Transportation Research Part C 95 (2018) 599–615

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Stochastic prediction of train delays in real-time using Bayesian networks☆

Francesco Corman[a,*], Pavle Kecman[b]

[a] Institute for Transport Planning and Systems, ETH Zurich, Switzerland
[b] Department of Science and Technology, Linköping University, Sweden

ARTICLE INFO

Keywords:
Bayesian networks
Prediction
Railway traffic
Stochastic processes
Train delays

ABSTRACT

In this paper we present a stochastic model for predicting the propagation of train delays based on Bayesian networks. This method can efficiently represent and compute the complex stochastic inference between random variables. Moreover, it allows updating the probability distributions and reducing the uncertainty of future train delays in real time under the assumption that more information continuously becomes available from the monitoring system. The dynamics of a train delay over time and space is presented as a stochastic process that describes the evolution of the time-dependent random variable. This approach is further extended by modelling the interdependence between trains that share the same infrastructure or have a scheduled passenger transfer. The model is applied on a set of historical traffic realisation data from the part of a busy corridor in Sweden. We present the results and analyse the accuracy of predictions as well as the evolution of probability distributions of event delays over time. The presented method is important for making better predictions for train traffic, that are not only based on static, offline collected data, but are able to positively include the dynamic characteristics of the continuously changing delays.

1. Introduction

Accurate prediction of train delays (deviations from timetable) is an important requirement for proactive and anticipative real-time control of railway traffic. Traffic controllers need to predict the arrival times of the trains within (or heading towards) their area in order to control the feasibility of timetable realisation. Similarly, the transport controllers on behalf of train operating companies may use the predictions to estimate the feasibility of planned passenger transfers, as well as rolling-stock and crew circulation plans. Valid estimates of arrival and departure times are therefore important for preventing or reducing delay propagation, managing connections, and providing reliable passenger information. The difficulty for predicting the train event times comes from the uncertainty and unpredictability of process times in railway traffic. The models for real-time traffic control have so far mostly focused on overcoming the great combinatorial complexity of train rescheduling (Corman et al., 2014b; Meng and Zhou, 2014; Törnquist and Persson, 2007), delay management (Dollevoet et al., 2014) and rolling-stock and crew rescheduling (Nielsen et al., 2012; Potthoff et al., 2010). The developed approaches are able to solve complex instances in real-time, however they typically assume perfect deterministic knowledge of the input traffic state and subsequent traffic evolution.

In recent years, the uncertainty of train event times has been recognised as one of the major obstacles for computing feasible and

☆ This article belongs to the Virtual Special Issue on “Big Data Railway”.

* Corresponding author.
E-mail addresses: francesco.corman@ivt.baug.ethz.ch (F. Corman), pavle.kecman@liu.se (P. Kecman).

https://doi.org/10.1016/j.trc.2018.08.003
Received 23 October 2017; Received in revised form 11 June 2018; Accepted 6 August 2018

implementable solutions for rescheduling problems in railway traffic (Corman and Meng, 2014; Quaglietta et al., 2013). The uncertainty of an event is usually represented by the probability distribution of its realisation. However, most of the existing approaches assume fixed probability distributions for train delays and do not consider the effect that real-time information on train positions and delays may have on (the parameters of) the corresponding distributions. In order to create realistic online tools for real-time traffic management, the dynamics of uncertainty of delays needs to be considered. When new information about train positions and delays becomes available, the uncertainty for predicting subsequent events is typically reduced. The main objective of this paper is to examine the effect that the prediction horizon and incoming information about a running train may have on the predictability of subsequent arrival and departure times of all trains. In other words, we try to give an answer to the question: how does the probability distribution of delay of an event change over time?

