

Reconstruction of train trajectories from track occupation data to determine the effects of a Driver Information System

T. Albrecht^{1,2}, R. M. P. Goverde¹, V. A. Weeda¹ & J. van Luipen³

¹*Faculty of Civil Engineering and Geosciences,
Delft University of Technology, The Netherlands*

²*“F. List” Faculty of Traffic and Transportation Sciences,
Dresden University of Technology, Germany*

³*Department Strategy & Innovation, ProRail, Utrecht, The Netherlands*

Abstract

The Dutch railway network is one of the most intensely utilized networks in the world and therefore many conflicts occur. The infrastructure manager ProRail developed and tested a Driver Information System (DIS), which gives information on current positions and delays of other trains on the same track. In a potential conflict situation, the driver is then able to anticipate to the behaviour of the previous train which should help avoiding unscheduled stops before red signals and reducing the consequences of a conflict (increased energy consumption and delay).

During the real-world test of the DIS, only few trains were equipped with accessible onboard measurement equipment which was insufficient for a profound analysis of the effects of it. It was therefore decided to reconstruct train trajectories by using track occupation data, which is available for all the trains with a precision of one second. A least squares approach is proposed to obtain the most probable train speed profile along the line. The reconstructed trajectories are used to support the manual observations of special events during the tests.

Keywords: driver information, traffic management, train detection system, speed measurement.



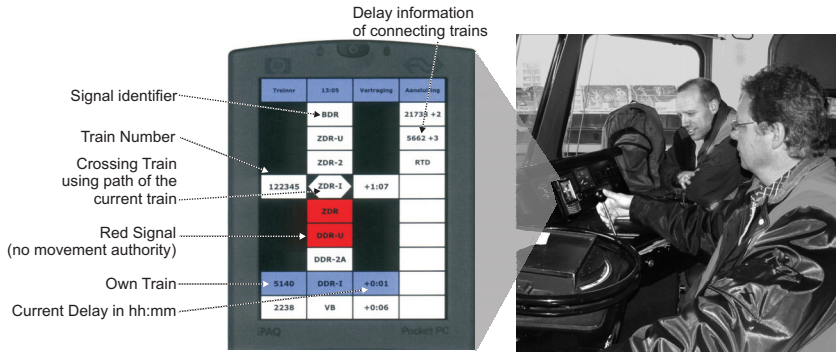


Figure 1: RouteLint: Extra information to the driver during the real-world tests.

1 Introduction

1.1 Motivation: the “Het Spoor Meester” project

Although timetables are constructed conflict-free with regard to schedule d train paths, it is impossible to avoid conflicts in practice due to unforeseeable random disturbances. Conflicts often lead to unplanned stops which cause delays to the trains that were forced to stop and also consecutive delays to other trains using the same route. Moreover, this also results in a higher energy consumption of the concerned trains due to unplanned re-acceleration.

In order to reduce the number and the consequences of those conflict situations, the Dutch railway infrastructure provider ProRail developed a Driver Information System (DIS) called RouteLint within the research project “Het Spoor Meester” (Dutch for: The Master of the Rail). This DIS displays train positions (per signal section) and delays of other trains as well as the already reserved path of the own train, cf. Fig. 1. With this additional information and their professional skills the drivers should be able to anticipate to conflicts before they occur and to adapt their way of driving in order to decrease the negative effects of conflict situations.

After very promising simulator tests of this system [1], it was decided together with the most important Dutch railway operators (NS Reizigers and NS Internationaal for the passenger trains and Railion for freight trains) to test the system in practical operation on the heavily utilized line between Rotterdam and Dordrecht during two weeks in November/December 2005. Before the tests, some train drivers were specially instructed using the DIS to be ambassadors of the project. During the tests, those ambassadors got onboard of selected trains travelling on the corridor and had to interpret the information given by the DIS to the driver in charge of the train (Fig. 1). Although the number of trains equipped with RouteLint was limited by the number of available ambassadors, more than 100 test runs have been done.

1.2 Measuring the effects

During the time the ambassadors were on board, they noted interesting events connected with the additional information given by the DIS.

