A concept note for railway timetabling to rationalize and improve capacity utilization

Prof. Narayan Rangaraj and Prof. Madhu Belur, Indian Institute of Technology Bombay
narayan.rangaraj@iitb.ac.in, belur@iitb.ac.in 8th January, 2018

Abstract

This report deals with some principles of timetabling and defines some notions of “good” quality of a timetable with respect to various stakeholders involved in a railway network (including passengers as end-users, freight customers and the network operator). This is followed by an elaboration on various parameters and how they affect the quality of a timetable. To the extent possible, the important rail section between Howrah and New Delhi (and more particularly, the congested section between Mughalsarai and Ghaziabad) is taken up for analysis to illustrate concepts. Recommendations for improvement in capacity utilization, junction-congestion, bottleneck section congestion and distribution of slacks and allowances are made: both specific to the Howrah-New-Delhi route and in general.

Table of contents

Abstract 1

1. Principles of rail-traffic timetabling 2

2. Current timetable observations: HWH-NDLS 7

3. Capacity utilization of sections handling mixed traffic 10

4. Analysis of how various parameters affect congestion 14

5. Junction congestion analysis 18

6. Recommendations 21

7. Conclusion 23

References 24

Acknowledgements

The study was sponsored with financial support of NITI Aayog, Government of India and conducted by Indian Institute of Technology Bombay, Mumbai. The content of this report (data collection, analysis, simulation and discussion) has been the outcome of the work carried out by many past and present M.Tech. students and project employees at IIT Bombay, in particular: Aayush Aggarwal, Soumya Dutta, Sudhanshu Shekhar, Ashutosh Khare, Vipin Maurya, Deepak Miglani, Manoj Patel, Anuj Sahu, Abhishek Singh, P. Sudarshan and R. Vidyadhar. We thank them for their contribution in this report.

Disclaimer: Indian Institute of Technology Bombay has received a grant under the Research Scheme of Niti Aayog, 2015 to conduct the study and compile this document. However, Niti-Aayog shall not be held responsible for findings or opinions expressed in the documents prepared. This responsibility rests with Indian Institute of Technology Bombay, Mumbai.

1. Principles of rail-traffic timetabling

The UIC 406 document [11] compiled by the International Union of Railways proposes a framework that is broadly applicable to many of the professional railway operating environments, worldwide. It identifies four considerations in capacity assessment:

- throughput (number of trains),
- heterogeneity (mix of trains),
- stability (robustness) and
- traversal time (service).

Out of these four considerations of capacity assessment, throughput and heterogeneity are directly conflicting (i.e., it is possible to have more number of trains if one reduces the variety and streamline the traffic and conversely, the more the types of trains one wishes to run, the smaller would be the number of trains that one can typically operate). Similarly, traversal time and stability are also directly conflicting - the less the (planned) traversal time, the less room there is for recovery, punctuality and stability of schedules, and vice versa. Similar considerations apply to timetable construction and the performance of traffic on a given part of the network.

The following figure shows this for two types of infrastructure usage:

A. metro trains (like suburban trains, dedicated usage by homogenous non-overtaking trains)
B. mixed traffic (i.e. trains with different running/halt characteristics)

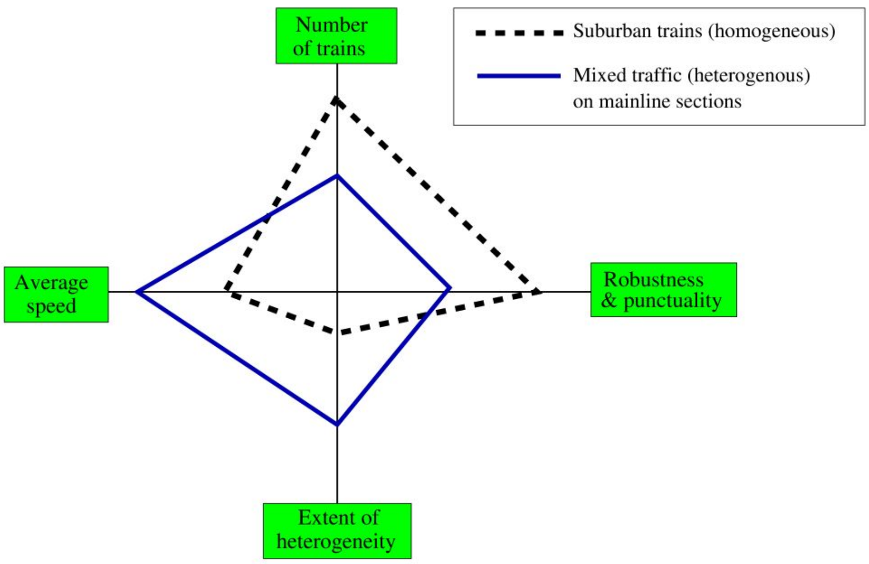

Fig 1.1: The balance of railway capacity (adapted from [11])

The interaction effect between other pairs of performance measures and the combined behaviour is less straightforward to quantify and depict. We summarize about these parameters and what they depend on.

Traversal time: Traversal time on a given section depends on

(a) length of section to be traversed
(b) achievable speed of rolling stock (moderated suitably to allow for realistic running)
(c) planned halts

(d) time for accelerating and decelerating to and from halts to allow for overtakes
(e) explicit allowances to account for temporary speed restrictions and traffic hindrances.

An analysis of these factors on important sections of Indian Railways is an ongoing exercise, which is done at different levels - divisional, zonal and at the railway board. This exercise can benefit much from data analysis and regular assessment of operational performance.

Congestion and Overtaking: The current method of estimating line congestion is using a variant of standard formulas that are relevant for a single stream of traffic, and are not relevant for mixed traffic where there are significant interactions between different types of trains (speeds, lengths and priorities). In an environment with different types of trains,

- either there is overtaking in which case there are slow movements to allow for this, resulting in some loss of capacity on critical sections,
- or there is no overtaking and then the faster trains are constrained to move at below their maximum speeds, possibly resulting in some loss of potential throughput.

For example, on extremely high-density sections such as Mumbai suburban, the latter strategy is followed and even the high priority trains capable of 130 kmph run at considerably lower speeds and overtaking is completely avoided. The resulting loss of traversal time for these high priority trains over these relatively small sections is considered acceptable because the overall throughput is improved due to the absence of overtaking in the high-density section.

Junction/Terminal Operations: A significant aspect of infrastructure that is under-studied is the impact of junction/terminal operations on sectional running. It is a commonly encountered phenomenon that trains (both passenger and freight trains) undergo substantial waiting just short of major junctions. It is worth emphasizing that it is not just platform or running line resources at junctions that are the bottleneck to throughput, but the crossovers and track resources which provide access to various parts of the junction that are often the bottleneck to throughput. The effect is most significant when there are reversals, loco changes and other less-streamlined operations at junctions. This analysis is pursued in more detail in Section 5 below.

