PRISM: A Macroscopic Monte Carlo Railway Simulation

Markus ZINSER[1], Torsten BETZ[1], Maurice BECKER[1], Michael GEILKE[1], Carla TERSCHLÜSEN[1], Adam KALUZA[1], Ingrid JOHANSSON[2], Jennifer WARG[2]

[1]Deutsche Bahn AG, DB Analytics, Poststraße 20, 60329 Frankfurt am Main, Germany
[2]KTH Royal Institute of Technology, Brinellvägen 23, 100 44 Stockholm, Sweden
Corresponding Author: Markus Zinser (markus.zinser@deutschebahn.com)

Abstract

The current state-of-the-art in timetable analysis in the presence of disturbances is to use railway microsimulation, which typically yields detailed results on infrastructure or timetable performance. However, microscopic simulation is time-consuming and requires a detailed infrastructure model and expert knowledge to run adequately. This paper outlines a macroscopic approach which aims at reducing execution time by restricting the level of detail. A first prototype was developed which is able simulate one day of operation on the German railway network with about 40,000 trains in less than a minute. The approach has been studied in detail on that network. Several years of operational data have been analyzed to generate input for the macroscopic simulation. In the results, we describe our analysis tools and compare distributions of the simulation to operational data. It was found that the simulation describes real operational train data very well.

Keywords: railway, simulation, timetable, disturbances, quality

1. Introduction

The ability to quickly evaluate the effects of different options is essential in railway timetable planning. However, difficulties arise in assessing the robustness of a proposed timetable with regard to possible infrastructure disturbances and non-availability. Simulating the impact of disturbances and infrastructure non-availability on train travel times is one approach to evaluate timetable robustness. Such a simulation can be performed at different levels of detail. In a microscopic simulation, exact train paths through the network are simulated and infrastructure disturbances can be modeled at the level of individual switches or signals.

By contrast, in a macroscopic simulation only certain aspects of the network are included, e.g. stations and stops, but no detailed track plan between these network nodes. In this case, a stochastic approach is required because failures of different infrastructure elements of the same type on a specific macroscopic network edge can have different consequences for train travel times. This approach ignores the causal connection between disturbance, re-routing, possible waiting times and the resulting delay, and simply assigns a probability to different delay times given a type of disturbance.

The current state of the art railway simulations are not able to simulate large networks within a short runtime. This paper describes a macroscopic Monte Carlo railway simulation named PRISM (Plasa Railway Interaction Simulation Model). It is both able to simulate a large network within a short runtime (~min) and at the same time leading to realistic results which represent a typical day of operation. Reliability is an important measure of a railway system´s performance. With the help of simulation, infrastructure, trains and timetables can be modeled and their operation be estimated. A review of available simulation models for train movement, power supply systems and traction drives was done by Goodman, Siu and Ho in 1998 [1], while a train simulator for general purposes was proposed by Ho et al. [2] four years later, and the development has continued.

Microscopic simulation models are commonly used. They have a high level of detail, typically including the exact station layouts, placement of switches and signals, etc., in order to represent reality well. Borndörfer et al. [3] recommend dynamic, synchronous, microscopic, stochastic simulation to represent

the system in the best way. There are several microscopic railway simulation tools available, e.g. the commercial alternatives RailSys [4], [5] and LUKS [6], while OpenTrack is an open source alternative [7]. The amount of detail in microscopic models, however, results in long simulation times when larger networks are considered and increases the complexity of coding and handling of the models. Therefore, macroscopic simulation models with a lower degree of detail can be preferable. For example, Büker and Seybold [8] presented a macroscopic model for delay propagation in large networks. The required accuracy of the model highly depends on the task. In many cases, macroscopic simulation is sufficient, while many others require proper modeling of the system. Cui et al. [9] present a model which can be adjusted according to the user´s and project´s needs, with the possibility of micro-, meso- and macroscopic simulation.
The model presented in this article describes a macroscopic approach, first presented by Zinser et al. [10] and later called PRISM. The level of details used in the simulation can be adjusted so that it can be called a micro-macro model combining the advantages of both, always in accordance with the needs. The paper is organized as follows. Section 2 describes the required input data for the model, while section 3 describes the principle for the simulation together with detailed descriptions of the different model components. Section 4 presents the results from simulation of the German railway network, while section 5 finishes the paper with conclusions.