It was nevertheless necessary to support this subjective information by objective data, e.g. train speed profiles or energy consumption. However, energy consumption is not measured per train but in the power substations. Due to the relatively low number of trains equipped with DIS it seemed impossible to use this data for an estimation of the effects of the system. The experiment outline (rolling stock equipped with DIS not known in advance) also did not allow to install onboard devices especially for the trials and neither were the data of the black boxes (train run recorders) accessible for non-safety purposes.

To overcome this dilemma Delft University of Technology proposed to make use of available track occupation times to reconstruct train speed profiles as suggested earlier by Goverde [2]. Those trajectories could partially be validated by measurements from energy meters installed onboard of a few trains.

This paper describes the theoretical background of the proposed method in section 2 and its practical application in section 3, which also discusses issues of data procurement, some sample trajectories and the validation of the approach with the energy meter data. Section 4 summarizes main results and gives an outlook on future research.

2 Theoretical background

2.1 Speed profile estimation

The motion of a train is a dynamic process with three variables that change over time: the position of a train, its speed, and the acceleration (or deceleration for negative values). In the sequel, it is assumed that the section passage times and section lengths are available. The passage times of sections on a route are generated by the VTL-tool, a tool similar to TNV-Prepare [2, 3]. The section lengths are provided by ProRail.

Consider a train run over a route of n track sections and assume that the n section entrance times are available from track circuit occupancy measurements. The train dynamics on this route can be described by the discrete Newton laws of motion

$$\begin{cases} x_{k+1} &= 3D \quad x_k + v_k(t_{k+1} - t_k) + \frac{1}{2}a_k(t_{k+1} - t_k)^2 \\ v_{k+1} &= 3D \quad v_k + a_k(t_{k+1} - t_k), \end{cases} \quad (1)$$

for $k = 3D1, \dots, n-1$. Here x_k is the (known) route position of the beginning of section k , v_k is the speed at this section boundary, a_k is the (average) acceleration on section k , and t_k is the measured entrance time of section k . The main problem here is that the measured passage times are rounded to seconds and therefore the



passage times are unknowns within some bounds. Problem (1) can be solved using a least-squares approach as explained next.

Let τ_k denote the realized entrance time of section k (in seconds). Hence, measurement t_k is just τ_k rounded down in seconds and so $\tau_k \in [t_k, t_k + 1)$. Then the problem is finding variables v_k ($k = 3D1, \dots, n$), a_k ($k = 3D1, \dots, n - 1$), and τ_k ($k = 3D1, \dots, n$) with $\tau_k \in [t_k, t_k + 1)$ such that

$$J_k(v_k, a_k, \tau_k, \tau_{k+1}; x_k) := 3Dx_k + v_k(\tau_{k+1} - \tau_k) + \frac{1}{2}a_k(\tau_{k+1} - \tau_k)^2$$

is equal to x_{k+1} for all $k = 3D1, \dots, n - 1$. Note that x_k is a known parameter in $J_k(\cdot)$ and not a variable. Posed this way, the problem has some redundancy: it is sufficient to find the initial speed v_1 and actual passage time τ_k at the first section, and then subsequently in each step k the appropriate variables a_k and τ_{k+1} , which then also fixes the speed v_{k+1} by the second equation in (1). The decision variables can therefore be reduced to the $2n$ variables $(\tau_1, \dots, \tau_n, v_1, a_1, \dots, a_{n-1})$.

The problem can now be defined as a least-squares optimization problem that minimizes the sum of squared errors of $x_{k+1} - J_k(v_k, a_k, \tau_k, \tau_{k+1})$, while satisfying $\tau_k \in [t_k, t_k + 1)$ for all k . Instead of explicitly adding the passage time constraints to the optimization problem it is more convenient from an algorithmic point of view to extend the objective function with additional terms that penalize large deviations from the actual passage times. Consider therefore the following least-squares optimization problem

