RTSE, a multi-component closed-loop control framework for railway networks

Raimond Wuest¹, Jonas Looser¹, Bernhard Seybold¹, Marco Laumanns², Juliane Dunkel², Daniel Huerlimann³, Samuel Roos⁴,
¹ Zurich University of Applied Sciences, Winterthur
² IBM Research - Zurich
³ OpenTrack Railway Technologies Ltd.
⁴ Emch + Berger AG

Abstract

Optimal operation of rail transport systems has become an increasingly challenging task over the last decades. To allow for a better understanding of the system dynamics in different operational states (including disruptions) and in order to evaluate and to improve control strategies, a multi-component simulation framework, representing a closed-loop operation environment for railway networks, is being developed. This framework is based on a time controlled and partially automated operational concept. Time control requires all operational processes to be continuously monitored with respect to the production schedule. Deviations exceeding some pre-determined tolerance thresholds will result in a re-adjustment of the production plan in real-time. A dedicated rescheduling algorithm is implemented to achieve this goal.

Involved parties (agents) are explicitly taken into account. For instance, train drivers might be technically enabled to follow new operational targets like re-adjusted train speeds while approaching conflict points. The framework, called Rail Transport Service Environment (RTSE), consists of three main modules: (i) a traffic simulation environment, (ii) a system state monitoring module, and (iii) the scheduling module. The modules are interconnected through standard communication interfaces so that each module can be exchanged easily depending on the user environment. Railway traffic simulations are carried out using the dedicated railway simulation tool OpenTrack. The simulated traffic situations are interpreted by an automated monitoring module including a threshold detection mechanism, which compares actual and planned process states and induces rescheduling actions executed by the
rescheduling algorithm, if required. Rescheduling actions take eventually reduced availability of resources into account.

**Keywords:** Rail traffic – Service intention – Rail traffic state monitoring – Real-time dispatching – Network performance – Rail traffic simulation

1 Introduction

Modern railway operational processes abandoning accumulated legacy cases, explicit description of the service to be delivered to the customer, and increased computational and communication performance allow to make rail transportation service more predictable as well as to orient operations to customer benefit even in case of incidents. Operation according to plan even in case of delay or disruption allows for a better use of contested capacity. With our proposed framework we aim to demonstrate and evaluate this opportunity.

1.1 Motivation

The identification of increasing capacity problems that ask for a redesign of railway operations is the main motivation of the project participants for developing the proposed framework. In order to derive the elements of the proposed approach, we will first have a closer look at the major problems.

Challenges in service delivery

The Swiss railway network and the services delivered through it can be characterized as a multiple hub-and-spoke network with integrated clock-face timetable and it is well known to be strongly interconnected. This means that there are lots of point to point services that require one or more train transitions (see for instance the Zurich area of the Swiss timetable 2013 in Figure 1). In these cases individual delays often have impacts on a large part of the network. This is the case if train dispatchers decide to hold back the connecting trains affected or if they decide not to keep communicated connections. In the first case, the initial delay of the feeding train not only causes a delay for the passengers of this train but also for the passengers that are supposed to connect to other trains. In the latter case in which connections are broken, all connecting passengers will suffer a much higher delay as they have to wait for the next connecting train to their intended destination.


1 Introduction

Figure 1: Clock-face timetable for the area of Zurich, Source: SMA and Partner AG Zurich, Timetable 2013.

On the other hand, the operating staff as well as the customers are facing typical problems that result from the lack of timely information on the system as illustrated in Figure 2.

Figure 2: Operating staff and customers confronted with decreasing service reliability, each with different responsibilities within the service process chain
Following the above explanations, the major service delivery challenges can be summarized as follows:

- A tight regular timetable due to an increasing gap between peak hour and off peak hour demand.
- Decreasing service reliability due to operational volatility and technical disturbances.
- Limited usability of public transport due to communication problems.
- Significant total passenger delays due to local dispatching decisions.

If we analyse specific cases and try to figure out what happened before major network delays, we observe similar patterns. We see that in most cases operational disturbances lead to blocked resource assignments and hence unusable production plans. As the production plan is the technical basis for operating the timetable, it is the main task of operational control to reassign resources such that the normal timetable and the planned services will be restored as soon as possible. This is a complex task that requires, even in the case of small delays, the consideration of numerous operational and technical constraints.

