Transportation Research Part C 68 (2016) 462–489

Contents lists available at ScienceDirect

Transportation Research Part C

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

Distributed model predictive control for railway traffic management

Bart Kersbergen *, Ton van den Boom, Bart De Schutter

Delft Center for Systems and Control, Delft University of Technology, The Netherlands

ARTICLE INFO

Article history:
Received 10 December 2015
Received in revised form 7 May 2016
Accepted 7 May 2016

Keywords:
Railway traffic management
Dispatching support
Mixed integer optimization
Distributed model predictive control
Disturbance management

ABSTRACT

Every day small delays occur in almost all railway networks. Such small delays are often called “disturbances” in literature. In order to deal with disturbances dispatchers reschedule and reroute trains, or break connections. We call this the railway management problem. In this paper we describe how the railway management problem can be solved using centralized model predictive control (MPC) and we propose several distributed model predictive control (DMPC) methods to solve the railway management problem for entire (national) railway networks. Furthermore, we propose an optimization method to determine a good partitioning of the network in an arbitrary number of sub-networks that is used for the DMPC methods. The DMPC methods are extensively tested in a case study using a model of the Dutch railway network and the trains of the Nederlandse Spoorwegen. From the case study it is clear that the DMPC methods can solve the railway traffic management problem, with the same reduction in delays, much faster than the centralized MPC method.

1. Introduction

In many countries in the world large, complex, and very busy railway networks have been built. Especially in North and West Europe, China, and Japan the railway networks are used near their maximum capacity. As a result, very little buffer time is available to recover from delays.

Every day small delays occur in almost all railway networks. Such small delays are often called “disturbances” in literature. In order to deal with disturbances dispatchers reschedule and reroute trains, or break connections. Currently most dispatchers take these decisions based on their experience, a given set of ground rules, and a limited overview of the network situation. To be able to handle disruptions trains may need to be canceled, or they may need to be rerouted through the entire network. These changes affect the rolling stock circulation and the personnel schedules. As a result, the rolling stock circulation needs to be recomputed and adjustments to the personnel schedules need to be made. Large perturbations, such as trains breaking down and tracks being blocked, causing trains to be canceled, are often called “disruptions” in literature. In this paper we focus on railway traffic management for disturbances, and therefore we will not consider the rolling stock circulation or the personnel schedules. For an overview of those research areas, literature on disruption management, and integrated approaches combining several of these research areas the reader is referred to the survey paper of Cacchiani et al. (2014). The literature on methods to railway traffic management for disturbances can, for the most part, be split up into two groups based on the size of the problem instances that are considered: there are approaches that focus on a small part

---

* Corresponding author.
E-mail addresses: b.kersbergen@tudelft.nl (B. Kersbergen), a.j.j.vandenboom@tudelft.nl (T. van den Boom), b.deschutter@tudelft.nl (B. De Schutter).

http://dx.doi.org/10.1016/j.trc.2016.05.006
0968-090X/© 2016 Elsevier Ltd. All rights reserved.

of the railway network and there are approaches that take the entire (national) network into account. We will call railway management for small parts of the railway network "local control" and railway management for the entire network "global control". Our work in the paper will focus on global control.

Most of the research on railway traffic management is on local control. There we will first look at the literature on this subject. In recent years many approaches for local railway traffic management have been proposed such as the methods of Caimi et al. (2012),Corman et al. (2010a, 2012b), D'Ariano and Pranzo (2009), D'Ariano et al. (2007), Meng and Zhou (2014), Murali et al. (2015), Pellegrini et al. (2014, 2013) and Rodriguez (2007).

Caimi et al. (2012) developed a railway traffic management method that tries to schedule and route all trains in an area in and around a large station. A model predictive control approach with a microscopic model of the railway operations is used based on blocking times. At each point where rescheduling of a train is possible, a set of possible blocking times for different routes and departure times at platforms or arrival times at the boundaries of the area are considered for that train. As many trains as possible are then assigned a route with corresponding blocking times such that all safety and operational constraints are still respected. The objective is to optimize the passenger satisfaction, measured by punctuality and reliability. The resulting optimization problem is a binary linear programming problem.