The distance-time chart below contains the time-axis as the horizontal axis and the distance along the vertical axis.

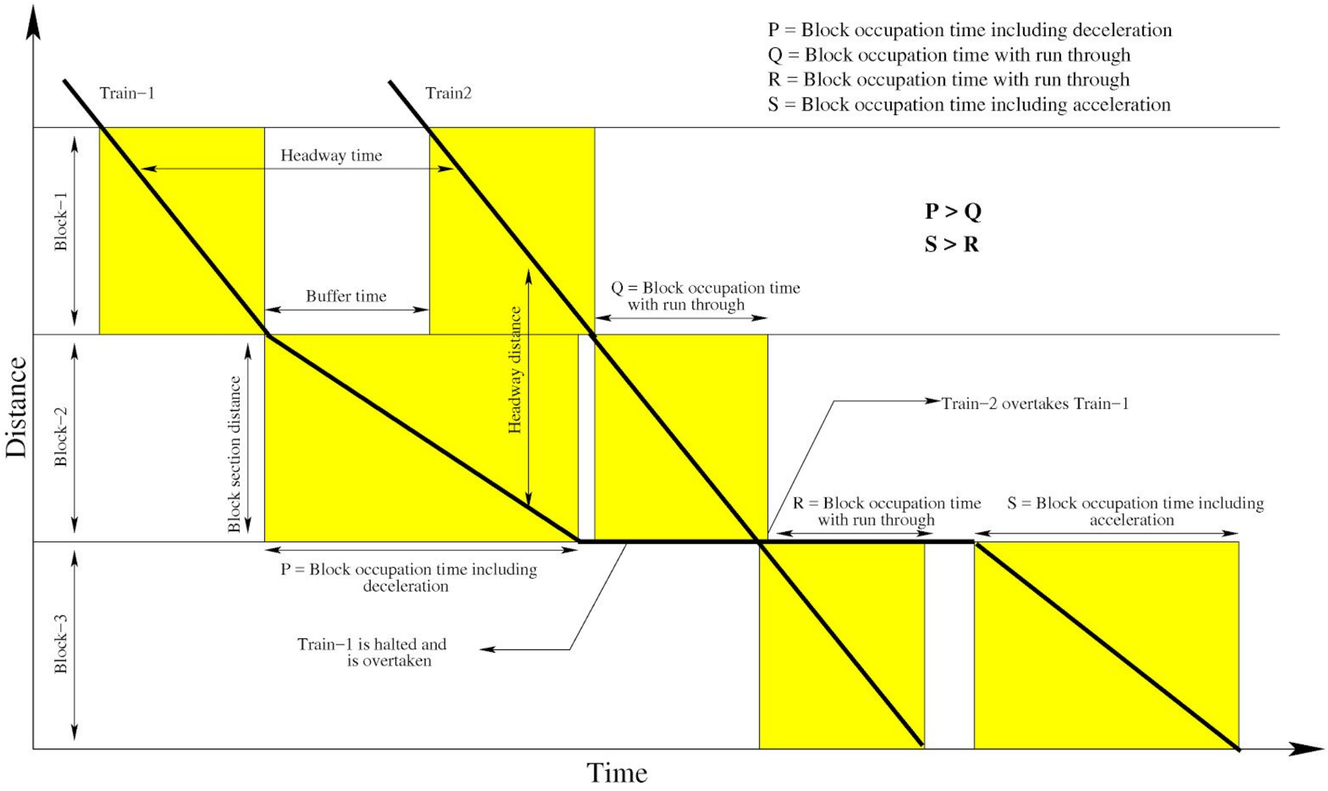

Fig 1.2: Distance-time chart showing headway distance, headway time, and overtake of Train 1 by Train 2

Good quality paths for <u>passenger</u> trains

With passengers as primary beneficiaries of the railways, below are some notions defining quality of a passenger train path.

- High traversal speed between stations
- Low overtakes by other trains
- Halts only at scheduled stations (and not for just being overtaken by other trains)
- Punctuality of departure

The above points are clearly motivated by the fact that railways is a *service*.

Good quality paths for <u>freight</u> trains

It is noteworthy that freight trains help reduce passenger fares since freight traffic earns revenue to the railways: thus enabling freight paths eventually helps all passengers. Further, prioritizing and enabling *rail* freight instead of *road* traffic helps a cleaner atmosphere for the country since rail usage is cleaner than other modes of transport. Thus the benefit of freight paths is eventually transferred to the country.

1. High traversal speed between stations
2. Low overtakes by other trains
3. Halts only at scheduled stations (and not for being overtaken by other trains)

The above points help achieve more number of freight paths and better delivery of freight service (by quicker delivery), and delivery at a lower cost (by lower inventory, lower usage-time of wagons, lower usage of locos and crew). A process of timetabling by enabling many good freight paths in a mixed-rail-traffic scenario is inevitable since, even after the Dedicated Freight Corridor is commissioned on select-sections, it is unreasonable and unrealistic to have separate infrastructure all across the network.

Robustness and punctuality

A railway system is influenced by many uncertainties and delays caused by malfunctions or deviations from plans. Delays directly caused by disturbances are called primary delays. The delays caused to a train due to primary delays of other trains are called secondary delays. Due to interdependencies in railway system, a large part of delays in congested systems are actually secondary delays.

Some of the major causes for disturbances and primary delays are

1. Planning - while planning the timetable, error could arise, such as the following: dwell time may be planned shorter than required, length of trains not accounted for and acceleration and deceleration parameters not properly accounted for.
2. Infrastructural failures - malfunctioning switches, broken catenary, failing signals, and maintenance works taking longer than the planned times.
3. Human factors - Driver behaviour and response time in manual operations of route setting and other actions add to stochastic primary delay.
4. Weather Condition and others - poor visibility in fog leads to increase in braking distance and decrease in acceleration rate. In some rail section, cattle run over and animal conflict are actually a serious issue and as of now, no effective solution has been found.

The above factors may contribute a small fraction of total delay but due to the delay propagation it spreads in the railway network in both time and space and eventually causes large secondary delays. Some of major causes of secondary delays are

1. High capacity utilization and thus smaller headways
2. When a train reaches terminal station with a delay such that time required to change rolling stock for the use of other subsequent train is not enough, then delay propagates.
3. When a train reaches late, the planned crew at the interchange point also gets late and can not serve their duties in full, thus leading to delay of subsequent train.

Train data usually related to delays include both primary and secondary delays. From robustness considerations, a good timetable is one in which the primary delays get absorbed quickly and do not impact other trains and cause secondary delays.

Criteria for a “quality timetable” for mixed rail traffic scenario

The above notions of quality of passenger trains and freight trains are to be kept in mind when defining the quality of a <u>timetable</u>. These criteria are to ensure higher throughput (in terms of passenger trains and freight trains), better quality paths for passenger and freight trains, and thus better usage of rail-infrastructure.