$$\begin{aligned} \text{Minimize} \quad & \sum_{k=3D1}^{n-1} (w_1 (x_{k+1} - J_k(v_k, a_k, \tau_k, \tau_{k+1})))^2 \\ & + \sum_{k=3D1}^n \left(w_2 \left(\tau_k - t_k - \frac{1}{2} \right) \right)^2 + \sum_{k=3D1}^{n-2} (w_3 (a_{k+1} - a_k))^2 \quad (2) \end{aligned}$$

subject to $v_{k+1} = 3Dv_k + a_k(\tau_{k+1} - \tau_k)$ for $k = 3D1, \dots, = n - 1$. This constraint implicitly determines the speeds v_k ($k = 3D2, \dots, n$) in terms of passage times and acceleration, and only serves for notational convenience. In the actual implementation these (equality) transformations are substituted directly in the cost function. The first term in the cost function penalizes the deviation of the section position x_{k+1} and the estimated train position at time τ_{k+1} given by $J_k(v_k, a_k, \tau_k, \tau_{k+1})$. The second term penalizes the deviation of the actual passage time τ_k from the expected passage time $t_k + 1/2$. Recall that a rounded measurement t_k is in fact a uniformly distributed stochastic variable on $[t_k, t_k + 1)$ with mean $t_k + 1/2$. The third term models that jumps in acceleration or deceleration are as small as possible over subsequent sections and is added to direct the least-squares algorithm towards satisfying solutions. The coefficients w_1, w_2 and w_3 are weight factors to balance the various terms. Note that an absolute deviation of 1 s of the passage time estimates τ_k from the measurements is allowed. However, a jump of 1 m/s² over subsequent accelerations is unacceptable. Therefore the acceleration jump terms can be multiplied by a weight of for instance



$w_3 = 3D100$. In this case a change in value of a passage time τ_k of 1 second is in the same order as a change of 0.01 m/s^2 in acceleration jump (assuming $w_2 = 3D1$). The error of the first term should be zero and therefore w_1 should be large compared to the other two weights.

2.2 Limitations of the proposed methodology

2.2.1 Measurement precision

Track circuit occupation times are rounded down to the full second. This means that the actual time difference between two adjacent sections, which determines the average speed on this section, may (in the extremes) be a second shorter or longer than the measured time difference. This has special importance on sections with very short occupation times, e.g. very short sections or sections passed with high speed.

With current train detection equipment available on the market, a higher precision can be achieved. It is however questionable whether this improvement can also be used in practice at reasonable cost or whether other sources of errors (e.g. transmission delays) will then dominate the quality of the data.

2.2.2 Assumption of a single acceleration rate per section

In the proposed approach, a single acceleration rate per section is assumed. In particular, it is assumed that a train does not stop at intermediate sections. In some situations this assumption causes severe instability to the filtering algorithm, one of which is illustrated in Fig. 2a. Here, a train slows down with constant deceleration rate at the beginning of section 2 and comes to a complete standstill, which corresponds to a second acceleration rate of 0 within this section.

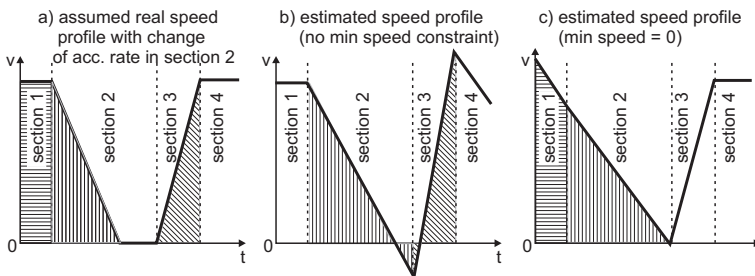


Figure 2: Three speed profiles which all fulfill the constraints of constant section lengths (areas under the curves) and passage times at section borders.

The algorithm then tries to match section occupation times and section lengths to a train behaviour with only one acceleration rate. Part b) and c) of Fig. 2 show solutions, which could have been produced by the algorithm, but do not correspond to the real train behaviour: In order to match section lengths and occupation times with only one acceleration rate, solution b) proposes negative speeds. If one tries

to avoid these irregular values by adding a penalty for negative speeds to the cost function, the estimated speed before the section with two acceleration rates will diverge significantly from the real train behaviour (cf. Fig. 2c). The wrong estimations can moreover influence sections close to the area where the switching occurs, although the oscillations generally become smaller at larger distance.