In the work of Pellegrini et al. (2014) the authors propose a mixed integer linear programming (MILP) approach to reschedule and reroute trains. They model the infrastructure with as many details possible. They even model the track circuits a block section is build up from separately so they can model sectional release in the railway traffic management and compare it to block or route release. Optimizations can be performed subsequently over a long time horizon, and the decisions resulting from each optimization are implemented as they are produced, resulting in a so-called receding or rolling horizon. They propose two objective functions for the MILP approach, namely minimizing the maximum secondary delay and minimizing the total (secondary) delay. They test their approach for two case studies: the triangle of Gagny and the Lille-Flandres station for different prediction horizon lengths. The first case study is used to determine the difference in delay reduction for the sectional and route release approaches. The second case study is used to test the real-life applicability of the approach. They try to determine the optimal solution for the various delay scenarios and horizon lengths. In many of the cases it proves to be quite difficult to determine the optimal solution, but in most cases good solutions can be found within 3 min. Our work differs from this work in several ways. The goal here is to reduce the delays in small areas such as stations, whereas we aim to reduce the delays on a network wide scale. As a result we do not consider rerouting, we consider station areas a black boxes where we know the arrival and departure times, but not the routes, or the layout of the station.

Meng and Zhou (2014) developed a method for the simultaneous rerouting and rescheduling of railway traffic for an N-track network. The network is described per block section and in each block section only one train can be present at the same time. They propose a novel Lagrangian relaxation solution framework to decompose the problem into easier to solve dual subproblems and then further decompose the problem into a set of single train routing and scheduling problems that can be solved in sequence. The solutions of the dual problems can be transformed into feasible solutions for the original problem.

Murali et al. (2015) present a macroscopic modeling strategy for medium sized railway networks for freight trains. They also propose a method to reduce the complexity of the model by aggregating portions of the network into single nodes. They formulate the routing and scheduling problem as an integer programming (IP) problem and propose two methods to find (partial feasible) solutions to the IP-problem called PFSLPR and PFSRC. For PFSLPR the IP-problem is solved using LP-relaxation. The solution of the LP-relaxation is rounded of to find an integer solution giving the routes and the departure times at the origin stations. With the PFSRC approach the possible routes are limited by adding constraints that enforce a periodic route schedule. The resulting problem is a mixed integer linear programming-problem. Solving it gives the routes for the trains and with the routes the maximal matching problem can be solved to find the departure times. The authors use a genetic algorithm to improve the solutions found with the two approaches. The computation performance and quality of the solutions of the approaches is compared for a set of networks with a different number of trains. Finally the original model is compared to the aggregated model for different levels of aggregation.

Rodriguez (2007) proposes a method for the routing and scheduling of trains through an area around the Pierrefitte-Gonesse junction north of Paris. Rodriguez uses a microscopic simulation to model the train and driver behavior and describes the routing and scheduling problem as a constraint programming problem. For the given case study the reduction of the delays ranges from 63% to 96%.

In the work of D'Ariano and Pranzo (2009) and D'Ariano et al. (2007) the railway operation is also modeled as a microscopic model based on blocking times with an Alternative Graph (AG) approach. In their AG approach for every train occupying a block section a node is created in a graph. The nodes of a single train are then connected to each other through running time constraints and for every pair of trains occupying the same block section headway/separation constraints are added. If the order in which the trains can occupy the block sections can be changed with rescheduling actions, then a pair of alternative arcs defining the two orders in which the trains can occupy the block section are added to the graph. A new schedule for the railway traffic is found when for each pair of alternative arcs only one arc is chosen and no circuits of positive length are present in the graph. The graph has an extra node to which all nodes are connected and the weights of the arcs from all nodes to this extra node are chosen such that minimizing the maximum weight of all paths from the starting to the ending node corresponds to minimizing the maximum consecutive delay. To solve this problem the authors use their own branch and bound algorithm.

Corman et al. (2010a, 2012b) have extended the work of D'Ariano et al. (2007) to also consider breaking connections and locally rerouting trains. They consider the effects of breaking connections by determining the maximum consecutive delay