Calculating the maximal number of additional freight trains in a railway network*

Christian Meirich, Nils Nießen

RWTH Aachen University, Institute of Transport Science
Chair of Railway Engineering and Transport Economics
Mies-van-der-Rohe-Straße 1
D - 52074 Aachen

Keywords:
railway capacity; assessment of railway infrastructure; network optimisation; routing; linear programming

Abstract

This paper describes an approach to maximise the number of train runs through global networks. Nowadays the capacity of railway lines as well as the capacity of railway nodes can be calculated by means of analytic algorithms, especially by applying queueing theory. Currently there is no generally accepted method to allocate the overall capacity of lines and nodes of such networks by using analytic algorithms. The interactions of these elements (e.g. a node might be the limiting factor for an adjacent line) are part of the approach presented in this paper. To maximise the number of train runs in a global network the passenger trains will be fixed on their train curses so that the remaining capacity could be used for freight train service. Existing or detected bottlenecks could be eliminated by means of sensible rerouting due to the optimisation, which will reveal the best train paths through the network. The paper concludes with an illustrative computation for a generic railway sub-network.

* This is the Author’s Accepted Manuscript of the following article: C. Meirich; N. Nießen, Calculating the maximal number of additional freight trains in a railway network, Journal of Rail Transport Planning & Management 6(3), pp. 200-217, 2016, which has been published in final form at https://doi.org/10.1016/j.jrtpm.2016.06.005
©2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/
Email address: niessen@via.rwth-aachen.de
1 Introduction

One aim of strategic network planning is to optimise the capacity available throughout the network. It is only currently possible to establish capacity separately for railway lines and railway nodes by means of analytical methods. Use is primarily made of the principles of queueing theory to this end. It has not been possible hitherto to apportion capacity for lines and nodes as would be required to itemise network capacity. Differing assumptions used for calculations can be regarded as being one of the reasons for this. The importance of facilitating such apportionment so as to justify suitable means of developing and maintaining the network in longer-term planning will nevertheless grow in the years ahead. The procedure presented here aims to use the capacity available on lines and nodes in a macroscopic model and to establish a global capacity maximum for the railway network in the form of maximizing possible train moves.

Establishing the maximum number of possible train moves in a network is of particular relevance with rail traffic forecast to rise in Europe. The procedure is additionally designed to serve as a basis for identifying bottlenecks in the network. Reducing the number of train moves on a railway line or node identified as being congested is conducive to removing bottlenecks. It is possible by rerouting trains along alternative running paths to reduce the loading on congested network sections. The scope for a free choice of train-paths is nevertheless constricted in passenger traffic due to the predefined scheduled stops involved. In freight traffic, by contrast, in which sources and sinks generally prevail, resort can be had to alternative running paths displaying a high level of residual capacity. This allows congested infrastructure areas to be circumnavigated.

The present work aims to describe the loading for a complete railway network. An algorithm capable of maximising capacity as a function of the number of train moves is to be developed to this end. It is additionally to be possible to detect bottlenecks and eliminate them by means of sensible rerouting. The procedure presented here will also, to conclude, permit analysis of a variety of infrastructural or operational scenarios.

The paper concludes with an illustrative computation for a generic railway sub-network.
2 Research Progress to Date

Set out in the following section is the current state of scientific progress or point of departure for the network-wide optimisation of capacity. It includes a description of the general composition of railway networks as well as of procedures for establishing the capacity of transport systems.

2.1 Composition of railway networks: lines, nodes and sets of tracks

A railway network can be divided into individual infrastructure elements that are explained in greater detail below and basically comprise nodes and lines.

Nodes are deemed to be either stations linking at least two lines with one another or junctions. A node may be further subdivided into route nodes (RN) and waiting positions\(^1\) (cf. Fig. 1). The UIC leaflet 406 describes the railway nodes as larger interlockings which are situated on the edge of train path line sections or line sections [1].

![Fig. 1: Route node and set of tracks](image)

Sets of tracks can take the form of platform, through and passing tracks as well as of handling tracks capable of serving as (scheduled or unscheduled) stopping or waiting positions.

A line is the connection between two nodes. Where a line under review is to incorporate the running paths between two major nodes, then it is possible to include intermediate stations as part of the line. It is the major nodes, however, that form the

---

\(^1\) May also be referred to as a set of tracks.
start and end of the line. Furthermore, a network is defined by the nodes and the lines, which usually cannot overlap [1].

2.2 Capacity of transport systems

The capacity of transport systems is generally taken to mean the ability to meet demand for the conveyance of persons, goods or information to a desired level of quality.

In [2] the capacity of infrastructure is defined as follows:

“The maximum number of trains that may in theory be operated over a defined part of the infrastructure over a certain period of time, through this limiting value is unlikely to be reached in practice.”

The transport system’s actual loading is compared and contrasted with a basic quality indicator as a rule. The higher the loading, the sooner the queue lengthens (gridlock formation) and hence the sooner waiting times arise. This correlation between traffic units and waiting time is shown in Fig. 2.

![Fig. 2: Correlation between loading and quality](image)

The maximum possible number of trains $n_{\text{max}}$ theoretically gives rise to infinite congestion (infinite waiting time, infinite delay). This number of trains is referred to as the “theoretical capacity”, a value that is of no relevance to practical applications, however. It is generally the case, instead, to compare and contrast the quality available (waiting time, knock-on delay) with an admissible level of service (LoS) when dimensioning railway infrastructure. The number of trains $n_{\text{opt}}$ can thus be used to establish the (economically) optimum loading, that at which congestion is at a defined level and the system as a whole is most profitably marketed for both infrastructure managers and train operating companies. This optimum number of trains is thus referred to as the “practical capacity”.
Of relevance when establishing capacity, therefore, are the infrastructure area (also the infrastructure elements) and its properties, the period under review, the operating schedule and the traffic quality defined.

A train move is defined as traversing a running path which includes minimum two nodes. An infrastructure element is equal the lines, route nodes and set of tracks.

The capacity of railway networks is influenced by the most diverse of factors. Hence attention needs to be paid when studying capacity to a plethora of operational requirements, rescheduling strategies, speeds, length of block sections, automatic train control systems and signalling facilities. Such influences are factored in with the aid of minimum headway times (cf. [2]). The use of minimum headway times when establishing capacity by analytical means is to the applicable technical standard.