2. Input data

The simulation model PRISM requires as minimal input an infrastructure representation, trains, and a timetable. These are described below.
The simulation is based on a macroscopic railway infrastructure representation consisting of nodes and edges. The level of detail in the railway representation can in principle be freely chosen. However, the nodes are typically chosen to be equivalent to the operating stations while the edges link these. This means that between two nodes there are edges with a length ranging from a few hundred meters up to several kilometers. In addition, basic infrastructure properties of the nodes and edges are needed. The nodes need information whether additional tracks for overtaking are available or not. The edges need information about the train protection system (e.g., which system, distances between signals, etc.) and basic information on how many tracks an edge has, whether it is electrified, and the maximum speed allowed.
Train properties are specified via so-called model trains. A model train describes a certain group of trains which have similar properties like maximum speed, safety systems, length, weight, and acceleration/deceleration.
Besides the infrastructure and train information, a timetable is necessary. This timetable needs to specify which type of train is passing which series of infrastructures nodes and what the planned times for arrival and departure are. Hereby, the granularity of the timetable needs to match the granularity of the infrastructure.

3. The methodology

In the following section, the simulation principle is described, followed by a more detailed description of the modelling of the various components of a train run.
The PRISM model is a discrete-event simulation. Train runs are modelled as consecutive departure and arrival events, where each event corresponds to the arrival or departure on one of the infrastructure nodes. When the event is processed, the time of the next event is already calculated. This means that for a train running on the edge (A, B) the arrival time at B is already calculated when the train departs at A. If no other train is interfering with the departure (see Section 3.1) and if the train is allowed to depart (see Section 3.2), the arrival time at the next station is calculated using the following formula:

$$arrival_{train1,nodeB} := departure_{train1,nodeA} + travel\_time(train1_{nodeA \rightarrow nodeB}) + \sum_{cause} disturbance(traintype, time, cause) - delay\_reduction(train1_{nodeA \rightarrow nodeB}) + random\_deviation(train1_{nodeA \rightarrow nodeB})$$

Here, the travel time is taken from the timetable. The additional terms which are leading to either delay or delay reduction are briefly discussed in the following paragraphs. If no other train is interfering with the arrival, the departure time of a train at a certain station is determined in a similar way:

$$departure_{train1,nodeB} := arrival_{train1,nodeB} + dwell\_time(train1_{nodeB}) + dwell\_time\_deviation(dwell\_time, traintype, time, node)$$

The dwell time is also taken from the timetable. The dwell time deviation, which covers both delays and dwell time reduction to reduce delay is discussed in the following paragraphs.

3.1 Modelling of train conflicts

Train conflicts (interferences) are modelled based on minimum headway times. When the arrival or departure time of a train is calculated, the distance to the previous train on the same segment is checked. The following train will be further delayed if the temporal distance is smaller than the minimum headway time. The calculation of minimum headway times is currently implemented for signal-based systems only and happens during runtime. It can be calculated from the components of the block occupation times, which consist of the route building time, the approaching time, the signal sight time, the travel time through the block, the evacuation time of the block, and the route releasing time. If a conflict is detected, the corresponding event will be repeatedly processed until no further delays are transferred.

3.2 Modelling of dispatching

The simulation is capable of switching the order of trains at nodes which are flagged as stations with an overtaking possibility, at which all trains that will pass the station within the next 30 minutes are listed. When processing the departure event of a train, the simulation performs a simple conflict detection by checking each train pair listed in the station for interference conflicts up to the end of the common path or until the station with the next overtaking possibility is reached. For two conflicting trains, several decision schemes are implemented to decide whether and, if yes, how the conflict should be resolved. The simplest scheme is based on dispatching ranks which can be given to each model train. Parameters like the minimum transferred delay until the end of the common path or the minimum transferred delay at the conflicting point can be set. Other dispatching schemes take into account more information of the timetable, the train types or the current delay of the trains. If the decision is taken that the departing train should wait for a conflicting train to pass, its departure will be delayed and reprocessed at a later stage. Currently, there is no possibility to change the train route via dispatching within the simulation.