Choosing a higher weight factor for acceleration changes w_3 , the algorithm can find solutions, which do not contain those strong oscillations. As a consequence, the obtained solution will most probably have higher (and sometimes negative) time differences (term 2 in eq. 2).

The shorter a section is, the less is the probability of a change of regime within this section, so the less the influence of this error becomes.

3 Application

3.1 Test line

The RouteLint DIS was tested on the four-track railway line between Rotterdam and Dordrecht. This 20 km long line is used by six intercity and international lines and three regional train lines and has an important number of freight trains due to the connection to the Kijfhoek freight terminal which is the most important of its kind in the Netherlands. There are two important elevation changes on this line: The Willemspoor tunnel near Rotterdam and the bridge over the Oude Maas just before Dordrecht, which both have gradients up to 17‰.

The whole line comprises more than 400 track circuits (including different variants around switches/ track crossings). In the larger stations, the average length of track circuits is 110 m, while on the open track (including 5 minor stations and several junctions to the freight terminal) this is around 320 m. The shortest track circuit length was 25 m and some aggregated sections were up to 1000 m length.

3.2 Software design

In order to be able to efficiently compute different speed profiles, a dedicated software tool was developed in MATLAB to

1. integrate all the different data sources and prepare the filtering,
2. filter the data according to the algorithm described in sect. 2.1,
3. output major results graphically and as aggregated data in html-format.

The information flows between the different data sources are illustrated in Fig. 3. Data procurement as well as logically connecting the different data sources represented the main part of the work. For instance, track circuit positions had to be identified manually from plans because they were not available elsewhere, including all the different paths at track switches and crossings. Gradient data was available digitally, but had to be matched to the circuit positions via a track number, because different tracks had a different height profile on the examined line. Finally, sections had to be identified, where trains were allowed to stop, because the filtering algorithm can not work correctly if stops are not correctly identified due to



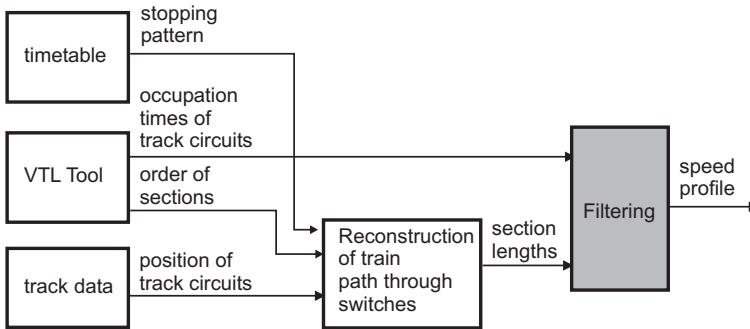


Figure 3: Implemented system components for speed reconstruction energy consumption estimation.

the assumed constant acceleration rate per section (cf. section 2.2). Those sections had to be matched with the stopping pattern of the train; in the case of long trains, the last section before the stop and the first section after the stop had to be identified.

3.3 Application for the RouteLint real-world tests

The proposed methodology was applied successfully for several hundred different trains riding on the corridor. The weighting factors have been chosen as $w_1 = 3D1000$ (per m), $w_2 = 3D1$ (per s) and $w_3 = 3D100$ (per m/s^2) respectively.

Because of the short section lengths of the examined line, the errors described in section 2.2 were only of minor importance: less than 5% of the examined trajectories contained the oscillations described there, mainly because of unplanned stops, but also because of very short sections (less than 100 m) passed with high speed (>120 km/h), that means occupation times of less than 4 secs (high influence of rounding error).

3.3.1 Passenger trains

Among the passenger trains, the Intercity trains which only stop in Rotterdam Centraal and Dordrecht are most likely to get into conflicts with the slower traffic on the route (regional passenger trains or freight trains). Fig. 4 shows the estimated trajectory of an IC travelling in southbound direction getting into a conflict situation with a single locomotive coming from the Kijfhoek.