**2.3 Procedure for establishing capacity**

There are various procedures for establishing the capacity of railway infrastructure. Common to most is that they map the infrastructure and train characteristics as well as specific national signalling systems and operating processes as input variables.

Literature on the subject summarises the approaches available in differing ways. Pouryousef, Lautala and White provide a fine overview of the various approaches [3]. [4] divides these approaches into analytical, optimisation or simulation methods. In [2], Pachl advocates a breakdown into analytical and simulation procedures, the approach echoed most frequently in the literature. A further option involves a division into timetable-based procedures and such as are not underpinned by any precise timetable [5].

Capacity can additionally be established adopting compilatory procedures. One widespread approach to establishing capacity draws on the blocking-time theory (cf. [6]). One means of establishing capacity adopted by the UIC (Code 406 Capacity [6])
operates by compressing what are known as stepped blocking-time series. To this end, the stepped blocking-time series compiled are pushed as closely together as possible without causing conflicts and it is then possible by comparing the amounts of (compensated and non-compensated) time consumed to determine the concatenated track-occupation ratio. Placing the concatenated track-occupation ratio in relation to quality allows potential capacity to be established (cf. [1]).

A further procedure involves establishing capacity by simulation. The system’s behaviour is replicated and the ensuing parameters measured. Attention primarily focuses on delay development here. Simulation exercises can be run either asynchronously or synchronously. Synchronous simulation only allows the running of traffic to be reviewed, however (cf. [7]). In order to derive maximum benefit from the two simulation procedures, they are now often run in hybrid formation (cf. [8]).

Capacity can similarly be established with analytical procedures. The railway system is viewed with the aid of queueing-theory approaches to this end. Railway lines and nodes are modelled as queueing systems. Train moves can then be grasped as demands for the service system. Service times are described by means of minimum headway times for diverse sequence-of-trains scenarios. Attention is drawn to [9], [10] and [11] for a detailed account of the approach underpinned by the queueing theory. The upshot of establishing capacity analytically is that the number of possible train moves for the infrastructure section under review is output as a ratio of a predefined level of service.

Assessing the overall capacity of a network different approaches exist. The macroscopic NEMO traffic generation and re-apportionment model is used to assess the profitability of a railway network [12]. In [13] the authors describe a solution to maximise the number of train paths in a network by using mixed integer program. However, Hertel describes in [14] the maximisation of transport services by railway lines. Martin explained in [15] and [16] two procedures how to determine the maximum capacity in an investigation area by using the simulation. Moreover, there are possibilities to generate regular interval timetables without conflicts for rail passenger services and also for rail freight services in networks [17].

With interest focusing on a direct correlation between capacity and quality in a railway network, analytical methods are most suitable for the procedure presented here.
3 Mode of Procedure and Method

There follows a detailed delineation of the mathematical model adopted and an outline of the general course of action for the optimisation targeted.

3.1 Basic model

For the purpose of optimising the number of train moves over an infrastructure element and for a sub-network on the basis thereof, use is made of a set of linear equations. These can then be solved with applicable software.

A set of linear equations comprise a linear (objective) function that is usually to be minimised or maximised. In addition, at least one equality constraint is needed per variable in the objective function.

The objective function is to be formulated as the maximisation of all trains/possible train moves \( n_T \) on all predefined (profitable) running paths in the sub-network (cf. Formula 3.1.)

\[
\text{max } c^T \cdot n_T \quad 3.1
\]

with

\[
n_T \geq 0 \quad 3.2
\]

\[
0 < c \leq 1 \quad 3.3
\]

\( n_T \) vector of the number of train moves  
\( c \) vector for the cost for weighting running paths  

Running paths are used as variables in this objective function. There may be a variety of running paths between an initial and a destination node, assuming these are negotiable by the commercial type in question and profitable. The further course of action for selecting alternative running paths is set out in greater detail in Subsection 3.3.

The constant \( c \) can be used for weighting running paths. These can be weighted to the following default requirements as laid down by Radtke in [18]:

- shortest path (by running path)  
- shortest path (by running time)  
- most profitable path (by pathing price)
- energy-efficient/less polluting path

Weighting the running path in line with capacity consumed is additionally conceivable.

Each running path is then weighted and apportioned in direct comparison with, for example, the shortest path on the corresponding point-to-point route. It is possible to weight a number of criteria simultaneously. With maximisation being posited, the reciprocal is used to determine the costs per running path, since otherwise the least profitable point-to-point route would be favoured.

An upper limit $n_{\text{max}}$ and, where applicable, a lower limit $n_{\text{min}}$ for capacity are required as input parameters for each infrastructure element. These capacity limit-values can, for instance, be arrived at by running capacity-quantification software (cf. Subsection 3.2). For simplicity’s sake, computations are confined to the upper limit $n_{\text{max}}$. Because currently capacity figures (maximum number of train moves) cannot be calculated analytically for a subnetwork, the capacity of the combined elements – lines, route nodes and set of tracks – are used. The interactions of these elements (e.g. a node might be the limiting factor for an adjacent line) are part of the approach presented in this paper.

One of the following conditions accordingly applies as an equality constraint for each infrastructure element.

Lines:  
\[ n_{T,L_{i,j}} \leq n_{\text{max},L} \]  
3.4

Set of tracks:  
\[ n_{T,SoT_{i}} \leq n_{\text{max},SoT} \]  
3.5

Route nodes:  
\[ n_{T,RN_{i,r}} \leq n_{\text{max},RN} \]  
3.6

Indices include the number or designation of the applicable starting element $i$, the additional index $j$ denoting, in the case of lines, the node in the direction of which the line runs. Index $r$ for route nodes indicates the location of the route node relative to the set of tracks (left; right; etc.)

Equality constraints corresponding to the variables for the objective function need to be put in place for the solver software so optimisation can be performed. They are required for each node and point-to-point route along a line as well as for each route node reviewed and are quantified as a function of applicable capacity (possible number
of train moves) for the element and of the number of train moves working these elements (operating schedule). A train move is, as set out above, defined with reference to its running path under this procedure. This means that a train move (variable) likewise consumes a unit of capacity per infrastructure element and thus also needs to factored into the equality constraints.