3.3 Modelling of disturbances

Disturbances in PRISM are simulated via a Monte Carlo method. The delay of a train at the start of its run (or entrance in the simulated network) is sampled from histograms which are obtained from operational data. These histograms are specified for every station of the infrastructure and can be further specified for every train type and different time windows. For every travel time between two nodes during a train run, random sampling determines whether the train will be affected by a disturbance and, if so, how much delay will be incurred. The simulation differentiates between disturbances that affect one specific train and disturbances that affect a piece of infrastructure and thus

all trains passing that infrastructure. For each infrastructure node and edge, the probability for the occurrence of a disturbance is determined from operational data. If a disturbance occurs, the resulting delay will be sampled from a Weibull distribution which has been fitted to operational data. These parameters can be easily modified to allow for scenario simulations.
In addition to delays caused by disturbances, a further small deviation from the planned travel time is sampled from a normal distribution, which has again been adjusted to operational data. This small deviation is meant to cover 'random' effects, meaning small deviations due to weather conditions or the like.
Similar to the start delays, dwell time deviations are sampled from histograms which are obtained from operational data. Hereby, only delayed trains are considered, reducing the data to the subset of trains which potentially reduce their dwell time. These dwell time deviation histograms are specified for every station and different ranges of planned dwell times. In addition, they can be further split into train types and time windows. The dwell time deviations are sampled prior to the simulation start, implying every train is trying to reduce its dwell time to reduce delay. If the sampled dwell time deviation is leading to a departure time prior to the planned departure time, it is discarded during simulation, and a new dwell time deviation is sampled from a normal distribution which is fitted to operational data.

3.4 Modelling of construction sites

Three types of construction sites are currently considered in the simulation: closure of one track on a double track line, reduction of the speed limit, and closure of all tracks. The latter is not explicitly modelled but expected to be covered by the timetable. The closure of a single track on a double track line is done by simply reducing the number of tracks to one, thereby only allowing for simultaneous railway traffic into one direction on that track. The reduction of the speed limit is simulated by calculating the time a train needs to travel the section. The calculation takes into account the time to break from the nominal to the reduced speed, to traverse the segment with the reduced speed limit, and to accelerate to nominal speed again. If this time is longer than the planned travel time, a delay is added.
A construction site is defined by referring to an infrastructure edge, having a length (in case that the part with the reduced speed is shorter than the actual infrastructure edge), and a time at which the infrastructure starts and ends. By using this definition of construction sites, unplanned infrastructure modifications such as segments with reduced speed limits due to infrastructure defects can also be simulated.

3.5 Modelling of delay reduction

A delay reduction component allows trains to utilize timetable buffers and supplements and potentially mitigate the effect of disturbances. The minimal technical driving time is necessary to calculate the amount of supplement a train can use to reduce its delay. This minimal technical driving time can be either calculated by a microscopic simulation or be obtained from operational data. In our approach the latter is done by using the percentile of the driving time distributions (the type of percentile is determined by the amount of data available). For each infrastructure edge, each train type, and different driving modes a distribution is estimated. The driving mode categorizes trains which are accelerating, decelerating or driving with constant speed by checking the distance to the last stop. The amount of available delay reduction potential is calculated by subtracting the minimal driving time from the planned time. This actual reduction potential is again smeared during runtime of the simulation by using a reversed truncated Weibull function. This covers the effect that the minimal driving time underlies minor fluctuations due to changing technical conditions (weather, train traction, train weight, driver behavior, etc.).