**Problems in service development**

Usually timetable development is an extensive, iterative planning process starting years ahead of the beginning of the actual timetable period. The main objective is to find train runs with departure, arrival and dwell times that simultaneously meet functional requirements (Service Intention, SI) connecting each origin and destination in the network with required travel times, frequencies and service levels as well as technical requirements (the utilization of available resources, such as track topology, rolling stock and operation staff). This is a strongly iterative process, which usually requires a lot of feasibility checking. It is considered to be unrealistic to try implementing such a process in operational environments. The main difficulty of this service development process is its sheer complexity (see Figure 3):

- The development process is sub-divided into three phases: 1. Service offer development, 2. Timetable development, and 3. Operations and Dispatching.
- Organisation and tools involved in each of the development phase are specific for each phase.
- Each phase transition is related to a certain shift of objectives (phase 1: demand, phase 2: capacity, phase 3: stability)
- When taking dispatching decisions during operation, it is very difficult to deploy
the original service intention in real-time.

<table>
<thead>
<tr>
<th>Planning level</th>
<th>Long term planning (Service offer and resources)</th>
<th>Mid-term planning (Timetable)</th>
<th>Short term planning (Operations and dispatching)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time horizon</td>
<td>20 to 5 years</td>
<td>5 years to 24 hours</td>
<td>24 to 0 hours</td>
</tr>
<tr>
<td>Interrelations</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Level of detail                 | low level of detail                              | high level of detail          |                                               |

**Figure 3:** Service offer value chain: Development process in three phases, distributed over many years. From this it becomes obvious, that it is very difficult to refer in all phases to the same objective (service intention)

## 2 Solution approach

### 2.1 Service Intention

The central part of the proposed approach is the assumption that any timetable is one out of several possible technical realizations of an underlying service intention (SI). Because the SI can be regarded as a functional requirement for the scheduling task, it is defined in terms of frequencies between origins and destinations, travel times and service levels rather than of exact departure and arrival times, rolling stock utilization or even trains numbers. Although these attributes are information typically provided by a public transport timetable, it is nothing else than the result of an assignment of processes and resources (consisting of track topology, rolling stock and operational staff) to these functional requirements, while considering numerous spatio-temporal constraints. As a consequence, operational irregularities or disruptions, which, for a certain period of time prohibit further operation of the timetable planned, require a new assignment of available resources to the functional requirements. To formalize this approach, we define the periodic service intention
RTSE, a multi-component closed-loop control framework for railway networks

similar to [Cai09] as

\[ SI = (Z, C, D, \bar{\rho}), \]

where \( Z \) denotes the set of all train runs \( z \) (with \( z \in Z \)), \( C \) is the set of all connections, \( D \) denotes the set of all technical and operational dependencies and \( \bar{\rho} \) is a given time period (for details see [Cai09]).

2.2 Process Integration

From the arguments provided in section 1.1 it becomes clear that a solution needs to focus on the service intention. To achieve this, the proposed integrated framework requires the following:

- a global data model for all objects (e.g. train runs, resources)
- global rules and parameters (e.g. resource conflict conditions)
- scalability of both traffic network and system architecture (to insure feasibility of the production plan with a global network scope)
- assessment of the process state (in case of operations),
- handling of technical conditions, and
- recalculation of an SI-based production plan in real time.

If we sketch this proposed solution in the process development scheme of Figure 3, we obtain the service offer value chain as illustrated in Figure 4.

![Figure 4: The development of a production plan with a particular focus on service intention enables the integration of processes, objectives and scopes.](image-url)
2.3 Process model

Operation processes, including resource state transitions are modeled as Timed Event Graphs. For instance, safety blocks are treated as resources which are assigned to train runs by the reschedule. This process model is shown in Figure 5 as a timed petri net representation of the departure process of a train run.

![Diagram of Timed Petri Net](image)

**Figure 5**: Integrated process model: example of departure process and its representation as timed event graph.

Given a certain SI, the integration of planning processes which is required for the closed loop development of the production plan of Figure 4 just means that the task to be executed is always the same: find a solution for the assignment of processes that satisfies the SI with the available resources.

2.4 Production plan

The result of this integrated planning process is the production plan, which contains all necessary information for operating the train service. Its formal representation is a list of assignments of event times $d_i(k)$ to triples $(E_i, L_i, S_i)$, where $E_i$ represents the event type (arrival, departure, passage), $L_i$ the train line, $S_i$ a location (station, signal, junction) and $k$ the $k^{th}$ occurrence of a periodic event $i$ within a given time period $T$.