The constraints for the set of tracks in nodes SoT\(_i\) under this approach are generally thus:

\[ \sum_{n_T \in n_T, SoT_i} n_T \leq C_{SoT_i} \]  

3.7

A general constraint can likewise be formulated for lines L\(_{ij}\):

\[ \sum_{n_T \in n_T, L_{ij}} n_T \leq C_{L_{ij}} \]  

3.8

Route nodes RN\(_{i,r}\) are calculated analogously with the aid of the following formula:

\[ \sum_{n_T \in n_T, RN_{i,r}} n_T \leq C_{RN_{i,r}} \]  

3.9

In formulae 3.7, 3.8 and 3.9, the sum of trains \(n_T\) relates to train moves negotiating the node, line or route node in the period under review. Available capacity is designated as \(C_{SoT}\) for set of tracks in a node (cf. formula 3.7) and \(C_L\) for lines (cf. formula 3.8). \(C_R\) in formula 3.9 denotes capacity for a route node.

### 3.2 Process chart

All the concepts presented in this Section can be readily assimilated in the following flowchart of the optimisation process (Fig. 4). First the general process is elucidated. The spheres of running-path search, optimisation, possible alteration of the mixing ratio and evaluation of results are covered in greater detail in the following Sections.

Optimising a suitable railway (sub-) network is predicated upon performing specified steps beforehand. It is first necessary to model a railway network to the desired degree of detail. The degree of detail opted for depends on whether a purely macroscopic network, i.e. nodes and lines\(^2\), is to be analysed or whether a finer level of modelling is to be selected.

---

\(^2\) In macroscopic networks, normally the term “link” is used instead of “line” [2].
Fig. 4: Optimisation flowchart
A further step involves establishing the composition of traffic (mixing ratios) and minimum headway times for each network section on the basis of generic trains with an appointed reference operating schedule working the previously defined sub-network. The composition of traffic depends on the types of trains which are sharing an infrastructure section. Furthermore the reference operating schedule is used as basis for the first calculation run. These indicators generally serve as input parameters for exercises to establish potential capacity on each infrastructure section (line, route node or set of tracks). All these points are necessary and could pool as the initial situation (input) to the described optimisation.

As depicted in Fig. 4, the next step in the procedure involves detecting rail passenger traffic. In the event of there being any passenger traffic, this is to be fixed and the number of passenger trains $n_P$ will be used to reduce the limit-values for capacity later in this progress. A free route search is usually difficult, if not impossible, to conduct, since this type of traffic is subject to mandatory and (chronologically) predefined stops in most cases.

The aim of the present approach to optimisation is to sensibly route freight traffic through the railway network, since in this way relief can be provided for congested railway infrastructure. It is likewise possible in this way to accommodate any additional freight traffic in the network if there is no longer any spare capacity on the conventional/original running paths. Profitable alternative routes between the freight service's departure and destination nodes are required for this purpose (cf. Subsection 3.3).

Given the various approaches to establishing capacity (cf. Subsection 2.3), a non-timetable-dependent approach is required, since a direct comparison is to be made between capacity and quality. Analytical computations are accordingly most suitable for the procedure presented here.

The algorithms to calculate capacities of each infrastructure element in a railway network are based on single- or multichannel service systems.

It is possible for the purpose of establishing the capacity of railway lines to model a line as a single-channel service system. The service times of this system are derived from the minimum headway times of the trains. For each line an average minimum headway time $\bar{z}$ can be established. The track-occupation ratio $\rho$, which equates to the loading, can be calculated by means of the average minimum headway time and the inter-arrival times [19]. Combining these input data with the average delay per train and the

---

3 Every line, set of tracks or route note can be pooled as a network section.
probability of a knock-on delays can be calculated using the Strele formula devised by Schwanhäußer [20] (also in [21]).

Schwanhäußer has published an analytical approach to calculate the unscheduled waiting time for railway lines in the year 1974. Unscheduled waiting times are a measurement to describe the anticipated knock-on delays. This approximation is based on a non-timetable dependent approach. In extension literature the procedure often is named as the Strele formula derived from the German term of capability on railway lines. The Strele formula is defined as:

\[
\sum t_{d2} = \left( p_{d1} - \frac{p_{d1}^2}{2} \right) \cdot \frac{\bar{t}_{d1}^2}{\bar{t}_p \cdot \left( 1 - e^{-\bar{t}_h/\bar{t}_{d1}} \right)} \cdot \left[ p_{er} \cdot \left( 1 - e^{-\bar{t}_{h,er}/\bar{t}_{d1}} \right) + (1 - p_{er}) \cdot \frac{\bar{t}_{h,er}}{\bar{t}_{d1}} \cdot \left( 1 - e^{-2\bar{t}_{h,er}/\bar{t}_{d1}} \right) + \frac{\bar{t}_h}{\bar{t}_p} \cdot \left( 1 - e^{-\bar{t}_h/\bar{t}_{d1}} \right)^2 \right]
\]

3.10

with:

- \( p_{d1} \) probability of an initial delay
- \( \bar{t}_{d1} \) mean initial delay
- \( \bar{t}_{d2} \) knock-on delays (waiting time)
- \( p_{er} \) probability of equal-ranking train succession scenarios occurring
- \( \bar{t}_p \) buffer time
- \( \bar{t}_h \) mean minimum headway time
- \( \bar{t}_{h,er} \) mean minimum headway time for equal-ranking trains
- \( \bar{t}_{h,dr} \) mean minimum headway time different-ranking trains

To calculate the knock-on delays the initial delays have to be considered, which are divided into the probability of an initial delay \( p_{d1} \) as well as the mean initial delay \( \bar{t}_{d1} \).

The approach is based on different priorities of the trains, too.

The given waiting time calculated by the Strele formula will be compared to the allowed waiting time by using quality parameters which lead to the possible number of trains (cf. subsection 2.2, especially Fig. 2 and Fig. 3). Further information for this approach can be found in [20] and [21]. The Strele formula is implemented in the self-titled module of the LUKS software tool widely used in Germany [22].

As already shown in Fig. 1, a railway node can be subdivided into route nodes and sets of tracks. Route-node capacity can be calculated either as a multi-resource queue after [23] or, more straightforwardly, as a single-channel pseudo-system after [24].
To estimate the capacity of route nodes the Hertel formula is usable too. Further details on assessing the theoretical capacity and waiting times of route nodes can be found in [25] as well as in [21].

The set of tracks is divided into a multi-channel service system in which the number of tracks corresponds to the number of parallel service channels. To this end, all tracks that can substitute each other in a station are collated into sets of tracks. In contrast, if there are some tracks which could not collated into the set of tracks, it will be possible to have another set of tracks in the same. When considering train moves, it is possible to determine average service times and inter-arrival times with their corresponding scatter parameters for each set of tracks. The procedure after Hertel [11] (also in [2]) enables the capacity of a set of tracks to be determined on this basis.