A. Low traversal times (time spent in the system)
B. High running speeds of trains
C. Minimizing overtakes between trains
D. Punctuality of <u>departure</u> at stations with scheduled halts
E. Punctuality of <u>arrival</u> at stations with scheduled halts
F. Resilience/robustness of the timetable due to unanticipated disruptions

Point E above about punctuality of arrival at a station is important to ensure that trains arrive neither late nor **<u>earlier</u>** than scheduled time: arriving **<u>earlier</u>** uses up valuable infrastructural resources - this detrimental effect of arriving too early appears to be underestimated. On the other hand, some allowances are required and inevitable to ensure Point F above and thus there is a trade-off between points E and F. This aspect is analyzed in more detail in Section 4 below.

To illustrate the above notions, we consider timetables and train operations on a congested route on the Indian Railway network, taking the example of Delhi - Howrah. Our objective is to come up with a set of principles for good timetabling on mainline sections handing mixed traffic and illustrate this to the extent possible on this particular section.

On the NDLS-HWH section, there are only a few trains that go end to end (i.e. from some origin in Delhi area to some destination in Kolkata area). We consider the route: Delhi area (DLI/NDLS/NZM/ANVT) and Howrah divided into 10 sections:

DLI/NDLS/NZM/ANVT - GZB (section 1),	GZB - TDL (section 2)	TDL - CNB (section 3)	CNB - ALD (section 4)	ALD - MGS (section 5)
MGS - Gaya (section 6)	Gaya-DHN (section 7)	DHN-ASL (section 8)	ASL-BWN (section 9)	BWN-HWH/SDAH (section 10)

Table 1.1: Subsections of the New-Delhi to Howrah route

Trains are divided into the following classes:

Class 1: Rajdhani, Duronto, Shatabdi, Garib Rath, Tejas etc. Max speed 130 kmph
Class 2: Express trains, Max speed 110/130 kmph
Class 3: Stopping passenger trains, Max speed 70/90 kmph
Class 4: Freight trains, Max speed 60-75 kmph

2. Current timetable observations: HWH-NDLS

We analyzed the train running data from http://indiarailinfo.com and we report some key parameters for the ten sections listed in Table 1.1. For each of these 10 sections, analysis of average traversal speeds for each class of trains, with halting times at stations included and without. Sections 1 and 10 above have significant suburban/commuter traffic and also have multiple line working and several termination options for passenger trains serving those areas (Delhi and Kolkata). A comprehensive analysis of those options, including terminal infrastructure, is beyond the scope of this report.

Analysis of allowances

Allowances are extra time values given within a timetable for each train at regular distances (or at the end): this is given to help maintain punctuality of operation of a given timetable. This is analyzed in detail later below. On each section 1-10, allowances given per kilometer tabulated, per class of train. This data is available through the working timetable for sections 2, 3, 4, and 5 and the allowances are likely to be similar for the remaining sections.

Analysis of planned overtakes

All overtakes tabulated in each section 1-10 - these are listed by “x/o” in the website http://indiarailinfo.com Of particular significance (and to be avoided) are the overtakes between trains of a similar type. Overtakes can also be derived from working timetable by analysis of arrival and departure times at a station. Note that overtakes can happen only at a station. Details about this in included in the following sections.

Analysis of bottleneck resource

For our purposes, a bottleneck resource is one whose improvement will result in the improvement of the entire section as a whole. There are different ways in which one could identify a bottleneck resource.

- Section with maximum utilization
- Section with minimum speed per hour (of freight trains, but also of passenger trains)
- After identifying the section definition and signaling regime, the block section that takes longest to traverse

A possible candidate for a bottleneck resource is the track section between Mughalsarai and Allahabad (about 150 km). An indication of this is the low *timetabled* speeds of all categories of trains - see below. This needs to be independently verified in as many dimensions as possible. It is also widely believed that junction areas (in particular, Allahabad) could be a bottleneck for the entire section. A separate junction analysis is included in Section 5 below.

Analysis of train composition

In order to maximize the benefits of paths, on bottleneck sections, as a general policy, trains of high capacity should run. We find that over the years, suboptimal train compositions are either absent or in a very small number (one or two) and stopping passenger trains have been reduced to a small number to maintain minimal connectivity of wayside stations.

Analysis of speeds, halts and overtakes

The tables below contain data for ten sections between NDLS area and HWH for various types of trains. For each type of train, we show average speeds over each section and the number of overtakes.

Rajdhani/Shatabdi/Garibrath	NDLS-GZB	GZB-TDL	TDL-CNB	CNB-ALD	ALD-MGS
Num. of trains	9	6	6	5	5
Average speed (kmph)	68	91	91	97	76
Overtaken by others (count)	0	0	0	0	0
Overtaking others (count)	6	9	8	6	7

Table 2.1: Some statistics for NDLS-MGS for train category: Rajdhani/Shatabdi/Garibrath
NDLS-GZB includes all terminals in the Delhi area.

Superfast/Express	NDLS-GZB	GZB-TDL	TDL-CNB	CNB-ALD	ALD-MGS
Num. of trains	49	29	40	30	31
Average speed (kmph)	39	64	69	70	47.4
Overtaken by others (count)	21	30	18	4	16
Overtaking others (count)	17	13	18	12	17

Table 2.2: Some statistics for NDLS-MGS for train category: Superfast and Express
NDLS-GZB includes all terminals in the Delhi area.

Passenger/Memu	NDLS-GZB	GZB-TDL	TDL-CNB	CNB-ALD	ALD-MGS
Num. of trains	16	2	1	1	2
Average speed (kmph)	24	46	37	38	37.5
Overtaken by others (count)	5	4	7	1	2
Overtaking others (count)	1	0	0	0	0

Table 2.3: Some statistics for NDLS-MGS for train category: Passenger and Memu
NDLS-GZB includes all terminals in the Delhi area.