4. Results

A fully functional simulation prototype is provided as a package for the programming language R. However, the core simulation is implemented in C++, since it underlies strong performance requirements due to the request to simulate large railway networks within short runtimes. With the

current implementation, runtimes of about a minute are achieved for simulating a whole operational day in Germany (about 40,000 trains). The majority of runtime is, hereby, caused by reading in and writing out the data. The actual simulation takes place in about ten seconds.
The PRISM results are compared in detail to operational data of the German railway system to evaluate its accuracy. The comparison is systematically used to detect shortcomings in the modelling and to evaluate how accurate the simulation can describe delay related observables. For the systematic analysis of the simulation results, a dashboard has been developed in which several analyses are integrated, ranging from very general results, like punctuality, down to train related results like average delays. Figure 1 shows two pictures of this dashboard.

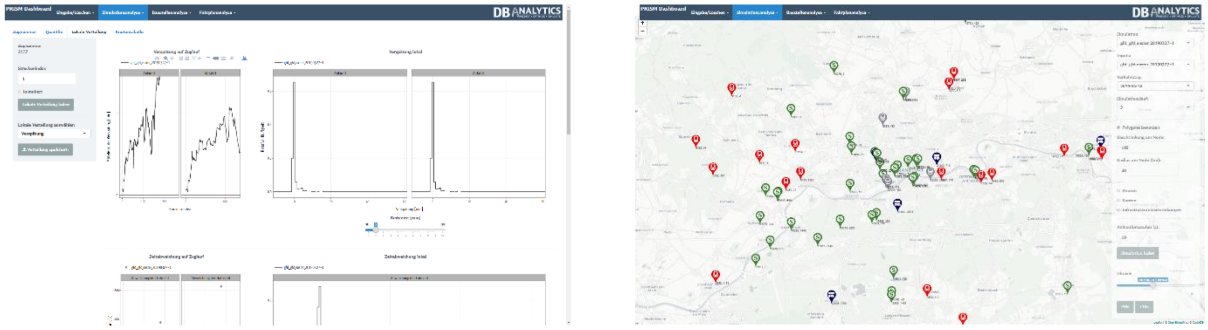

Fig. 1: Dashboard to analyze the simulation results. The left-hand side shows a picture of a detailed view which allows to analyze individual trains and how their average delays develop along the train run. The right-hand side shows a map which can replay a specific simulation run.

Figure 2 shows two examples of comparisons between PRISM simulation results and operational data. On the left-hand side, a histogram of the average build-up of extra delay for long-distance trains on infrastructure edges is shown. On the right-hand side, a similar histogram is shown for the average delay reduction. Both operational data (black) and the simulation results (red) are shown. The simulated data is the result of simulating all trains on the entire German network for all days in 2017. Both distributions are in good agreement, showing the ability of PRISM to yield realistic results.