To calculate the capacity of a set of tracks an multichannel queueing system has to be applied [21]. The number of channels corresponds to the number of tracks \( n_{Tr} \). In [11] Hertel developed a heavy-traffic model to calculate the waiting probability and the mean waiting times of such a system by using queueing theory.

The existing waiting time \( \bar{\tau}_W \) for the set of tracks is defined as follows:

\[
\bar{\tau}_W = \bar{\tau}_A \cdot p_0 \cdot \frac{\rho_{nTr} \gamma \phi}{n_{Tr}! (1-\phi)^2}
\]

where

\[
\phi := \left( \frac{\rho}{n_{Tr}} \right)^{\gamma} < 1 \quad 3.12
\]

\[
\gamma := \frac{2}{v_A^2 + c v_B^2} \quad 3.13
\]

\[
c := \left\{ \begin{array}{ll}
1 & ; v_A \geq 1 \\
\left( \frac{\rho}{n_{Tr}} \right)^{1-v_A^2} \cdot (1 - v_A^2) - v_A^2 & ; v_A < 1
\end{array} \right.
\]

\[
p_0 := \left[ \sum_{i=0}^{n_{Tr} \rho} \frac{\rho_{nTr} \gamma \phi}{n_{Tr}! (1-\phi)} \right]^{-1} \quad 3.15
\]

with:

- \( \bar{\tau}_W \) mean existing waiting time
- \( \bar{\tau}_A \) mean inter-arrival time
- \( p_0 \) represent inter-arrival probability
- \( \rho \) is defined as being the level of utilisation of the system
- \( n_{Tr} \) number of tracks
- \( v_A \) coefficient of variations of the inter-arrival times
- \( v_B \) coefficient of variations of the service times
To derive the capacity of a set of tracks an allowed limit for the waiting probability \( (p_{w,\text{all}}) \) exists. Under the background that a maximum number of train runs should be determined, the allowed and the existing \( (p_W) \) waiting probability have to be nearly equal.

\[
p_W \leq p_{w,\text{all}}
\]

and

\[
p_W = p_0 \cdot \frac{\rho^n}{(n-1)!} \cdot \frac{\gamma \cdot \phi}{1-\phi}
\]

To calculate the upper limits of the capacity, the Strele formula will be used to calculate the capacity of lines and also for the capacity of route notes separately [21], [26] and [27]. To calculate the set of track capacity the method of Hertel can be used [11], [21].

With the aid of predefined levels of service, it is then possible to disclose the corresponding capacity ceiling for each infrastructure element. Capacity established by means of the reference operating schedule in this way serves as a ceiling for the subsequent optimisation process (cf. Subsection 3.1). It is then already possible on the strength of the input data to stipulate further restrictions or exceptions. Congested railway infrastructure, for instance, may be subjected to the condition of being negotiated in a debased state (risk/deficient), whilst the rest of the network is to be dimensioned with a view to achieving optimum quality. This allows existing bottlenecks to be relieved in the run-up to optimisation, as impaired quality is generally accompanied by a higher capacity ceiling.

In the event that there being any passenger traffic, the calculated capacity has to be reduced to the number of passenger trains \( n_p \) on the relevant infrastructure elements. In addition, the limits for the optimisation of the sub-network are available now.

The optimisation procedure uses the formulae presented in Subsection 3.1 to establish the maximum number of train moves/train-paths in the sub-network, in the process identifying the route selection for freight trains that makes the best sense for the sub-network’s overall capacity.

There is a need, following optimisation, to check which routes freight trains have been accorded through the sub-network as a means of guaranteeing the maximum possible loading for the network assuming adherence to the previously defined quality criteria. Increasing the amount of freight traffic on infrastructure elements worked heterogeneously (i.e. by passenger and freight traffic) causes the composition of traffic (mixing ratio) to alter relative to the reference operating schedule. It is to be examined whether capacity ceilings have been lowered or raised as a result of the new train mix.
The more critical scenario obtains where they have been lowered and is further pursued below for illustrative purposes. A rising of the ceiling merely results in the maximum total number of trains being checked.

Where altering the mixing ratio on mixed-traffic lines (passenger and freight) yields a capacity ceiling that is no longer sufficient, it is necessary to remedy this and run a fresh optimisation loop.

An iterative operation is then run until the capacity limits cease to be subject to alterations. The algorithm stops when there is no further change in the limits of the infrastructure elements due to a change in the mixing ratio on this respective section. So the termination criterion is, if the change in the mixing ratio is negligible. This occurs by means of a limit-value $\varepsilon$. If the deviation is lower than this limit-value, then the result is the output. Due to tests on different railway lines we prefer that the limit-value $\varepsilon$ should be set to 2 %. If the change of the mixing ratio is higher this might lead to an adaption of more than 0.5 trains which results in a different upper limit of the capacity. Nevertheless it cannot be guaranteed that the algorithm terminates in all cases. This might occur of a single train is switched between two lines constantly. Therefore the algorithm stops after a fixed number $n$ of iterations. We suggest to set $k = 5$. In the investigated examples there were no changes in the upper limits or in the routing process of the solver after the fifth iteration.

For practical reasons, especially on larger networks it has to be checked, if the termination criterions are sufficient enough. In the considered network the iteration stops after the third run (cf. subsection 4.3) so the termination will not be exhausted, but on larger networks it cannot be excluded that another solution exist with better results within the defined limits.

### 3.3 Running-path search and profitable alternatives for rail passenger/rail freight traffic

Potential running paths between sources and sinks in the sub-network are used as variables for the targeted optimisation.

Suhl and Mellouli summarise their definition of a path as follows in [28]. A path in a directed graph is a chain $[e_1, e_2, ..., e_t]$ whose directed edges point in one direction only; to put it more precisely, there exist $t+1$ nodes $[\ldots]$. Hence, a path is conclusively represented by its sequence of nodes $[i_0, i_1, ..., i_k]$. $[\ldots]$ It is usually tacitly assumed that chains and paths are elementary, i.e., apart from any relation of equality between the initial and final nodes, nodes shall not recur in chains and paths.
As set out above, there is only limited scope for conducting running-path searches in rail passenger traffic. Rerouting local rail passenger services is severely hampered by the frequency of stops. There is greater scope for selecting alternative routes in long-distance passenger traffic paying due regard to running times. Where rerouting gives rise to longer running times, this may lead to the rail mode becoming less acceptable to passengers.