In this way the production plan can be used directly for monitoring purposes by comparing planned event times with recorded event times and estimating impacts, e.g. according to [Gov09], as well as for a railway timetable stability analysis, e.g. using max-plus algebra according to [Gov07].
RTSE, a multi-component closed-loop control framework for railway networks

In case of larger disruptions, where no production plan which satisfies the current SI can be found, the dispatcher has to relax the SI in an iterative process in the management layer. In this iteration process the (re-)scheduler creates production plans for several variants of SI that are provided by the dispatcher (see process loop in the Management Layer of Figure 6).

### 3 System Design

To address the problems outlined in chapter 1 and to incorporate the fundamental conceptual elements described in detail in chapter 2, we developed the general framework shown in Figure 6. The system consists of three layers, each with dedicated functionalities:

**Figure 6:** The functional components of the system, the feedback loop and the corresponding interfaces (interface numbers are shown in blue circles): (I0) configuration data, (I1) service intention, (I2) production plan and unfeasible train runs (if any), (I3) timetable and actor instructions for real-time process configuration, (I4) position and process state messages, (I5) new scheduling constraints related to the available resources.

The Management Layer is at the top. One of its components is the dispatcher who represents the highest decision level in the Management Layer. The dispatcher’s task is to manage the functional requirements, i.e. the SI. The SI is entered into the
Management Layer component (re-)scheduler (interface number 1) as a basis for the calculation of the normal production plan. In case of an operational rescheduling requirement that cannot satisfy original SI-requirements, the dispatcher has to relax the SI such that the (re-)scheduler can find a feasible solution for a new temporary production plan (interface 4). The production plan (see 2.4) is displayed either as a list (see Figure 7 a) or as a timed petri net (see Figure 7 b).

![Figure 7: production plan representation a) list, b) timed petri net.](image)

Components of the **Logical Layer** are responsible for the message generation corresponding to the production plan (interface 3, see Figure 8).
RTSE, a multi-component closed-loop control framework for railway networks

Figure 8: Actor instructions to train drivers, operators (resp. traffic management system) and train guards affected by the (re-)scheduling action (interface I3)

After actor specific formatting, the production plan is used by the Logical Layer for (re-)configuration of the railway network. During the run time of the railway network there are position and process state messages (interface 2) evaluated permanently by the operation control part in order to detect threshold exceedings and new constraints and to initiate rescheduling (interface 3) in the Management Layer. The message interfaces of the Logical Layer (interfaces number 2, 3, 4 and 5) are easy to standardize. This ensures that conclusions that can be drawn from the simulation environment, are transferable to real world conditions. On the other hand, different rescheduling modules (Management Layer) can be used without changing anything in the Logical Layer.

The Physical Layer consists of the operational environment, represented either by the simulation environment (in our case the railway simulation tool OpenTrack [OpT12]) or by the real-world system. It carries out all transport operation, safety, customer and disruption processes and contains detailed information about resources (interface 0).

4 System Behaviour

The system behaviour is strongly influenced by the precision, with which the dynamics of train runs can be approximated by the (re-)scheduling algorithm of the Management Layer. But as the behaviour of the simulated train runs is only
controlled via data configuration and message instructions (train departure, train speed etc.) it is essential to generate the right speed instructions at the right time. An example of an intercity train run in a given test scenario (see Figure 9) is illustrated in Figure 10. It shows that our framework is able to generate a reasonable dynamic precision in normal (undisturbed) situations.

Figure 9: Topology of test scenario with 12 stations. Path of intercity train indicated in red

Figure 10: Deviation of planned vs. recorded event times of intercity train run
5 Conclusion and Outlook

The behaviour of the RTSE framework, which we present in this paper is of sufficient accuracy to simulate train operations in a closed loop. Our next steps will be to implement the entire system setup in an online architecture and to use this to investigate the dynamic closed loop behaviour of a railway network in real time. Thus we will establish a system environment that allows us to benchmark dispatching decisions in terms of good choices for relaxed SI’s under conditions of disturbed train operations.

References


Corresponding author: Raimond Wuest, Zurich University of Applied Sciences, Institute for Data Analysis and Process Design, CH-8401 Winterthur, Switzerland, phone: +41 58 934 6581, e-mail: Raimond.Wuest@zhaw.ch