Rajdhani/Shatabdi/Garibrath	MGS-GAYA	GAYA-DHN	DHN-ASN	ASN-BWN	BWN-HWH
Num. of trains	3	3	1	2	2
Average speed (kmph)	97	74.3	78	72.5	65.5
Overtaken by others (count)	0	0	0	0	0
Overtaking others (count)	3	1	0	3	2

Table 2.4: Some statistics for MGS-HWH for train category: Rajdhani/Shatabdi/Garibrath
BWN-HWH includes trains to Kolkata area

Superfast/Express	**MGS-GAYA**	**GAYA-DHN**	**DHN-ASN**	**ASN-BWN**	**BWN-HWH**
Num. of trains	18	23	12	28	49
Average speed (kmph)	62.7	55.4	57	61	57
Overtaken by others (count)	9	3	0	5	15
Overtaking others (count)	8	4	3	7	20

Table 2.5: Some statistics for MGS-HWH for train category: Superfast and Express
BWN-HWH includes trains to Kolkata area

Passenger/Memu	**MGS-GAYA**	**GAYA-DHN**	**DHN-ASN**	**ASN-BWN**	**BWN-HWH**
Num. of trains	8	8	11	16	9
Average speed (kmph)	42	41	37	53.2	39.5
Overtaken by others (count)	7	4	1	3	7
Overtaking others (count)	0	0	0	1	2

Table 2.6: Some statistics for MGS-HWH for train category: Passenger and Memu
BWN-HWH includes trains to Kolkata area

The above tables indicate how lower priority trains are overtaken more by higher-priority trains. Further, there are also some (though not many) scheduled overtakes between trains of the *same* priority. As elaborated later below, if a section is identified as a “bottleneck section”, it is crucial to eliminate as many overtakes as possible on that section: this helps improve overall throughput. The above statistics are computed from the data available on http://indiarailinfo.com it is to be noted that allowances (sometimes excessive amounts) have been included and this reflects in smaller average speeds. A more realistic picture of section-congestion is reflected in the quality of freight traffic paths and not only the data from passenger train schedules.

3. Capacity utilization of sections handling mixed traffic

Capacity is often specified as the ‘‘maximum number” of trains per day a section can accommodate, and a typical utilization or “throughput” is compared with the capacity to obtain “efficiency” of the timetabling process or implementation/operation. There are at least three different definitions of capacity and capacity utilization. We will pursue the first one in this report, but also describe three other notions in this section, in which the third one (i.e. 3C) is the one related to “Scott’s formula” and is familiar in the Indian Railways context.

3A: Mixed-traffic-ideal-grouping notion of capacity

With homogenous traffic, capacity is the number of trains that can pass through a section in a given unit of time: here per day. The reality in *mixed* traffic is much more complex, as the operation has to consider a specific number of trains of different characteristics and some specific (or at least typical) sequence of these trains on the given section. In such a situation, given sufficient traffic, if there is no overtaking, faster trains are forced to travel at lower than their maximum or rated speeds OR slower trains will be further slowed down as they decelerate, stop, get-overtaken, and then accelerate in order to give way to faster (usually higher priority) trains. In either case, there is a significant loss of capacity.

Earlier studies by Indian Railways have highlighted the impact of speed differential on section capacity. This report provides some initial basis for objective and meaningful computations of capacity in such situations. This has a significant impact on timetabling.

Total capacity utilization under the situation of <u>typical-grouping of trains</u>:
This capacity is calculated by considering the time loss due to overtakes (acceleration, deceleration) for a typical sequence of train running on the section.

Total capacity utilization under the situation of <u>ideal-grouping-of-trains</u>:
This capacity is calculated by considering the full grouping of trains (no overtakes) but with the desired mix (numbers) of each type of train.

Brief explanation of capacity calculation:

- Time to travel the bottleneck block (longest block) in a section is calculated.
- Actual time utilized is calculated by sum of travel time by trains, prior and later headway for each train and time for overtakes (acceleration time, deceleration time and time to cross the loop).
- Available time is 840 minutes after excluding 240 minutes (i.e. 4 hours) for maintenance and 70% efficiency from 1440 minutes (the number of minutes per day).
- Capacity utilization is the ratio of actual time utilized to available time.

A low value of capacity utilization here would suggest trying a proper grouping of trains, thus calling for more careful timetabling, rather than automatically justifying more investment in infrastructure.

The above definition is used later in this report. We describe three other notions briefly.

3B: Distance-time-chart-occupied notion of capacity utilization

A second way to define capacity is the fraction of area “occupied” by trains with respect to the total area in the distance-time chart. It is essential to include safety related headways into the notion of “occupied”. This definition allows a 100% capacity utilization by trains as long as they have the same running characteristics, independent of their running speed. One can introduce the 70% efficiency and maintenance block margins here too. Though this definition has its merits and has been studied in the literature, we do not pursue this definition in this report.
A low value of capacity utilization obtained by this definition would also suggest some leeway to be achieved through proper grouping of the trains in the mixed traffic scenario.

3C: Bottleneck section: slowest-train-based notion of capacity

A third notion of capacity would be the “throughput”: here the bottleneck section is identified and the number of trains with slowest speed is calculated and the actual number is compared with respect to this number. This definition has perhaps been used in the data from Indian Railways provided by NITI-Aayog. Conventionally, line capacity has been calculated in the Indian Railways context using the so-called “Scott’s formula” which involves the number of trains that can pass a section: the deciding factor being the slowest train and the bottleneck section. Headway constraint is also taken into account, and a factor of about 70% to 85% is included to ensure resilience. This definition of capacity is widely used in Indian Railways and several references in Indian Railways documents contain an elaboration: see [8], for example.

This notion suffers from the drawback that since quite a few trains run at speeds faster than the slowest train, we get an exaggerated capacity utilization percentage (often much higher than 100%); the obtained value of the capacity utilization is too sensitive to the fraction of the slowest trains in the entire day’s train composition. We note that this definition is meaningful to some extent, in homogeneous traffic conditions, such as suburban sections or freight only sections.

A low value of capacity as per this definition would call for splitting of the longest block and/or increase of the speed of the slowest train. Investment in infrastructure to reduce the traversal time over this bottleneck would help in capacity improvement. A low percentage capacity utilization would mean much heterogeneity and improper grouping of trains.

3D: Bottleneck section: fastest-train-based notion of capacity

One can consider the fastest train for the purpose of capacity calculation: we then have actual percent utilization (in the presence of mixed traffic) as lower than 100%. If the number of the fastest trains is much lower than the rest, then this definition would give a very low value of capacity utilization.

A low value of capacity here would mean that the bottleneck section needs more infrastructure upgradation and a low utilization percentage would mean much heterogeneity and improper grouping of trains.

It must be noted that “throughput”, as a quantitative measure of how much a given capacity is utilized, and “congestion” is a standard term across many areas involving flows over networks: air-traffic, internet data-traffic (see [7], [4] and references within for some examples). A *qualitative measure* of capacity utilization is as important too: this ensures that customers are satisfied by the Quality of Service (here, the customers are railway passengers and freight traffic beneficiaries.)

While a detailed investigation into the correct notion of capacity utilization is further needed, for the purpose of this report, we continue with the first notion of capacity elaborated above, i.e. the mixed-traffic, ideal-grouping based capacity calculation and utilization. We list some numerical values used for the calculation.

Headway Time (minutes)

The following numbers in minutes have been assumed to be safe operating headways between different categories of trains. The prior headway reflects the importance in maintaining punctuality of that category of train and the post headway reflects the requirement of the lower priority category of train to be able to follow the leading train at sufficient speed. We note that trains are typically scheduled in the priority reflected in table 3.1 and in a typical sequence, it is the prior headway for the first two categories and the post headway for the last two that are important, more often than not.