For the reasons just set out, attention is confined to rail passenger traffic in this procedure. The remit additionally entails reducing in advance, i.e. prior to optimisation, the maximum capacity established on the various sections by the available number of passenger trains. Radtke explains the course of action for modelling passenger traffic under the macroscopic NEMO traffic generation and re-apportionment model as follows in [18]: “With passenger traffic not always being adapted to pure demand, frequently due to political constraints, in reality low-ridership passenger trains are, for instance, run for operational reasons or to maintain a clockface pattern timetable throughout the day even though demand only actually peaks in the morning and evening. In a purely demand-based approach, by contrast, low-ridership passenger trains would not be generated at all and would not, therefore, consume train-paths either.”

In [29], Konanz states: *As is largely the case in real traffic management, rail passenger traffic is processed on a priority basis, thus correspondingly diminishing the output capacity of line sections. The remaining free output capacity then provides the latitude for running freight traffic as profitably as possible.*

The subsequent optimisation is thus performed, drawing on Radtke and Konanz, solely for freight traffic, with any rail passenger traffic being mapped as being prioritised.

In the following optimisation, study is to be given to the extent to which it is worthwhile working freight traffic through the entire sub-network as a means of achieving a maximum number of train moves. It is necessary to identify the points at which point-to-point freight traffic routes begin and end beforehand, as it is these that form the point of departure for the route search. In generalised terms, any running path is conceivable between these points and can be used at will as long as sufficient capacity is available on the running path (sub-segments negotiated).

Various route-search procedures exist with which to establish the shortest or fastest link between two points in a network. One of the most well-known is the Dijkstra algorithm [30]. It is necessary here, however, to output all possible routes below the highest boundary condition. All routes have in the process to be acyclical, hence there is no duplicate negotiation of nodes or lines.
At the same time, it needs to be examined whether the proposed point-to-point route or alternative running path is actually admissible in railway operations. Traction or line-related restrictions such as traction type (diesel or electric), tractive effort available (equilibrium speed over rising gradients), axleloads etc. need to be borne in mind (cf. [18]). Furthermore, there is a dynamic process for the running-path search depending on the parameters of the trains (e.g. length) related to the available length of tracks. It is possible to deal with. If a train cannot use a track e.g. because of the lengths, the running path cannot defined as an economic alternative. With other words, the running-path search will not detect this alternative. Otherwise if there are other train moves, with other train characteristics, to this track the capacity/limits have been checked after the optimisation by a new Strele/Hertel calculation.

Assuming the above boundary conditions are met, it is then additionally necessary to assess the proposed route’s profitability. This involves, amongst things, citing the shortest paths by distance or time, since these are the basis upon which the train operating company is charged for train-paths. Selecting the most energy-efficient or least polluting path may also, moreover, be of relevance to the assessment of profitability (cf. [18]).

All profitable alternatives found are to be used as a basis for optimisation. Should it not be possible to find alternatives, then there is no scope for rerouting freight trains on this point-to-point route.

### 3.4 Optimisation

The actual optimisation operation is performed by a commercial solver. The number of possible train moves on a sub-network is now maximised by forming a set of linear equations. The mathematical foundations for this are set out in Subsection 3.1 above.

The solver comes up with a maximum possible number of train moves on the sub-network. The outcome also covers the running paths selected and the number of train moves on these running paths. This means the extra freight trains on every infrastructure element. Maximisation does not, however, necessarily entail only selecting a single running path per source-destination route. It is additionally possible to output the loading for each individual infrastructure element. Finally, the new mixing ratio on the infrastructure elements can be calculated.

### 3.5 Altering the mixing ratio and adapting capacity

As already illustrated in Fig. 3, train moves or, respectively, the mixing ratio have a significant impact on capacity. With the number of freight trains rising as a result of
optimisation, the mixing ratio is inevitably altered (assuming the line is also worked by passenger traffic), which may lead to capacity limits being adjusted. A process of interpolation is set out in the following Subsections.

Any alteration of the train mix generally only occurs on mixed-traffic lines, since on dedicated passenger or freight lines the mixing ratio of passenger trains $p_P$ continues to stand at 100 % (passenger lines) or 0 % (freight lines) following optimisation.

Fig. 5 shows illustrative percentage variations in the mixing-ratio, arrived at for a generic line. Computations were conducted for each interpolation node (at 20% stages). The total number of trains involved in workings on this line section was kept constant and only the mixing-ratio percentage was altered. The upper curve represents the pattern of line capacity assuming only local-passenger and freight traffic were run, the lower curve relating instead to a mix of long-distance-passenger and freight traffic. Where both long-distance and local passenger services are run, capacity lies between the two curves. Capacity is represented as a function of the passenger-train fraction.

\[ \text{capacity} = \begin{cases} 
\text{long-distance-passenger traffic} \\
\text{local-passenger traffic} 
\end{cases} 
\]

**Fig. 5: Illustrative representation of output capacity for given mixing ratios**

The exact limit of the capacity can be calculated again (Strele/Hertel) by changing the mixing ratio after the optimisation. It also is possible by interpolating the values given to establish output capacity with reference to the mixing ratio$^4$.

It becomes apparent that mixing fast and slow services yields differing capacity curves. With freight and local-passenger traffic travelling at comparatively homogeneous speeds, this curve lies above the curve for the applicable comparison of these types of traffic with long-distance passenger traffic.

$^4$ The trend line posits a third-degree polynomial.
3.6 Evaluation of results

Once optimisation has been concluded, the results need to be evaluated. First the number of implementable train moves/train-paths in the overall network can be adopted as a key indicator. This is made up of the optimised number of (freight) train moves and the base loading (passenger trains). It is additionally possible to establish the optimum freight-train running paths for the network, even if these do not necessarily equate to the shortest running paths (as a function of time and/or distance).

Loading per infrastructure element can be arrived at by contrasting the capacity available for optimisation purposes with that actually consumed. Use can thus be made of recognised indicators (quality and track-occupation ratio) to point up congested railway infrastructure and compare them with deviating infrastructure or operating-schedule variants.

Further indicators can take the form of comparing and gauging the number of train-paths in the sub-network with the aid of length or run-timing factors. If consideration is additionally given to the sub-segments in a network, the capacity they consume can for instance be itemised relative to consumption on the network as a whole.

