Fact-checking of timetabling principles: a case study on the relationship between planned headways and delays

Fabrizio Cerreto [a,1], Megan Holt Jonasson [a]
[a] Traffic Division, Rail Net Denmark
Carsten Niebuhrs gade 49, 1577 Copenhagen, Denmark
[1] E-mail: fceo@bane.dk, Phone: +45 82349172

Summary

Railway schedules are often planned after timetabling principles derived from practical experience and based on the macroscopic limitations of a system, rather than the microscopic conflicts inherent in its signaling system. This inaccuracy in planning principles can lead to infeasible timetables, which induce delays and thus reduce the service reliability of railway transport. The objective of this paper is to support the design of fact-based timetables by introducing a systematic analysis of historical data from operations on the Danish railway, which includes testing the principles used in the timetable design, specifically regarding the allocation of headway time, and identifying possible improvements to the scheduling of trains according to their specific characteristics. The data records used in this analysis are generated by the signaling system and the automatic train detection system. The records state the scheduled and realized times of each train at every timing point on the network, as well as, categorical information about the trains and the measuring points. The timestamps are rearranged by an automatic algorithm to calculate the scheduled headways and the change in deviation between consecutive trains at the timing points. The minimum feasible headways between pairs of movements are then identified through linear regression of the largest reduction in delay between consecutive trains, as a function of the planned headway. This analysis supports the identification of the adjustments to the timetabling principles. As a result, the generation of secondary delays in operation is reduced, and the possibility of delay recovery under disturbed operation is improved.

Keywords
Railway Delays, Headways, Timetables, Data Analysis, Train detection systems

Extended abstract

1.1 Problem

Railway schedules are often planned after timetabling principles derived from practical experience. The knowledge of the practitioners and general rules-of-thumb drive the design of the structural parameters of a timetable. For instance, the headway on a railway line is often given as a standard value of minimum time separation between consecutive services from the origin station (Andersson, et al., 2011). However, this value does not consider the actual blocking times of the line sections and, therefore, the minimum technical feasible headway. This approach feeds the risk of systematic interferences between trains as the potential conflicts are only considered at a macroscopic level (i.e. between stations) in lieu of the block section level. Indeed, this inaccuracy typically leads to infeasible timetables, and these hidden microscopic conflicts induce delays that reduce the service reliability of the railway transport and its attractiveness for the passengers (Palmqvist, et al., 2018). The increasing availability of data collected by automated systems provides the opportunity for verification of the experience-based timetabling principles. Thus, the amount of hidden conflicts in the timetables can be reduced according to the observed performance of realized operations.

The objective of this paper is to introduce a method of systematic analysis of historical data from operations to support the design of fact-based timetables. As a result, the robustness of the schedules can be improved, together with the service reliability and its attractiveness for passengers.

1.2 Literature

The analysis of records from realized operations found several applications in railways. The managerial approach typically focuses on the identification of critical patterns that should be corrected. Richter and Schittenhelm provided a series of articles describing the data collection system in Denmark and methods and measures for the data analysis. The authors highlighted the potential of the integration of historical records from past operations into the planning phase of future schedules. Such a feedback loop would improve the reliability of the railway service (Schittenhelm & Richter, 2009). A ranking method was proposed to identify the worst performing trains and line sections, while other statistics were used to describe the development of delays along these train journeys (Richter, 2012). Cerreto et al. introduced systematic analyses for large datasets to study the development of delays along the train journeys, highlighting reoccurring delay patterns that should be tackled to improve the service reliability of the line section (Cerreto, et al., 2016; 2018). Big data techniques (i.e. association rules) have also been utilized in mining the causes of delays in urban railways, (Yabuki, et al., 2015). Detailed data has also been used in literature to identify knock-on delays due to conflicts in the schedules. (Goverde & Meng, 2011; Daamen Winnie and Goverde, 2009)

Research in urban-bus transportation was carried out by van Oort et al. (2015), where the development of delays en-route was compared to the realized headways to identify congestion. However, the research on the relationship between the planned headways and the realized delay distributions is still rather limited in railways.