Train type	Prior headway	Post headway
Rajdhani (130 kmph)	5	1
Express (110 kmph)	3	2
Passenger (100 kmph)	1	3
Freight (60 kmph)	1	5

Table 3.1: Headway values in minutes for various types of trains

Time for overtakes (minutes)

Train type	Time required to cross a crossover point on the network	Time for acceleration	Time for deceleration
Rajdhani	2	2	1
Express	2	3	2
Passenger	2	4	2.5
Freight	2	10	5

Table 3.2: Time (in minutes) lost due to being-overtaken for various types of trains

Capacity utilization of some sections

Using the parameters above and some representative numbers for the traffic on various sections and typical sequences that are observed, we compute capacity utilization in the table below. Most of these refer to the important Mughalsarai - Allahabad section. We have considered 40 freight trains a day as an ambitious daily figure for the entire section, as per the data provided to us.

Section	Total number of passenger trains per day	Total number of freight trains per day	Total capacity utilization under typical condition (%)	Total capacity utilization under ideal condition (%)
Mughalsarai - Jeonathpur	43	40	94.58	66.10
Jeonathpur - Chunar	43	40	68.33	56.42

Chunar - Mirzapur	40	40	71.48	55.65
Mirzapur -Chheoki	43	40	81.82	71.42
Chheoki - Naini	43	40	82.40	68.12
Naini - Allahabad	52	40	104.30*	72.39
Allahabad - Bamrauli	39	40	115.23*	60.26

Table 3.3: Capacity utilization for various sections. Typical: typical grouping, Ideal: perfect grouping

The starred (*) numbers are not sustainable with typical sequences and headway times. Such traffic can be handled only in exceptional situations for short periods of time. The numbers in the extreme right column indicate that this is possible with resequencing in order to streamline traffic.

4. Analysis of how various parameters affect congestion

This section contains an analysis about various parameters that affect a timetable and the congestion caused due to primary and secondary delays.

Maintenance block considerations

1. As per current practices, in order to ensure safe running of trains, it is essential to have maintenance blocks of duration between 2-4 hours, contiguous in time (usually in day time hours) and for an entire section at a time. This has a significant impact on available capacity.
2. On congested sections, there is a conflict between the requirements of maintenance and traffic goals: an improper planning leads to congestion: though the congestion is local and temporary, it typically takes much time to resume to normal operation and recover the lost time.
3. After any substantial maintenance work, there are speed restriction on the track due to which train travel time increases (average speed decreases). This is part of the allowances built into the operating timetable on any section. Generally speaking, these are proportional to the length of track that has to be operated.
4. It is noteworthy that at many areas in the world, there are devices that are deployed to collect and monitor health of the track and other infrastructural components in a faster way and the data is analyzed in real-time at a railway-data-center for fault detection and diagnosis. While expertise for deployment and monitoring requires to be developed, this is the systematic and future way, and moreover, this is not large infrastructural investment but cheap and a viable tool. The <u>monitoring</u> aspects of maintenance blocks is avoided and capacity can thus be improved.

Slacks and allowances

In order to obtain good punctuality of railway operations, small disturbances and delays need to be accommodated during the operation: a timetable needs to absorb such durations of time. This is usually explicitly listed in a working timetable as ER/OR (recovery time values). Since the calculation sometimes tends to be ad hoc, and can have a very negative impact on both capacity and punctuality, we dwell on the calculation procedure and then on the consequence improper allowance in this subsection. The allowance can be calculated in two ways:

(1) by scheduling trains at speeds <u>lower</u> than technically achievable - we refer to this as *slack* and

(2) by keeping scheduled running times <u>longer</u> than the technically minimum running time - this difference is known as *allowance* or running time supplement.

We assume that slacks are planned with respect to driver behaviour and equipment performance in safe regimes.

For allowances, there are two steps which need to be followed to have punctuality in the operations of a timetable.

1. Determination of total allocation to be allocated to a train throughout its journey

Total allowance should neither be too high (as this leads to longer scheduled travel time, which in turn consumes the capacity of the system) nor should the allowance be too little (as this will not help achieve punctuality in the operation due to the presence of disturbances and daily exigencies). Generally, allowances are given in terms of percentage of the minimum travel time of the train. Internationally, the total allowance varies between 3% to about 10% of the running time, depending on the train and route characteristics. We find that this amount is 15% or more for trains in Indian Railways, which is on the higher side.

2. Optimal distribution of the total allowance to all stations throughout the journey

Once the total allowance to be allocated is decided, the next step is to distribute it among all stations in a rational way. One can allocate the allowance at the end of the section or alternatively distribute it at intermediate stations. There are merits and demerits to both as summarized below.

Though the chance that the allowance will be used effectively if all the slack is given to downstream stations is high, the train often gets delayed at earlier stations also and this causes knock-on (i.e. ripple effect) delay to other trains and thus affect the punctuality of the other trains. On the contrary, if more allowance is given to earlier stations and if there are no early disturbances, then the early allowances are lost and cannot be used. Since the delay reduction at a station not only reduces the delay of that particular train but it also reduces the knock on delays for subsequent trains, so instead of evaluating the delay at last station only, the sum of delay at each station should me minimized. Systematic methods need to be developed to rationally apportion slacks and allowances.

Grouping of trains based on speeds

Grouping of trains means to schedule trains of similar running characteristics in a bunch rather than interlace them. Grouping of trains leads to minimum overtakes and thus less traversal time. Grouping thus also leads to better capacity utilization, this means one can enable the running of more trains for a given infrastructure. Grouping however sometimes cannot be done to the desired level due to requirement of passenger convenience of timings of scheduled trains. Notwithstanding this requirement, the principle of grouping should be followed to the extent possible and should be a major consideration at the time of preparing timetables, especially for newer varieties of rolling stock.

IIT Bombay-based mixed-traffic simulator experiments to study effect of grouping

Theoretically we know that the grouping of train in section(s) can reduce congestion by reducing overtakes but this hypothesis can established more firmly by performing some experiments. We perform two types of simulation experiments:

- Effect of grouping on number of overtakes
- Effect of grouping on total traversal time for a given total of mixed traffic

<u>Effect on number of overtakes</u>

Consider two cases for grouping:

Case 1 (One 24 hour window):

Take full 24 hour time window and first schedule all 130 kmph trains with some headway time then schedule all 110 kmph trains and then same as above, schedule all 100/96 kmph trains and then schedule all 60/75 kmph trains. In this experiment we do not include freight trains.