These potential means of evaluating results can also be used to justify the longer-term expansion, redesign or downsizing of the network citing cost/benefit ratios as well as to ensure resources are effectively deployed.
4 Illustrative Computation

The methods and theories presented thus far are elucidated in greater depth with the aid of an illustrative computation in this Section.

4.1 Sub-network and running paths

A small model network consisting of seven nodes linked by the lines represented in Fig. 6 was set up as an illustrative network. The respective running times and distances between and within nodes are further posited as being known.

![Illustrative network](image)

The respective sets of tracks are numbered and the overall route node is denoted by means of the letters “a” and “b” (if there are any other route nodes in the same node, it will be possible to spend some more letters to define them). Continuous lines symbolise double-track railway lines, whereas the dotted line between nodes 5 and 6 indicates single-track working only. With the exception of the lines between nodes 2 and 6 and between nodes 5 and 6, which are dedicated freight lines, all lines considered here are regarded as being mixed-traffic lines.

With passenger traffic being given preferential treatment as set out in Subsection 3.3, the base loading is predefined on the respective lines and nodes.

<table>
<thead>
<tr>
<th>Long-Distance Services</th>
<th>Local Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point-to-point route</td>
<td>Point-to-point route</td>
</tr>
<tr>
<td>Nodes</td>
<td>Nodes</td>
</tr>
<tr>
<td>1</td>
<td>4,6,7</td>
</tr>
<tr>
<td>2</td>
<td>7,6,4</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1: Running paths of point-to-point routes for passenger traffic
The sources and attendant sinks for freight traffic were taken to be as follows:

- Point-to-point route 1: from node 4 to node 3
- Point-to-point route 2: from node 4 to node 7
- Point-to-point route 3: from node 7 to node 2

The number of train moves is to be conducted in this sub-network on the point-to-point routes detailed above, with various alternatives being predefined for each point-to-point route (cf. Table 2).

<table>
<thead>
<tr>
<th>Alternative</th>
<th>Nodes</th>
<th>Alternative</th>
<th>Nodes</th>
<th>Alternative</th>
<th>Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4,6,5,3</td>
<td>1</td>
<td>4,6,7</td>
<td>1</td>
<td>7,6,2</td>
</tr>
<tr>
<td>2</td>
<td>4,6,7,5,3</td>
<td>2</td>
<td>4,6,5,7</td>
<td>2</td>
<td>7,5,3,2</td>
</tr>
<tr>
<td>3</td>
<td>4,6,2,3</td>
<td>3</td>
<td>4,1,2,6,7</td>
<td>3</td>
<td>7,5,6,2</td>
</tr>
<tr>
<td>4</td>
<td>4,1,2,3</td>
<td>4</td>
<td>4,1,2,3,5,7</td>
<td>4</td>
<td>7,6,4,1,2</td>
</tr>
</tbody>
</table>

Table 2: Running paths of point-to-point routes for freight traffic

In this example, four alternatives are listed for each of three point-to-point routes, hence twelve variables are used in the objective function. For simplicity’s sake, the indices for the point-to-point route (1; 2 or 3) and the number of the alternative (1; 2; 3 or 4) are assigned to each variable (cf. Table 2).

The speed of the trains/on the line is also known and can used as input for the minimum headway times, so that there is a simplified approach (time = length/speed) to calculate the running time on the alternative routes.

4.2 Course of action

Given that the aim here is to maximise train moves, the objective function in general is given in subsection 3.1 (cf. formula 3.1).

In this example only the length of the train courses is used to weight the objective function. The weighted factors of the formula 4.1 can be calculated by the given distances between the nodes (cf. Fig. 6) the length of the route nodes and the length of the set of tracks. Simplified in this example every route node has a length of 0.5 km and the length of the set of tracks will be assumed to 1 km.

For each alternative (cf. Table 2) the distance between source and attendant sink can be calculated by adding the distances of the lines, the route nodes and the set of

---

5A train negotiating running path 1,2 thus passes through nodes 4, 6 and 7 in that order.
tracks. Furthermore the shortest train course of every point-to-point route can be appointed and due to this the ratio for the alternative to this shortest train course can be calculated. Because of the maximisation the reciprocal will be needed so that the shortest path of the point-to-point route will have the highest weight (equal 1.0). The other weights will be less than 1.0.

The weighted objective function for this set of equations under formula 3.1, the given distances and also to the running paths of Table 2 is thus:

\[ \max_c \mathbf{c}^T \cdot \mathbf{n}_T = 0.878 n_{r_{1,1}} + n_{r_{1,2}} + 0.951 n_{r_{1,3}} + 0.707 n_{r_{1,4}} + 0.433 n_{r_{2,1}} + 0.619 n_{r_{2,2}} + 0.928 n_{r_{2,3}} + n_{r_{2,4}} + 0.650 n_{r_{3,1}} + 0.750 n_{r_{3,2}} + 0.875 n_{r_{3,3}} + n_{r_{3,4}} \]

\[ n_{r_{n,m}} \] Vector of the number of train moves on the point-to-point route \( n \) and the alternative \( m \) (cf. Table 2)

It needs to be ensured for all equality constraints that the sum of all train moves over the respective infrastructure element is lower than the capacity ceiling for this element. Due to this the equation has the following basic form:

\[ A \cdot \mathbf{n}_{r_{n,m}} \leq C \]

with

\[ n_{r_{n,m}} \geq 0 \]

\[ C \geq 0 \]

\( C \) is defined as the upper bound for the capacity of each infrastructure element (\( C_L \) for lines, \( C_{SoT} \) for set of tracks and \( C_{RN} \) for route nodes). The matrix \( A \) is segmented to the following sub-matrices:

\[ A := \begin{pmatrix} A_L \\ A_{SoT} \\ A_{RN} \end{pmatrix} \]