In this paper we first describe a method for modelling uncertainty of train delays based on Bayesian networks. Railway traffic is modelled by means of a probabilistic graphical model which exploits conditional independences between events to allow the efficient computation of their joint distribution (Koller and Friedman, 2009). An important advantage of this method in the context of real-time prediction of train traffic is that it allows the information or evidence about a certain event to be propagated. In other words, evidence about realisation of one event affects (reduces) the uncertainty of other events. Therefore, probability distribution of e.g. an arrival delay in a station changes over time in discrete steps as more information becomes available. This can be used by traffic controllers to estimate the probability of a route conflict in their area, probability of arrival delay of a feeder train for a passenger transfer, etc. Moreover, having a better estimate of train delays could be greatly beneficial for validation and evaluation purposes of the state-of-the-art online traffic models. In particular, this approach enables the estimation of delay dynamics for the closed-loop (Corman and Quaglietta, 2015; Caimi et al., 2012), online rescheduling (Gatto et al., 2007; Bauer and Schöbel, 2014) and simulation (Nash and Huerlimann, 2004; Quaglietta, 2014) tools. Finally, even though we focus on the real-time prediction of railway traffic, the modelling framework and methodology presented in this paper could be extended to handle predictions in other scheduled and constrained systems, such as public transport, logistic networks, and supply chains.

The next section gives the description of the problem and a comprehensive literature review. The methodological framework is presented in Section 3, followed by the description of the case study (Section 4) and analysis of results (Section 5). Section 6 summarises the main findings and gives the recommendations for future research.

2. Problem description and literature review

The real-time prediction of railway traffic is one of the main tasks on the operational traffic control level. Trains are operated according to a timetable and a daily process plan. Due to inevitable disturbances and deviations from the planned schedule, train runs need to be continuously monitored. By monitoring we assume keeping track of all performance indicators such as the actual train positions, delays, realised running and dwell times of all trains, etc. Monitoring therefore provides the actual traffic state that can be used to predict the future evolution of traffic on the network. A predictive traffic model needs to continuously provide the controller level with the information about the expected traffic conditions. Moreover, it should enable the controller to evaluate the impact of potential dispatching actions.

Railway prediction models can be classified to static (offline) and dynamic (online), and deterministic and stochastic, depending on the time span between the time they are run, and the operations they aim to predict; and on how they tackle uncertainty, respectively. Deterministic prediction models assume full knowledge of the future traffic evolution (Dolder et al., 2009). Some approaches (see for instance (Burdett and Kozan, 2014; Wei et al., 2015) focus on simulating the traffic based on current state, and determine most likely conflicts, with limited usage of data on past operations. Even though the more advanced data-driven deterministic models are able to explain a large percentage of process time variability using the values of explanatory variables, a certain degree of uncertainty, especially for dwell times, still remains unresolved (Kecman and Goverde, 2015).

Stochastic models attribute each event with a probability distribution in order to model the uncertainty of its realisation. They can be classified based on how they use the real-time information to update their predictions, into static and dynamic. Whereas static prediction models are based on the offline computed probability distributions and their parameters, dynamic models are updated in real-time as new information becomes available. Most of the stochastic delay propagation models (Büker and Seybold, 2012; Medeossi et al., 2011; Meester and Muns, 2007) were used for offline analyses of timetables. A-Posteriori analysis (Yaghini et al., 2013; Lee et al., 2016) focuses instead on understanding factors and root causes of realised delays, and incorporate those into strategic planning and changes to the timetable. A recent contribution that focuses on predicting delays in the process of appraisal of infrastructure investments was presented by Marković et al. (2015). For all those offline approaches, their goal is to determine factors describing the influence of design parameters to a railway system, in a planning phase, i.e. many months or years before operations. The mentioned approaches are inherently static and do not consider the effect that the real-time information obtained from the monitoring system may have on reducing the uncertainty of the future events.

The concept of delay dynamics refers to the degree of information one has available about future events, when online real time information is considered, within the prediction horizon. The idea of delay dynamics is illustrated in Fig. 1. With every update of train delay (arrivals to station A and B) probability distributions of arrival times to subsequent stations (C and D) are updated.

We summarize in Table 1 the most related works of the literature, which are discussed in what follows. In column 2, the works are categorized according to their *approach*, i.e. basic algorithm used to determine the conditional probability of future events based on past data and current events. Column 3 describes the consideration of *dynamic* nature of the problem, i.e. whether the different online predictions over time are linked somehow by parametric relation inherent in the model. For the works mentioned, often dynamics is not mentioned (Böhmová et al., 2015) or is defined a priori based on rules (Bauer and Schöbel, 2014); in general, dynamics within the