Case 2 (Three windows each of about 8 hours duration):
- First take a 9 hour time window (00:00 - 9:00) and first schedule some (approx 35% of total) 130 kmph trains with some headway time then schedule some (approx 35% of total) 110 kmph trains and then same as above, schedule some (approx 50% of total) 100/96 kmph trains and then schedule some (approx 35% of total) 60/75 kmph trains.
- Same as above, take 9 hour time window (9:00 - 18:00) and first schedule some (approx 35% of total) 130 kmph trains with some headway time then schedule some (approx 35% of total) 110 kmph trains and then same as above, schedule some (approx 50% of total) 100/96 kmph trains and then schedule some (approx 35% of total) 60/75 kmph trains.
- Then take rest 6 hour time window (18:00 - 23:59) and first schedule rest (approx 30% of total) 130 kmph trains with some headway time then schedule rest (approx 30% of total) 110 kmph trains and then schedule rest (approx 30% of total) 60/75 kmph trains.
Note that it is required to keep a one-hour time gap at the end of each window.

Description of the experiment:

- Select any one direction (up or down) and take same number of trains as per Working Timetable (WTT) 2016.
- Reschedule the trains, this means changing their starting time according to case 1 or case 2 and make scheduled time table and unscheduled time table.
- Simulate the timetables of case 1 or case 2 using IITB simulator.
- Then calculate overtakes and average travel time of each type of train in both case 1 and 2 and compare with WTT 2016.

One can consider more cases, in addition to cases 1 and 2 and study the effect of grouping. This report contains comparison of Cases 1 and 2, and Working TimeTable (WTT) 2016.

Case 1 is the best possible case for grouping of trains according to speed of trains, in this case there are zero number of overtakes. However, this is not realistic case due to passenger convenience requirements.

We study the effect of the grouping on number of overtakes. We also include comparison with the WTT 2016. For this comparison with case 2, we consider 36 trains starting from GZB and proceeding towards MGS (for which we have the WTT 16). These trains are of 130 kmph, 110 kmph and 100/96 kmph. Conclusion of this experiment is that there were total 7 overtakes according to WTT 2016 and only 4 overtakes according to Case 2 by the three-window grouping. This means that by grouping of these 36 trains, according to case 2, the number of overtakes was reduced by 3: (i.e. reduced from 7 overtakes in the WTT 2016 to 4 overtakes as per grouping described in the three windows each of about 8-hour duration of Case 2).

Effect on total traversal time across all trains

For various possible groupings of a given breakup of heterogeneous trains, it is possible to generate timetables for each sequence of these trains. We analyze the effect of grouping on the total duration for sending the total of 40 trains (with breakup given below) from GZB to MGS. For all the 30 randomly generated timetables (each being a shuffle of the pattern containing the following composition), the start time of the first train from GZB is 00:00 midnight, but the arrival of the last train at MGS varies from 25 hours 14 minutes to 42 hours 41 minutes. This variation is an outcome of different groupings between these heterogeneous trains. In the following experiment, we checked various sequences of 7 Rajdhani trains (of speed 130 kmph), 6 express trains (of speed 110 kmph), 7 passenger trains (of speed 100 kmph) and 20 freight trains (with maximum permissible speed 50 kmph). “P” stands for Passenger trains, “E” for Express, “R” for Rajdhani and “F” for Freight train, and a pattern “PREEF” means P followed by R, followed by two E-trains, and then a F-train.

Sr. Num.	Best and worst patterns (amongst 30 randomly generated patterns)	Time for last train to reach final destination
1	PRFFRFRFFPEPEPFEFRFPFFERPFERFFFFRFFFPFEF (Relatively best pattern)	25hrs, 14min
2	FFEFFEEPFFRFPPFFPPFERRERFFFRFRFFFFEPFFRP (Median)	27hrs, 21min
3	EFFRPPFFPRFEFFFPFFPFREFFFFERPFFPREFERFFR (Relatively worst pattern)	42hrs, 41min

Table 4.1: Effect of train grouping on total traversal time

We conclude from Table 4.1 that the net traversal time can vary over a wide range. Given that the current simulator schedules trains based on both the *sequence* of trains in a proposed timetable and the *priorities* of the trains (with freight train being the least priority), a more thorough investigation is needed to associate various congestion parameters with the grouping/pattern. We consolidate some conclusions from the two types of experiments above:

A systematic grouping of patterns helps reduce overtakes, and also helps decrease net traversal time. In practice, grouping is governed by passenger requirements: grouping should be done as ideally as possible on the section that is identified as a bottleneck: improving performance (quantitative and qualitative) of the bottleneck helps the entire section performance.

5. Junction congestion analysis

The analysis of congestion at a junction is quite different than the analysis of a section. At a junction, various simultaneous or near-simultaneous movements are possible and resources have to be shared more carefully. There are currently no well-defined frameworks or planning tools to assess junction resources in a systematic manner. Some of our studies have been reported in [6, 9, 10] and this still remains a topic of research worldwide.

In order to quantify how much a entry/exit layout at a junction can affect the "throughput" of a junction: one can use a so-called resource-to-resource hindrance analysis. This method is not as applicable to a *section* analysis since in a section, near simultaneous movements are not allowed (except in opposite directions, or during an overtake).

The case of Allahabad junction (ALD) with 19 lines (both up and down) and 10 platforms is a good example of how hindrances on a train movement due to movements of other trains can use-up a lot of time and resource. This is shown in Figure 5.1. When a train arrives on a line, it not only causes hindrances to the platform or line on which it is halting/passing but also to other lines as the lines need to share various common linkages during the process of entering/exiting the station. A hindrance matrix is constructed to tabulate these hindrances. (Details of the approach can be found in [5, 6].) Analysis of the hindrance matrix along with the timetable of passenger trains at Allahabad junction gives the output graph below. It represents the availability, occupancy and hindrance for every loop line.

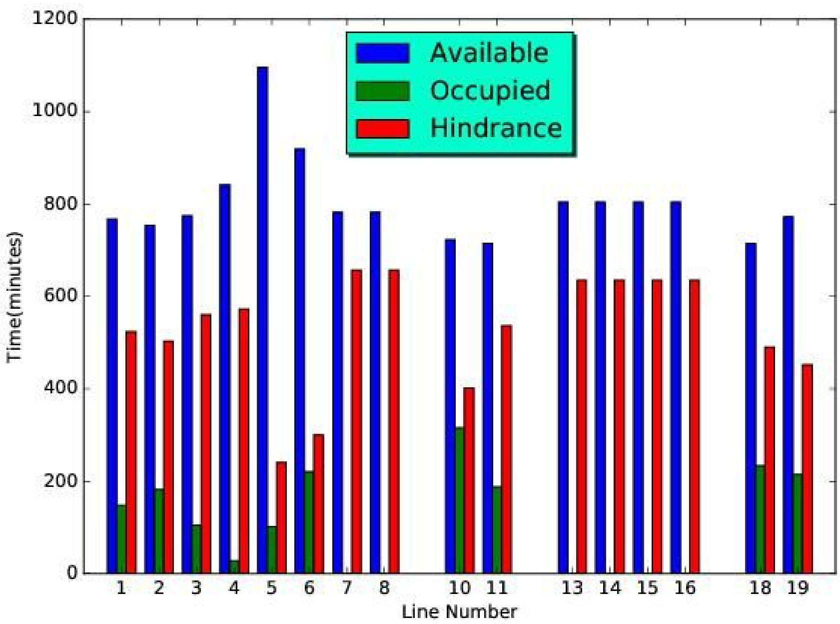

Figure 5.1: Time lost due to hindrances caused due to cross-movements in a junction area: ALD

Allocation of platforms to passenger trains in an optimal way can also help in decreasing the hindrances. One can allocate platforms to passenger trains (at least to some trains) in such a way that the train's entry/exit-movements produce very little hindrance to the other trains entering/exiting at that time. In this report, we have allocated platforms at ALD to most of the trains according to the data mentioned on http://indiarailinfo.com.