The sub-matrices involve the subjects for the lines \( A_L \), for the set of tracks \( A_{SoT} \) and for the route nodes \( A_{RN} \). These sub-matrices have the entry 1 if the infrastructure element is assigned by a train run. By terminal stations or double assigning of route nodes the entry is 2. Otherwise, if there is no train run through the infrastructure element the entry is 0.
If the subjects for the lines are considered, the following equality constraint needs to be observed:

\[
A_L \cdot n_T \leq C_L
\]

\[
\begin{pmatrix}
L_{1,2} & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
L_{2,3} & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
L_{3,5} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
L_{2,6} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
L_{4,6} & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
L_{5,7} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
L_{6,7} & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
L_{4,1} & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\
L_{3,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
L_{6,2} & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
L_{5,3} & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
L_{6,4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
L_{7,5} & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
L_{7,6} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
L_{5,6} & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}
\Rightarrow
\begin{pmatrix}
33 \\
15 \\
40 \\
27 \\
28 \\
36 \\
20 \\
28 \\
15 \\
27 \\
40 \\
28 \\
36 \\
19
\end{pmatrix}
\leq \begin{pmatrix}
4.6
\end{pmatrix}

with

\[L_{i,j}\] variable for the line between node \(i\) and node \(j\)

The set of tracks in a node thus contain all variables for the train moves that work them:

\[
A_{soT} \cdot n_T \leq C_{soT}
\]

\[
\begin{pmatrix}
SoT_1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
SoT_2 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
SoT_3 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
SoT_4 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\
SoT_5 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
SoT_6 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\
SoT_7 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0
\end{pmatrix}
\Rightarrow
\begin{pmatrix}
50 \\
80 \\
75 \\
55 \\
72 \\
95
\end{pmatrix}
\leq \begin{pmatrix}
4.7
\end{pmatrix}

with
$\text{SoT}_i$ variable for the set of track of the node $i$

With regard to the station throat in the nodes, the following conditions are to be met:

$$A_{RN} \cdot n_T \leq C_{RN}$$

\[
\begin{pmatrix}
RN_{1,a} & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\
RN_{2,a} & 0 & 0 & 1 & 1 & 0 & 0 & 2 & 1 & 1 & 0 & 1 & 0 \\
RN_{3,a} & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
RN_{4,a} & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\
RN_{5,a} & 2 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\
RN_{6,a} & 1 & 1 & 2 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\
RN_{7,a} & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\
RN_{8,a} & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
RN_{9,a} & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
RN_{10,a} & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
RN_{11,a} & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
RN_{12,a} & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\
RN_{13,a} & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\
RN_{14,a} & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\
\end{pmatrix}
\underbrace{\begin{pmatrix}
n_{T,1,1} \\
n_{T,1,2} \\
n_{T,1,3} \\
n_{T,1,4} \\
n_{T,2,1} \\
n_{T,2,2} \\
n_{T,2,3} \\
n_{T,2,4} \\
n_{T,3,1} \\
n_{T,3,2} \\
n_{T,3,3} \\
n_{T,3,4} \\
\end{pmatrix}}_{\text{}} \leq \begin{pmatrix} 58 \\ 90 \\ 70 \\ 54 \\ 100 \\ 65 \\ 75 \\ 48 \\ 80 \\ 92 \\ 53 \\ 100 \\ 58 \\ 72 \end{pmatrix}^{4.8}
\]

$RN_{i,r}$ variable for the route node $r$ of the node $i$

A further special feature that can be mapped when taking this course of action is the reversing of train moves in a station. The additional condition then is that there must be scope for double negotiation of a station throat (RN) (important for the extra dwell time in nodes). For example a train runs into a station, changed the direction in the set of tracks and leave the station in the same direction, so that there are two train runs from the same train in the route node.

This is thus a course of action that enables all equality constraints to be imposed for lines, sets of tracks and route nodes (cf. formula 4.6 - 4.8).

### 4.3 Capacity limits, results following optimisation and iteration

As can be seen in Formulae 4.6, 4.7 and 4.8, the capacity ceilings (analytically determined capacity minus the base loading for passenger traffic) are placed on the righthand side of the equation\(^6\). The freight trains working the infrastructure element must be fewer in number or equal to the capacity available. These capacity ceilings can be gleaned for each line, node and route node from the applicable Tables (Table 3; Table 4; Table 5). Furthermore the system of equations has to be solved. The solution

---

\(^6\) For reasons of simplicity capacity ceilings are given in whole numbers for the illustrative computation. These capacity values can be decimal numbers as well.

24
space can be restricted to integer, which will be a harder task, or it also is possible to find continuous solutions. It is thus possible for illustrative purposes to set an available capacity of 33 trains per period under review for the equations covering the line from node 1 to node 2 shown in the previous Sub-section. 75 trains/train moves can be processed in the period under review over the set of tracks in node 3 and 70 train moves over the attendant route node on the lefthand side. The lines L_{1,4} and L_{2,1} are not required in the running paths and are not therefore included in Table 3.

![Table 3: Capacity ceilings and results of optimisation on lines](image)

![Table 4: Capacity ceilings and results of optimisation in sets of tracks](image)

![Table 5: Capacity ceilings and results of optimisation in route nodes](image)
The results of, or capacity consumed following, optimisation are displayed in the third row of each category (lines, set of tracks, route nodes). The predefined ceilings for the infrastructure mean a total of 88 freight train moves are possible in the sub-network.

It is possible, that the algorithm does not find a solution. This means that there is no possibility to reroute freight trains on an infrastructure element or in the whole network. Furthermore, all entries in the matrix \( A \) (cf. formula 4.2), the number of trains \( n_T \) (cf. formula 4.3) as well as the limits of capacity are more or equal than 0. In the case, that there are linear dependences the set of solutions would be infinite. In the illustrated example there is no linear dependence between the columns in the matrix \( A \) and as a result the system of equations is solvable.

The next step involves checking individually for each infrastructure element whether its mixing ratio has been altered. Assuming that Fig. 5 mirrors the output capacity under the various mixing ratios for line section \( L_{4,6} \) and that a base loading for long-distance traffic is available alongside optimised freight train moves on this section, then it is necessary to examine the consequences this has for capacity limits. As can be seen in Table 1, 4 long-distance trains work this line section. For the purposes of the reference operating schedule, additional 16 freight trains were specified. This means that the passenger-train fraction \( r_{PT,pres.} = 20 \% \). A total of 20 train-paths/train moves are run on this section, though a capacity calculation has demonstrated that 32 train moves would be possible given this reference mixing ratio (cf. Fig. 5). It is assumed on the basis of this computation that a total of 28 train-paths are available for the purpose of optimising freight traffic (cf. Table 3; \( L_{4,6} \)). Following optimisation, 27 of a possible 28 train-paths were required for freight traffic. Together with the base loading for long-distance traffic, a total of 31 trains now work the line section \( L_{4,6} \). The mixing ratio \( r_{PT,new} \) now stands at 12.9 % (cf. Table 6).