The trajectory corresponds to the notes of the accompanying project ambassador. It says: "After a crossing train was visible on RouteLint, the driver slowed down the train." This happens at 35.5 km approximately: The train is slowed down and can approach the conflict area with a speed of 30-50 km/h. Not anticipating, it would probably have continued to go with maximal speed and later have come to a full stop, which would have caused more delay and a higher energy consumption.

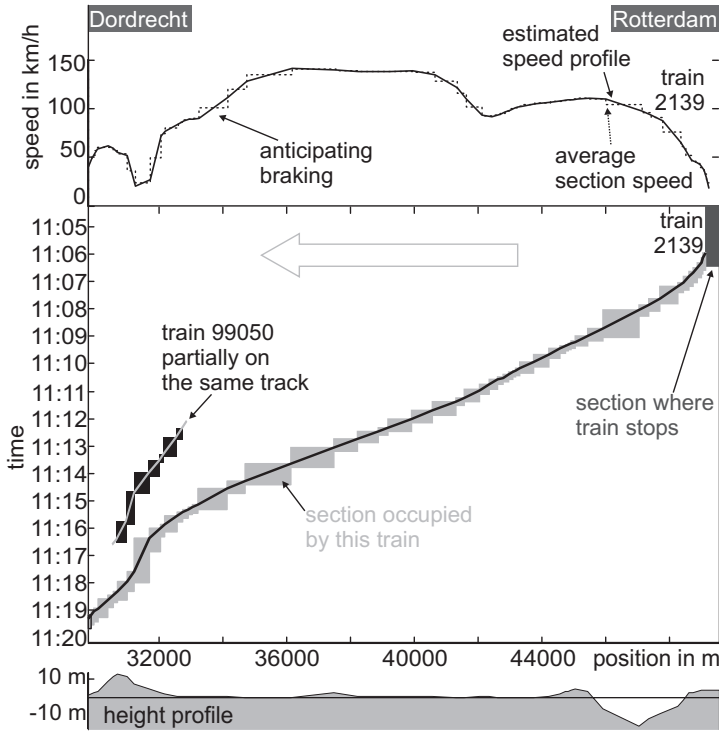


Figure 4: Intercity train with unplanned speed decrease.

3.3.2 Freight train

Fig. 5 shows trajectories of the local freight train 55007 for one train ride with and one without DIS. Due to the track geometry and bad signal visibility at some points of the line, the drivers without information on the signal states further on the line usually drive very conservatively in order not to have to brake the train from a high speed within a short distance. This can be seen from the trajectory without DIS, where the train comes through Dordrecht Station with a speed of around 40 km/h.

The DIS allows to look further ahead and to see that the route of the train is already clear: so with DIS, the driver did not have to ride that conservatively and could pass Dordrecht almost with the maximum allowed speed of 60 km/h saving as much as one minute travel time over the 3 km long section until the Kijfhoek freight terminal, which represents a saving of 20%.

The possibility of having faster freight trains can not be underestimated within the whole, very complex railway system: the infrastructure is occupied less with faster trains, less delays of freight trains lead to less conflicts with other trains, which itself allows reducing consecutive delays and therefore increasing service quality and reducing energy consumption at the same time.

