Journal of Rail Transport Planning & Management 2 (2012) 34–50

Contents lists available at SciVerse ScienceDirect

Journal of Rail Transport Planning & Management

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

Stochastic modelling of delay propagation in large networks

Thorsten Büker [a,*], Bernhard Seybold [b]

[a] VIA Consulting & Development GmbH, Frankenberger Str. 30, 52066 Aachen, Germany
[b] trafIT solutions gmbh, Heinrichstr. 48, 8005 Zürich, Switzerland

ARTICLE INFO

Article history:
Received 31 July 2012
Revised 14 October 2012
Accepted 16 October 2012
Available online 22 November 2012

Keywords:
Delay propagation
Delay distributions
Timetable robustness
Estimated time of arrival
Activity graph

ABSTRACT

Using analytical procedures to compute the propagation of delays on major railway networks yields sizeable computing time advantages over Monte Carlo simulations. The key objectives of this paper are to present a formalisation of delay propagation by means of an activity graph, to outline the required mathematical operations to traverse the graph and to elaborate a suitable class of distribution functions to describe the delays as random variables. These cumulative distribution functions allow to be speedily computed but also allows the quality of the computing process to be controlled. Last but not least, issues of procedural theory that arise in the context of networks are elaborated and the translation of the approach to a software tool is presented.

1. Introduction

In order to assess the robustness of railway timetables, it is necessary to forecast the propagation of potential delays in the railway network. Various types of delay (primary delay, secondary delay, current delay) belong to the random variables and can be modelled by means of cumulative distribution functions (cdfs). It is possible with the aid of analytical procedures to reliably compute delay propagation within a very short time, including for major (country-wide) networks. Unlike Monte Carlo simulation, in which concrete implementations of the delay situation are generated for each simulation run, analytical procedures directly process the cdfs of the underlying random variables. Events that in reality lead to changes in the delay (increase in delay due to a sequence-of-trains conflict, decrease in delay due to recovery margins in place etc.) are mapped in the model by manipulating the distribution functions. Use is specifically made of both conditional and unconditional convolution and of "excess beyond" operations.

To ensure the precision and efficiency of the procedure, it is necessary to select a class of distribution functions that guarantees adequately accurate adaptation to the real world whilst not giving rise to running-time and storage-space problems when use is made of the manipulation functions. Since operations on random variables are subject to rather strict mathematical constraints which cannot always be met while computing delay propagation, it is moreover mandatory to analyse, to which extent an injury of these constraints impacts the validity of the results.

This paper is an update of Büker and Wendler (2009) which is extended by a summary of the research results achieved later and being published in Büker (2010). Beside the theoretic background, also the considerations on the approach's transfer to practical application, as described in Franke et al. (2012), are given below.

The delay distributions adopted hitherto for robustness analyses do not fully meet these boundary conditions. The state of the science is set out in Section 2 below. After introducing the transformation of all required input data to an activity graph in Section 3 and having described the requirements for cdfs in Section 4, a distribution function is presented in Section 5 that permits very good adaptation to real conditions. The case is set out there for a piecewise continuous distribution function, with each continuous segment being represented by an act of distribution involving several parallel phases. To guarantee the efficiency of the computational algorithm, a complexity-reduction algorithm is as well outlined. After outlining mathematical constraints that have to be considered if stochastic operations are applied in a network context in Section 6, Section 7 sketches the implementation of the overall concept into the productive software tool and outlines aspects of its practical application. Last but not least the paper is closed by an outlook on further need for research.

2. State of the science

Before the operations to be conducted are set out in detail in the next two sections, existing methods of mapping random variables

* Corresponding author.
E-mail addresses: th.bueker@via-con.de (T. Büker), bernhard.seybold@trafit.ch (B. Seybold).

2210-9706/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jrtpm.2012.10.001

in the context of delay propagation are described and their suitability is appraised. As well, a brief summary of existing approaches to assess delay propagation is provided. A broader overview of existing approaches to assess timetable robustness and to estimate the time-of-arrival is given in Büker (2010).

2.1. Modified exponential distributions

Schwanhäusser arrived at the following distribution function by evaluating delays at entry onto sections of line Schwanhäußer (1974). Parameter a denotes the share of delayed trains, whilst λ is the inverse value of mean delay $\bar{t}$ for the trains delayed:

$$F_V(t) = \begin{cases} 0 & \text{if} \quad t < 0, \\ 1 - ae^{-\lambda t} & \text{if} \quad t \geqslant 0. \end{cases} \tag{1}$$

Fig. 1 illustrates the resultant distribution function. The distribution pattern reveals a jump discontinuity at $t = 0$. Proof of this distribution function being a suitable means of describing delays in train running has been forthcoming from several sources. Besides empirical approaches involving checking the fit on random samples (e.g. in Hermann (1996), Wendler and Naehrig (2004) and Yuan (2006)), analytical evidence has also been furnished by Engelhardt-Funke and Kolonko (2001).