With the passenger-train fraction now lower than in the reference operating schedule, it can be discerned from the curve pattern in Fig. 5 that there is no conflict with the upper capacity limit. If the attendant output capacity is read out/calculated, a value of around 35 train moves is arrived at on this section, which corroborates this assumption.

If, as a means of benefiting total output capacity for the network as a whole, freight trains are routed in such a way that the proportion of freight trains on mixed-traffic lines falls, capacity limits may nevertheless deteriorate. It is possible for comparative purposes to consider the opposite direction between nodes 4 and 6 (\( L_{6,4} \)). The same reference operating schedule applies (4 long-distance trains; 16 freight trains; \( r_{PT,pres.} = 20 \% \)). Following optimisation, however, only 5 freight trains are routed over this section of line (cf. Table 3). This means that a mere 9 trains work the line and the fraction \( r_{PT,new} \) now stands at 44.4 %. The line’s output capacity now moves in the
opposite direction and only 25 train moves can be conducted in all. With just 9 trains negotiating this section following optimisation, however, this is likewise no hindrance in the present example.

There will be an iteration process to the whole network. For this example the iteration depended on Fig. 5 will be considered in detail for the following lines (both directions):

- Line from node 2 to node 6 with 10 local service trains and 23 freight trains
- Line from node 6 to node 7 with 4 long-distance trains and 4 freight trains

Table 7 shows the results after the first optimisation run again and in addition the ascertained limits as well as the consumed capacities after the second and third loop of optimisation. Like outlined above, due the appointed reference operating program the mix ratio can be calculated (cf. Table 6) and therefore the given capacities can be calculated for the highlighted lines (cf. Table 7 values in bold).

After the optimisation loops the mix ratio for the passenger trains \( r_{PT} \) has been changed as follows.

<table>
<thead>
<tr>
<th>Line</th>
<th>( L_{2,6} )</th>
<th>( L_{6,2} )</th>
<th>( L_{4,6} )</th>
<th>( L_{6,4} )</th>
<th>( L_{6,7} )</th>
<th>( L_{7,6} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_{PT,1} )</td>
<td>30 %</td>
<td>30 %</td>
<td>20 %</td>
<td>20 %</td>
<td>50 %</td>
<td>50 %</td>
</tr>
<tr>
<td>( r_{PT,2} )</td>
<td>62.5 %</td>
<td>33.3 %</td>
<td>12.9 %</td>
<td>44.4 %</td>
<td>30 %</td>
<td>30 %</td>
</tr>
<tr>
<td>( r_{PT,3} )</td>
<td>100 %</td>
<td>33.3 %</td>
<td>12.9 %</td>
<td>21 %</td>
<td>22.2 %</td>
<td>13.3 %</td>
</tr>
<tr>
<td>( r_{PT,4} )</td>
<td>100 %</td>
<td>33.3 %</td>
<td>12.9 %</td>
<td>21 %</td>
<td>22.2 %</td>
<td>13.3 %</td>
</tr>
</tbody>
</table>

**Table 6: Alteration of the percentage of passenger trains**

The mix ratios in Table 6 are used to interpolate again for every new optimisation loop and through the changing of the mix ratios the upper limits have to be adjusted. Table 7 shows the results after the different iteration loops. After the second optimisation it is shown that the solver has rerouted some freight train through another line e.g. on \( L_{2,6} \) (cf. \( C_{L,cons_2} \)).
Table 7: Capacity ceilings and results of optimisation on lines after 2 more loops

After the third optimisation loop the solver has not routed the trains again and furthermore the required capacities on each line are the same. The values are the consumed capacities marked between the red lines. The optimiser can be stopped at this moment.

This course of action is to be verified following optimisation for each line – exemplary it was outlined above for three lines – set of tracks and route node. A process of iteration with the new ceiling is to be launched should limiting factors be encountered. Analysis is then to be conducted in such an instance into whether a solution to the equation problem tends towards one value or else into the degree of accuracy $\varepsilon$ (cf. Fig. 4) following which iteration can be ended. A suggestion of five iteration loops and $\varepsilon < 2\%$ was given in subsection 3.2.

As an alternative in the case of a decrease of the capacity limits it could be analysed if another quality step could be used to realise the optimized number of trains (e.g. from optimal down to risk or deficient).

5 Summary and Outlook

Establishing network-wide capacity constitutes a key developmental goal of strategic network planning. It is possible by addressing the capacity of railway lines and railway nodes to effect the meaningful optimisation by automated means of running paths and capacity consumption for entire networks.

In order to attain this goal, a macroscopic network is described by means of sets of linear equations and its maximum determined with the aid of a solver. In addition, alternative freight-traffic running paths between sources and sinks in the sub-network are to be output in accordance with previously defined boundary conditions such as lengths or running times. The solver uses these running paths to globally optimise the network in respect of the objective function.

Optimisation provides data on possible train-paths within the sub-network. Building upon these it is possible to evaluate train-path consumption, the loading of individual infrastructure elements and the most worthwhile running paths through the sub-
network. There is scope for comparing differing infrastructure or operating-schedule scenarios, thus allowing pronouncements to be made regarding current or future bottlenecks and their elimination or prevention.

Given that, in most cases, resources for engineering works are limited, inferences can be drawn with the approach presented here as regards the most worthwhile deployment of such resources. The approach presented is capable of pointing up the direct impact that expansion, redesign or downsizing measures have on a railway network’s total output capacity. Hence, though developing a line’s output may enhance capacity locally, there are circumstances, such as when an adjoining node limits the number of train moves, under which it will not benefit the network as a whole.

The present work is modular in composition so as to enable further boundary conditions and evaluations to be added to the basic system at any time. Additional areas of application worthy of note include an extension of the path-pricing system whereby the infrastructure manager subsidises “unattractive” running paths for the train operating company until such time as alternative routes are selected. This, too, is a means of bringing relief to congested railway infrastructure.

6 References


MICHAEL KÜMMLING, PETER GROßMANN, JENS OPITZ: Maximisation of homogenous rail freight train paths at a given level of quality - EURO 2015 Glasgow – Überprüfungsdatum 2016-02-15


RADTKE, Alfons: EDV-Verfahren zur Modellierung des Eisenbahnbetriebs. Hanover: Institut für Verkehrswesen, Eisenbahnbau und -betrieb Universität Hanover; Eurailpress, 2005