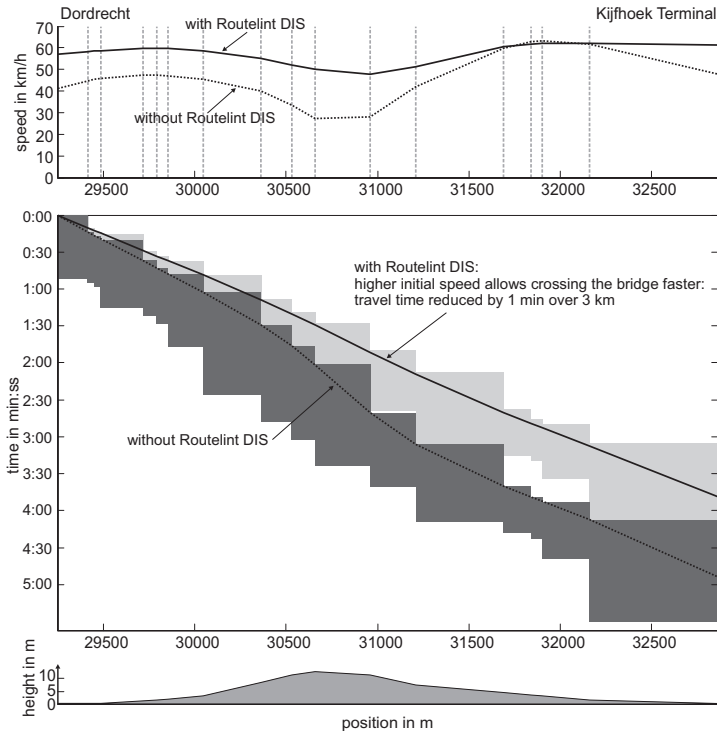


Figure 5: Freight train No. 55007 on 25-11-2005 (without DIS) and 01-12-2005 (with DIS).

This positive effect was also subjectively perceived by the freight train drivers who were most enthusiastic about the project.

3.4 Validation using energy meters

During the tests, five trains were travelling on the corridor which were equipped with energy meters: These devices delivered values for traction power every three seconds. This data has been used to demonstrate the correctness of the speed reconstruction approach. Fig. 6 shows one example of the reconstructed train speed over time for a fast regional train together with the measured traction power of the same train.

It can be seen, that both lines correspond well most of the time, however, there are some points (changes of regimes) where small time differences of less than 10 s can be observed (e.g. start of coasting between Rtd-Lombardijen and Rtd-Blaak).



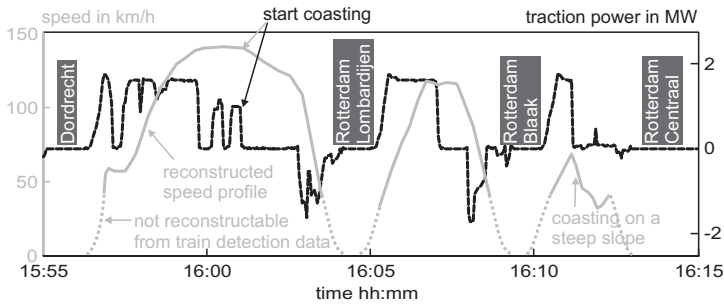


Figure 6: Fast regional train 2250 on 01-12-2005: Comparison of reconstructed speed profile with measured traction power.

4 Conclusions and outlook

The proposed algorithm delivered reasonable speed profiles, which corresponded perfectly with the observations of the accompanying project ambassadors in most of the cases.

For trains with high acceleration and braking rates, the given limitation of a single acceleration rate per section may cause oscillations which can be corrected with expert knowledge. Future research should focus on how to take that into account (varying number of acceleration changes per section, possibly under utilization of evolutionary algorithms as proposed for the computation of an unknown number of switching points for energy optimal train operation on sections with multiple gradients/ speed limits [4]).

Except for the reconstruction of train trajectories, the proposed methodology could also be used as an online application to compute current train speeds in a traffic control centre: This information should be valuable to the traffic controllers for train disposition in conflict situations and is available without having to install expensive measurement and communication equipment on board of every train.

References

- [1] Albrecht, T. & van Luipen, J., What role can a driver information system play in railway conflicts? *Paper submitted for the 10th IFAC Symposium Control in Transportation Systems*, Delft, The Netherlands, 2006.
- [2] Goverde, R.M.P., *Punctuality of Railway Operations and Timetable Stability Analysis*. Ph.D. thesis, Delft University of Technology, 2005.
- [3] Goverde, R.M.P. & I.A. Hansen, 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*, pp. 779-788, WIT Press, Southampton, 2000.
- [4] Chang, C. & Sim, S., Optimising train movements through coast control using genetic algorithms. *IEE Proc.-Electr. Power Appl.*, **144(1)**, pp. 65–73, 1997.