A junction entry/exit line layout can be termed perfect if for each platform, we have either availability (blue) or the platform is occupied (green): the hindrance amounts are ideally zero. The hindrance amount indicates the extent of non-availability of a platform due to a cross-movement of another train. For example, line-5 of ALD (see Figure 5.1) is best from a layout viewpoint since the hindrance (15%) into this line is least caused due to other line entry/exit movements. On the other hand, lines 3, 4, 7 and 8 are unavailable for 42% of the day just due to hindrances caused from other line entry/exit movements.

From the figure, one can also deduce that for each platform about 15% - 42% of a day is neither available for occupation nor is the platform occupied; this is solely due to the hindrances created by passenger trains moving in or out of ***other*** platforms. Note that low hindrance on Loop line 6 (about 17%) is only because it is a “dock line” from Naini, Allahabad City and Prayag and it does not experience any hindrance in the UP direction (towards Subedarganj). Loop lines 9, 12 and 17 are not shown as they are Engine/Stabling Lines.

FREIGHT TRAIN ANALYSIS AT ALLAHABAD JUNCTION

After simulating all the passenger trains at ALD junction, we have obtained the time for which the lines are occupied or free or hindered. So one can now simulate freight trains through the junction. We obtained the different amount of delays observed by the freight train at different time duration for the whole day.

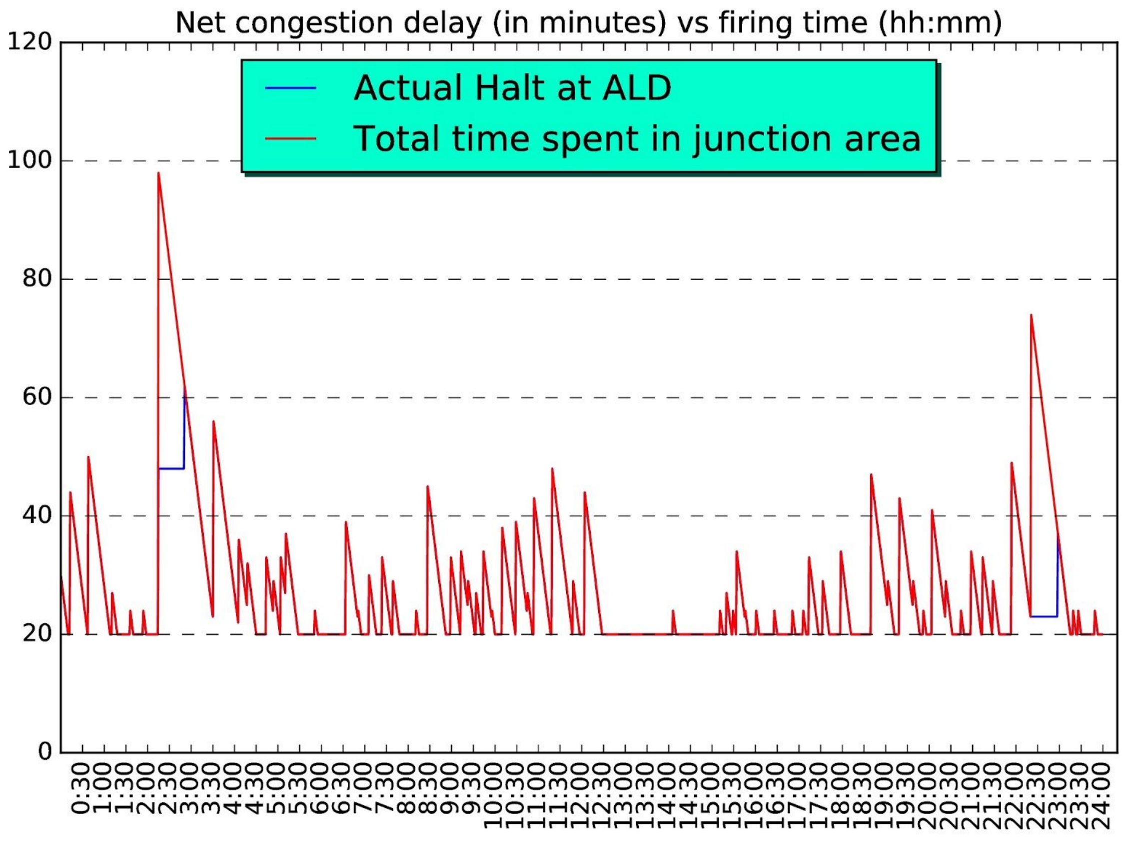

Figure 5.2: Delay experienced by freight trains while entering a junction at various times of the day

It is clear from the graph that freight trains coming between 12:30 to 14:30 experience least amount of hindrance due to the movements of other passenger trains. Also, halts less than 40 minutes are justifiable in junction areas keeping in mind their infrastructural complexities. We propose the following for freight trains at a junction.

1. Keep a variable freight halt so that one can get a more realistic picture of the freight trains waiting/movements at large junctions.
2. Backtrack the freight trains and start them from their source station such that they experience small amounts of hindrance due to other train movements while entering/exiting the junction.

An analysis of a junction using this approach, followed by a routine sensitivity analysis, can help identify the bottleneck links at the entry/exit of a junction: it would be incorrect to term the shortage of platforms as the bottleneck since, as is visible from the above plot, most platforms are occupied by less than 25% of the day.

A careful planning of short linking lines (to cause small hindrance to cross-movements) combined with a systematic platform allocation can reduce the hindrances in the hindrance matrix. This approach has been demonstrated above for the ALD junction and has also been pursued for Kanpur area in [5]. This approach is a possible line of analysis for understanding a complex operational issue: this issue is acknowledged across the world as complex and is a matter of ongoing research.

6. Recommendations

The recommendations below arise from a combination of tools and analyses performed from the data that was used in this study. Each generic recommendation is followed by a specific recommendation on the NDLS-HWH section to the extent possible, from the limited study that has been done here. This area deserves a much more extensive and continuously updated efforts on the part of Indian Railways, as conditions are bound to be evolve in the years to come.

