Transportation Research Part C 56 (2015) 251–262

Contents lists available at ScienceDirect

Transportation Research Part C

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

Analyzing passenger train arrival delays with support vector regression

Nikola Marković[a,*], Sanjin Milinković[b], Konstantin S. Tikhonov[c], Paul Schonfeld[a]

[a] Department of Civil and Environmental Engineering, University of Maryland, College Park, MD, USA
[b] Faculty of Transport and Traffic Engineering, University of Belgrade, Belgrade, Serbia
[c] Department of Physics and Astronomy, Texas A&M University, College Station, USA

ARTICLE INFO

Article history:
Received 28 November 2014
Received in revised form 5 April 2015
Accepted 5 April 2015

Keywords:
Train arrival delays
Support vector regression
Artificial neural networks
Machine learning
Infrastructure
Statistics

ABSTRACT

We propose machine learning models that capture the relation between passenger train arrival delays and various characteristics of a railway system. Such models can be used at the tactical level to evaluate effects of various changes in a railway system on train delays. We present the first application of support vector regression in the analysis of train delays and compare its performance with the artificial neural networks which have been commonly used for such problems. Statistical comparison of the two models indicates that the support vector regression outperforms the artificial neural networks. Data for this analysis are collected from Serbian Railways and include expert opinions about the influence of infrastructure along different routes on train arrival delays.

1. Introduction

Train delays are causing substantial losses to railway operators and passengers. The National Audit Office in United Kingdom reported that there were about 800,000 delays on the British national rail network during 2006–2007. This led to 14 million train-minutes of delays, which cost the passengers about £1 billion in lost time (Burr et al., 2008; Preston et al., 2009). Such numbers provide evidence that reducing train delays in an economically viable way would be highly desirable. An important step in this direction is to establish a functional relation between train delays and various characteristics of a railway system (i.e., infrastructure, timetables and trains). Such a functional relation would allow planners to evaluate how changes in the system would affect delays, and thereby help them determine the changes that would reduce delays in the most economical way.

Let us consider two specific situations in which a functional relation between train delays and characteristics of the railway system would be useful for tactical planning:

- *Investment planning*. Managers are interested in reducing train *arrival* delays at a station by investing in the infrastructure. The investments may include track renewal to increase the speed profile along some routes, building an overpass to avoid train conflicts, or introducing more advanced signaling systems that resolve rights-of-way more efficiently. A functional relation between train delays and characteristics of the infrastructure would enable the planners to evaluate the effects that different infrastructure projects (or their combination) would have on delays. This would assist the management in selecting the most effective investment plan, given the available budget.

* Corresponding author.

http://dx.doi.org/10.1016/j.trc.2015.04.004
0968-090X/

- *Simulation analysis*. Planners are interested in examining effects of changes in the railway system on bottlenecks, capacity utilization, or delay propagation. Such analysis is typically conducted by simulating specific traffic scenarios, given the information about the infrastructure and train movements (Hansen and Pachl, 2008). Since trains often deviate from their timetables, simulation models must account for delays to validly represent the real-world operations. A functional relation between train delays and characteristics of the system can be used to estimate initial arrival delays to the observed section or intersection. These initial delays can be used further as inputs to a simulation model to obtain more realistic estimates of the capacity utilization, delay propagation, or bottleneck effects at downstream locations.

The described situations motivate research on determining the functional relation between passenger train delays and various characteristics of the railway system. We shall refer to this functional relation as a train delay model that, given a set of inputs (i.e., information about the system), returns train delays. As implied in the motivating examples, we focus on the analysis of historical data and develop train delay models that can be used for tactical planning. This should not be confused with the analysis of train delays at the operational level and development of models for online predictions which are used for real-time traffic management. We review below these two separate lines of research:

- *Operational level*. Models for real-time prediction of train delays are used to detect potential instabilities in the timetable and alert the dispatchers to reschedule some trains. Train delays are predicted from data describing arrivals and departures, which are gathered in real-time Kecman (2014). Train movement data are collected from the infrastructure track occupation records, sensors in rolling-stock, or mobile GPS devices. Flier et al. (2009a) apply an algorithm that sweeps through the delay of the incoming trains and efficiently finds systematic dependencies in large-scale railway delay data. A delay propagation algorithm proposed by Hansen et al. (2010) uses an online model based on timed event graphs to predict running and arrival times. This approach is extended at the microscopic level by Kecman and Goverde (2014), who propose a model based on a timed event graph with dynamic arc weights that accurately captures train behavior and interactions on open track sections and in stations. Berger et al. (2011) propose a stochastic model for delay propagation in large transportation networks, which is suited to process massive streams of real-time data. Delay propagation across the network is modeled with max-plus algebra and analyzed based on timetable properties, such as realizability and stability (Goverde, 2010). A related line of research in the area of traffic predictions develops multivariate state space models that use state variables to describe a system by a set of first-order differential equations. In traffic research, state space models are often applied to multiple input–output systems of stochastic variables (Stathopoulos and Karlaftis, 2003; Dong et al., 2014). State space models for train delays could also process train describer messages and historical data to compute train delays using a Kalman filter. An advantage of state space models would be their ability to handle large networks effectively.
  As mentioned above, information from online prediction models can be used for real-time traffic management. For example, when delays or disturbances are predicted (or observed), it is vital to quickly retrieve a feasible train schedule. Krasemann (2012) achieves this with a greedy algorithm which efficiently finds good feasible schedules. To minimize train delays and missed connections due to disturbances, Corman et al. (2012) apply two heuristic algorithms to select the connections to be enforced. To reschedule duties of train drivers, Veelenturf et al. (2012) use column generation to solve the crew rescheduling problem with retiming. Corman and Quaglietta (2015) propose a rescheduling tool coupled with a railway traffic simulation model and analyze influence of different rescheduling parameters, prediction models and control delays on system performance. A comprehensive review of the recent literature on rescheduling tools and their applications in train operations is presented by Corman and Meng (2014). The relevant models are classified by their problem scope, model type and solution approach, with a special focus on current train positions and traffic state prediction.
- *Tactical level*. At the tactical level, train delay models have been applied to both timetabling and resource planning. Timetables are tested for robustness using the probability distributions of process durations which are derived from historical traffic realization data. Conclusions from these tests are later used to improve timetable robustness (Medeossi et al., 2011). In a study on passenger trains in the Netherlands, a stochastic approach to cyclic timetable optimization is used to minimize the average delay of trains (Kroon et al., 2008). Büker and Seybold (2012) propose a model with an activity graph where delay propagation is captured analytically, thus enabling precise and efficient calculations even for large networks. Goverde et al. (2013) compare and quantify the infrastructure occupancy both under scheduled and disturbed conditions. Their approach takes into account stochastic delays, dynamic responses to the signaling system, and typical traffic control actions for dealing with conflicts and possibly reducing delay propagation. A comprehensive review of railway disruption management for resource planning (timetable adjustment, rolling stock and crew rescheduling) is presented by Jespersen-Groth et al. (2009).

Various techniques have been used to analyze train delays at both operational and tactical level. They can be roughly classified into three groups:

- *Analytic models*. Probability distributions of train delays are derived from track occupancy and release records, or other data containing information on train movements. The obtained distributions of the underlying random variables are then used within stochastic queuing models (Huisman et al., 2002; Schwanhäußer, 1974). An analytic stochastic model is used