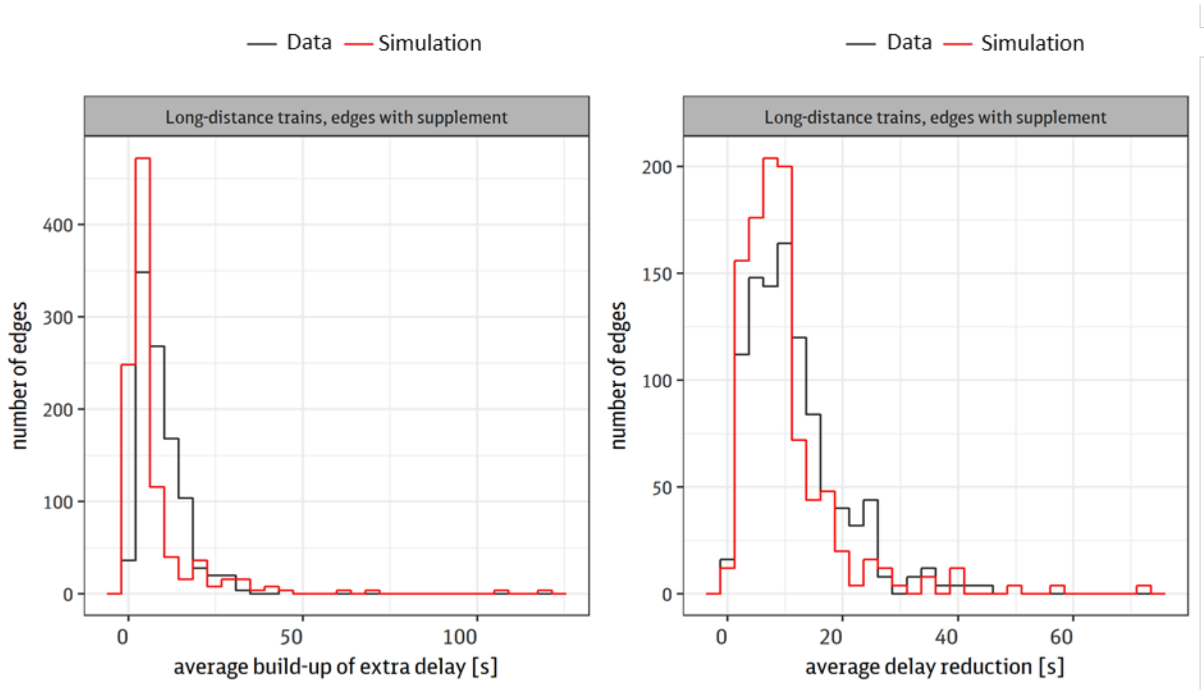

Fig. 2: Average build-up and reduction of delay of long-distance trains.

5. Conclusions

The ability to quickly evaluate the effects of different options is essential in railway timetable planning. However, current simulation methods are often complex or have too long runtimes to allow for quick responses. A novel macroscopic railway simulation method which makes strong use of operational data has been presented. Due to short runtimes of less than a minute for 24 hours of train operation for a large network (~40,000 trains), this method aims to close this gap.
This railway simulator can be used to easily study different scenarios in large railway networks. The impact of construction plans, reduced disturbance probabilities or alternative dispatching rules on KPIs like punctuality can be studied. Having a good estimation of the impact of a measure on the operational performance helps to prioritize these and, therefore, to improve and understand railway operations in general.
The method has been implemented and studied in detail for the German railway network. A detailed comparison to operational data was performed to prove its functionality.

Acknowledgment

This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation program under grant agreement No 730814 and No 826151.

References

[1] C. Goodman, L. Siu and T. Ho, "A review of simulation models for railway systems," in *International Conference on Development in Mass Transit Systems*, London, 1998.

[2] T. Ho, B. Mao, Z. Yuan, H. Liu and Y. Fung, "Computer simulation and modeling in railway applications," *Computer Physics Communications,* vol. 143, no. 1, pp. 1-10, 2002.

[3] R. Borndörfer, T. Klug, L. Lamorgese, C. Mannino, M. Reuther and T. Schlechte, Eds., Handbook of Optimization in the Railway Industry, Springer, 2018.

[4] J.-P. Bendfeldt, U. Mohr and L. Müller, "RailSys, a system to plan future railway needs," *WIT Transactions on the Built Environment,* vol. 50, pp. 249-255, 2000.

[5] A. Radtke and D. Hauptmann, "Automated planning of timetables in large railway networks using a microscopic data basis and railway simulation techniques," *WIT Transactions on The Built Environment,* vol. 74, 2004.

[6] D. Janecek and F. Weymann, "LUKS-Analysis of lines and junctions," in *Proceedings of the 12th World Conference on Transport Research (WCTR)*, Lisbon, 2010.

[7] A. Nash and D. Hürlimann, "Railroad Simulation using OpenTrack," *Computers in Railways IX,* pp. 45-54, 2004.

[8] T. Büker and B. Seybold, "Stochastic modelling of delay propagation in large networks," *Journal of Rail Transport Planning and Management,* vol. 2, no. 1, pp. 34-50, 2012.

[9] Y. Cui, U. Martin and J. Liang, "PULSim: user-based adaptable simulation tool for railway planning and operations," *Journal of Advanced Transportation,* vol. 2018, 2018.

[10] M. Zinser, T. Betz, J. Warg, E. Solinen and M. Bohlin, "Comparison of microscopic and macroscopic approaches to simulating the effects of infrastructure disruptions on railway networks," in *Proceedings of 7th Transport Research Arena TRA 2018*, Vienna, Austria, 2018.