Weidner extended the cdf by a third parameter to enable differentiation between systematic delays t_{sys} and delays of random nature (Weidner and Büker, 2007). The corresponding function is as follows:

$$F_V(t) = \begin{cases} 0 & \text{if} \quad t < t_{sys}, \\ 1 - ae^{-\lambda(t - t_{sys})} & \text{if} \quad t \geqslant t_{sys}. \end{cases} \tag{2}$$

While the low number of parameters permits a very efficient data handling on the one hand, each operation on this class results in a cdf that cannot be represented by the same class on the other hand. In the further course, such a property is denoted as missing *closedness* under the required operations. Approaches to bypass the missing closedness by approximations are introduced by Mühlhans (1990) and Weidner and Büker (2007). They are mainly based on enforcing an exact expectation value after each operation, but a proof on the quality of the approximation is not given.

2.2. Phase-type distributions

A phase-type distribution results from a system of one or more inter-related Poisson processes occurring in phases. The sequence in which each of the phases occur may itself be a stochastic process. The distribution can be represented by a random variable describing the time until absorption of a Markov process with one absorbing state. Each of the states of the Markov process

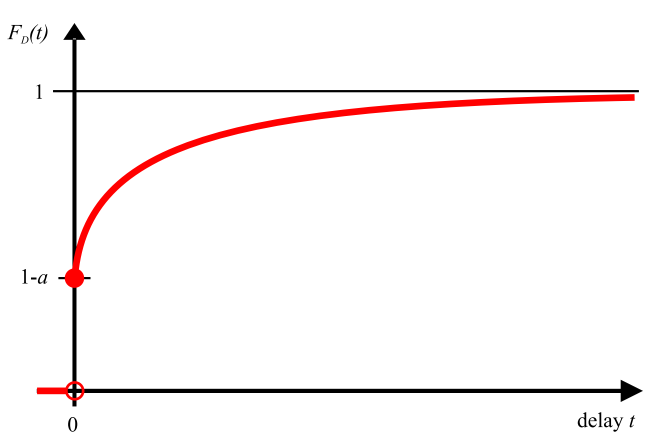

Fig. 1. Delay distribution after Schwanhäußer.

represents one of the phases. This class of cdf is closed under certain operations, in particular the unconditional convolution of cdf (the summation of random variables) and the summation of cdf. This property permits usage of phase-type distributions in models, where compositions and mixtures of phase-type distributed random variables occur (Neuts, 1981). Mainly depending on the number of used phases, the class of phase-type distributions allows a very close fitting as well to cdfs as to the samples of data. From the viewpoint of railway operation one has to keep in mind anyway, that phase-type distributions are solely defined for $t \geqslant 0$.

In Meester and Muns (2007) suggest using this distribution class to model delay propagation on networks and cite two properties to underpin their proposal: firstly, phase-type distributions permit a very good fit with given distribution functions and, secondly, this class facilitates closed execution of the requisite mathematical operations. Maximum formation and "excess beyond" operations are cited as needing to be factored in along with unconditional convolution. The operations shown in Section 4.3 refute the latter point, however: in the absence of free segment boundaries it is not possible on the basis of phase-type distributions to conclusively map secondary delays. As a consequence it is not possible to factor in elementary changes of sequence between trains.

2.3. Theta-exponential polynomials

Kolonko proposes adopting extended exponential polynomials in delay modelling (Kolonko, 2007). This kind of distribution function is described by means of an extremely flexible distribution function with singularities at points ϑ_i:

$$F_V(t) = \sum_{i=1}^{n} a_i \cdot (t - \vartheta_i)^{d_i} \cdot e^{\lambda_i \cdot (t - \vartheta_i)} 1_{[\vartheta_i, \infty)}(t) \tag{3}$$

The distribution function embraces an arbitrary number n of theta-exponential polynoms $a \cdot t^d \cdot e^{\lambda \cdot t}$, each one i defined on $\vartheta_i \leqslant t$. Unlike the distribution functions alluded to thus far, this function rule facilitates conclusive mapping of delay propagation on networks, as all requisite operations can be conducted in a closed manner. It needs to be borne in mind, however, that the number of parameters describing the random variables increases greatly with each operation, leading to numerical problems in data maintenance due to an incidence of very small values. Its use in electronic data processing is accordingly predicated upon the development of complexity-reduction algorithms.

2.4. Discretisation of random variables

Yuan delineates a method for computing delay propagation at a major station (Yuan, 2006; Yuan and Hansen, 2007). His approach discards analytical computation in favour of discretisation of random variables alongside the use of numerical convolution. In the form of representation selected, the quality of adaptation to existing distribution functions is determined solely by means of the set of points required to this end and can be arbitrarily well achieved. The numeric representation is furthermore closed under all necessary operations.

In contrast to a symbolic representation as presented by (1)–(3), the computation effort per operation is almost constant in the numeric representation. Vice versa, practical tests underpin, that the required computation time is considerably higher than in case of a symbolic representation of the cdf.

2.5. Overview of systems to assess timetable robustness

The qualified assessment of the operating quality of timetables is currently performed – if at all – with the aid of microscopic