 10. Linear regression model $D_{1500f} = 27 + 0.86 \max(A_{900}, A_{1500f})$

The interconnection between the Westbound IC 800 and IR 1900 trains results in a mutual dependence of the departure delays on the arrival delays. The departing route of the IC 800 and IR 1900 trains coincides from Eindhoven to Boxtel and is scheduled in this order. In only 3 cases (out of 61) this order is changed and the IR 1900 train departs first, which results in a missed connection in only 1 of these cases. Also, in only 3 (other) cases the IC 800 trains do not wait on a large delayed IR 1900 train. Clearly, the departure of the IC 800 trains to Haarlem then depends on the latest of both arrivals. And this also holds for the departure of the IR 1900 trains to Rotterdam, which departs after the IC 800 train.

The departure of an Eastbound IC 1500 train depends on both its own arrival and the arrival of the feeder IC 900 train. Again, both trains share the approaching route from Boxtel, where the IC 900 trains are scheduled after the IC 1500 trains. This order is changed in only 3 (out of 103) cases. In only 1 case an IC 1500 train does not wait on a large delayed IC 900 train. Still the regression model in both arrival delays outperforms the one in the feeder arrival delay only. As shown in Figure 10 the arrival delay of the feeder train is however in most cases the largest.

The IC 900 trains to Haarlem have to wait on IC 1500 trains from Heerlen with a tight transfer time of 2 minutes. Only in 3 (out of 95) cases the IC 900 trains do not wait on a delayed IC 1500 train. The variation in IC 900 departure delays can however not be explained satisfactory from arrival delays of the feeder trains as the bad performance of the regression model shows. Eindhoven is the terminal station of the IC 900 trains and hence the trains come from shunting tracks. A train is ready-to-depart for a new round-trip from its terminal station only if some necessary processes have been completed such as the shunting process, arrival of personnel, and several test procedures. This apparently results in departure delays that are larger than the arrival delays of the IC 1500 trains.

The intercept in the regression models, i.e. the constant (first) term in the regression model, gives the actual overload (or buffer for negative values) to the scheduled dwell/transfer time when the arrival delays are zero. Since here all intercepts are positive, ranging from about 30 to 90 seconds, each interconnected train pair has a too tight dwell time and/or transfer time. The predictor variable in the regression models is either the arrival delay of one of the trains or the maximum of both arrival delays, respectively. The first case implies domination of the used arrival delay over the arrival delay of the other train. In the latter case both trains propagate their arrival delay to the connecting train. The coefficients of the predictor variables are all positive and below unity, which indicate that the departure delays grow with the arrival delays but with a lesser rate. So from a certain arrival delay (determined by the intercept and the predictor coefficient) the departure delay gets smaller than the arrival delay. This implies that a small arrival delay result in an increased departure delay (the intercept dominates), whereas for a large arrival delays the (departure) delay is reduced. The latter behaviour corresponds to the practice that trains with large arrival delays are given priority and stop only for a minimum necessary time.

The regression models given in Table 2, with the exception of the last row, predict the departure delays from the arrival delays with an error less than 1 minute. These models can hence be used online by anticipating on predicted demands for outbound routes, which speeds up the departure process.

9 Conclusions

Based on the recent developed software tool TNV-Prepare, that amongst others recovers successive track occupancies and clearances of trains from train describer (TNV-system) records, the historical course of events in a railway station is analysed. Using regression analysis interrelationships between train arrivals and departures are identified and departure time prediction models from independent event times are derived. Possible residuals of the departure time predictions can be evaluated in order to explain remaining noise, resulting in timetable modifications or suitable process management procedures to reduce variation in process times and hindrance between trains approaching and departing from stations.

Accurate prediction of departure times and associated train path demands can support immediate route-setting schemes, when simultaneous requests for conflicting train paths can be expected and resolved well in advance. Moreover, evaluation and adaptive learning of realtime forecasts is an essential first step in the development of advanced automated process management systems that handle routine tasks automatically and report unusual situations to the process operators for intelligent decision-making.

References

Goverde, R.M.P. (2000), "Delay Estimation and Filtering of Train Detection Data", In: *Proc. TRAIL 6th Annual Congress 2000*, Part 2, No. P2000/3, TRAIL Research School, Delft.

Goverde, R.M.P., Hooghiemstra, G., Lopuhaä, H.P. (2001), *Statistical Analysis of Train Traffic: The Eindhoven Case*, TRAIL Studies in Transportation Science, No. S2001/1, Delft University Press, Delft.

Goverde, R.M.P., Hansen, I.A. (2000), "TNV-Prepare: Analysis of Dutch Railway Operations Based on Train detection Data", In: Allan, J., Hill, R.J., Brebbia, C.A., Sciutto, G., Sone, S. (eds.), *Computers in Railways VII*, WIT Press, Southampton, pp. 779-788.