1.3 Methodology

The relationships between planned headways, minimum feasible headways, headway buffers and delay changes, are investigated in this paper. The relationship between the planned headways and the changes in the deviation identify the headway buffer as a function of the planned headway. Such headway buffer should be large enough to reduce the transfer of delays to the consecutive trains, and small enough to allow for the desired service frequency to be scheduled. The minimum feasible headways are found through linear regression of the headway buffers as a function of the planned headway, where the minimum feasible headway corresponds to the planned headway that gives zero buffer.

The data manipulation and the statistical analyses are performed through the open source software R 3.5.1 and RStudio 1.1.456, by The R Foundation for Statistical Computing, Vienna, Austria, and by RStudio, Inc, Boston, MA, USA, respectively.

1.4 Kind and source of data

A dataset of train timestamps from January to November 2018 on the Danish railway network is investigated in this research. The records are generated by the signaling system and the automatic train detection system, and state the scheduled and realized times of the trains at every timing point on the network. The records include information about the operations and the measuring points, such as timing point name, specific blocking section, train ID, train category, scheduled time, and recorded deviation. Under the automatic train detection system, the arrival time of trains is given by the occupation of the platform track circuit, while the departure time is given by the occupation of the first main line track circuit. The time granularity system is 10s and includes a correction factor to account for the time elapsed between the occupation of a platform section and the actual stop of the train for passenger service. Similarly, correction factors for departure and pass-through records are provided after calibration with GPS measurements.

The timestamps are rearranged by an automatic algorithm to calculate the scheduled and realized running times, the scheduled and realized headways at the timing points, and the change in deviation between consecutive timing points for each train run.

1.5 Practical relevance of the findings

This research uses historical data to test the principles used in practical timetable design. The relationship between planned headways, minimum feasible headways, headway buffers and recorded delays is investigated from the train timestamps automatically generated by the signaling system. The results support the improvement of the schedules through a fact-based planning of the process times and buffers, as opposed to the current tradition of experience-based planning. It will be possible, therefore, to tailor the schedules to the individual trains, according to their specific characteristics, and to those of the preceding trains. The expectation is to reduce the generation of secondary delays in operation, and to improve the possibility of delay recovery under disturbed operation.

References

Andersson, E., Peterson, A. & Törnquist Krasemann, J., 2011. *Robustness in Swedish Railway Traffic Timetables.* Rome, s.n., pp. 1-18.

Cerreto, F., Nielsen, B. F., Nielsen, O. A. & Harrod, S. S., 2018. Application of Data Clustering to Railway Delay Pattern Recognition. *Journal of Advanced Transportation,* Volume 2018.

Cerreto, F., Nielsen, O. A., Harrod, S. & Nielsen, B. F., 2016. *Causal Analysis of Railway Running Delays.* Milan, Italy, World Congress on Railway Research, pp. 1-7.

Daamen Winnie and Goverde, R. M. P. H. I. A., 2009. Non-Discriminatory Automatic Registration of Knock-On Train Delays. *Networks and Spatial Economics,* 3, 9(1), pp. 47-61.

Goverde, R. M. P. & Meng, L., 2011. Advanced monitoring and management information of railway operations. *Journal of Rail Transport Planning and Management,* 1(2), pp. 69-79.

Palmqvist, C.-W., Olsson, N. O. E. & Hiselius, L. W., 2018. The Planners' Perspective on Train Timetable Errors in Sweden. *Journal of Advanced Transportation,* Volume 2018, p. 17.

Richter, T., 2012. *Data aggregation for detailed analysis of train delays.* s.l., WITPress, pp. 239-250.

Schittenhelm, B. & Richter, T., 2009. *Railway Timetabling Based on Systematic Follow-up on Realized Railway Operations.* Aalborg, Traffic Research Group, Aalborg University.

van Oort, N., Sparing, D., Brands, T. & Goverde, R. M. P., 2015. Data driven improvements in public transport: the Dutch example. *Public Transport,* 7(3), pp. 369-389.

Yabuki, H., Ageishi, T. & Tomii, N., 2015. *Mining the Cause of Delays in Urban Railways based on Association Rules.* Rotterdam, s.n., pp. 1-16.