Capacity-utilization: **General recommendation :** Capacity utilization measures are used as an indicator of congestion and high values are used to justify additional investments and possibly as a justification to provide for more traffic allowances in timetabling. While the general principle is acceptable, the capacity measure itself needs significant refinement and consensus in use before any valid claim can be made in mixed traffic sections, such as the Howrah - Delhi main line section on Indian Railways.

We suggest a simple extension that is formula based, to start with, but which considers both the mix of traffic on a section and to some extent, the sequence of traffic on a section. Further, since congestion is a direct consequence of both:

- timetabling (grouping, allowances and other scheduling aspects)
- inadequacy of infrastructure (additional main/loop lines, signalling upgradation needs, etc),

decisions about additional investments must be backed by a simulation and comparison of the simulated timetable before and after adding infrastructure: IIT Bombay's mixed-rail traffic simulator is one of the available tools, another tool, for example, is Satsang by CRIS.

Specific recommendation: The sections just short of Allahabad Junction, in both directions are crucial to improve the overall performance. The down direction has three lines from Subedarganj. The up direction needs streamlining and an increase in the number of freight train paths that avoid halts at the already congested Allahabad junction area. If the crew change points can be shifted west of Allahabad, some of the congestion can be eased.

Bottleneck-scheduling: **General recommendation:** A principle in bottleneck scheduling in multi-resource environments is that the maximal throughput strategy on the bottleneck resource should drive the schedule on the rest of the network. It is widely felt that the Mughalsarai-Allahabad section on the Howrah-Delhi section of Indian Railways is a potential bottleneck section. If this is so, the traffic on this section should be streamlined, with as few overtakes as possible, and also as ideal grouping as possible, so as to achieve maximum throughput and overall traversal time performance.

Specific recommendation: From an analysis of speeds, it appears that the ALD-MGS section is significantly more congested than the other major sections (reflected in lower speeds of passenger trains and also indirect evidence of average freight train speeds on this section). This is one indication of a bottleneck resource. This needs careful validation and if verified, it needs focused attention in the following ways (a) prioritizing investments on this section - the completion of automatic signaling on this important section was achieved only recently (b) streamlining movements by avoiding overtakes and slower movements on this section and c) placing of pinpointed facilities such as additional loop lines - simulation suggests that an additional one or two loop lines in Naini Junction in the UP direction would be useful. Where space and other resources permit, a third line for one or two stations short of major junctions would be useful and twin single line working can also be considered to debottleneck locally.

Junction-analysis: **General recommendation** The impact of junction movements is very significant and causes significant cascading impacts on sectional running as well. This area of analysis in railway operations is a very challenging one worldwide and proper tools and techniques need to be developed to analyse this. We have suggested some beginnings in this note.

Specific recommendation: It is hypothesised that Allahabad junction is the major junction causing delays and measures to debottleneck this area are imperative. Some have already been done and more needs to be done in the period till the Dedicated Freight Corridor infrastructure (which will bypasses Allahabad altogether) comes up for freight movement.

Slack-and-allowance-distribution: **General recommendation:** Divisional measures of punctuality should be reworked to have a more continuous unit of measurement (rather than a slab based measure) and should be based on the resource available in each part. The current practice of loading all allowances at the end of a section before interchange is detrimental to punctuality of operation and should be re-evaluated.

As elaborated in the previous sections, while some allowance is reasonable to ensure punctuality, excessive allowance is, in fact, detrimental to punctuality and causes more congestion: both on sections and at junctions. We summarize the effects here.

- When allowances are excessive, some trains arrive at a station or just short of a station, before the scheduled time. The waiting of these trains consumes resources - and has adverse implications. Main line waiting prevents other trains which are on time and/or prevents freight trains from moving. Loop line waiting decreases fluidity in the system and constrains freight movements by restricting overtake options.
- On the other hand, stopping just short of junction areas can be even more detrimental as multiple cross movements typically affected.

It is recommended that only reasonable amounts of allowance is provided for section-congestion, and the provision be focussed on easing section-congestion and not on easing junction-congestion. For guarding against junction congestion, given the occupancy pattern of platforms at junctions, it appears that additional halt times at junctions is a viable option. The option of cutting those durations short when needed is available.

Specific recommendation: Allowances are skewed towards the Howrah end of the Howrah New Delhi section and should be more rationally apportioned. An initial analysis suggests that the allowances for Rajdhani trains are of the order of 90 minutes over the 787 km from Delhi - Mughalsarai and of the order of 114 minutes for the 664 section from Mughalsarai to Howrah. The corresponding figures for Express trains are 248 minutes and 237 minutes. This is anomalous and should be corrected by reducing this allowance in the Mughalsarai Howrah section for almost all trains. A possible reason is that there are 2 divisions that operate trains between Delhi and Mughalsarai and 4 divisions after, and there is a provision of allowances at the interface of all divisions (in addition to more operationally logical ones before major junctions and terminals). This should be corrected.

7. Conclusion

Timetables serve as performance targets for train operation and are constructed carefully and with much planning. With the increasing size and number of planning units of Indian Railways, it is essential that this task is done in a well thought-out and planned manner, with as much data-based input and evidence-based decision making as possible, rather than only the subjective inputs of experienced timetablers. For example, the platform provided by Satsang, the timetabling tool of CRIS (Center for Railway Information Systems) is a sound basis for timetabling and should be used more effectively. It appears that this is not currently being done for reasons that are unclear.

Worldwide, railway organizations make use of ongoing developments in planning technology ranging from optimization and operations research to machine learning and data analysis to improve operations. This is true for main line operations, including freight services, as well as niche but important areas such as suburban rail services and metro services. Indian Railways should be the leader in this area, given the complexity and volume of services that it offers its customers and the role it has to play in the socio-economic development of this part of the world.

A landmark study to do with timetabling and railway operations as a whole is the complete revamp of the Netherlands Railway timetable (including sectional timings, terminals, rolling stock and crew) [2]. This significant plan was brought about by many years of cooperative work between railway personnel, some of whom had university appointments, academics and software professionals. The core of an ecosystem and skills for such a concerted effort for the admittedly much larger railway system in India exists and should be cultivated, as the impact can be significant.

References

[1] S. Dutta, N. Rangaraj, M.N. Belur, S. Dangayach and K.N. Singh, Construction of periodic timetables on a suburban rail network-case study from Mumbai, *Proceedings of the International Conference on Railway Transport Technology,* Lille, France, April 4-7, 2017.

[2] L. Kroon, D. Huisman, E. Abbink, P.-J. Fioole, M. Fischetti, G. Maroti, A. Schrijver, A. Steenbeek, R. Ybema, The new Dutch timetable: the OR revolution, *Interfaces*, vol. 39, no.1, pages 6-17, 2009.

[3] A. Landex, and A.H. Kaas, Planning the most suitable travel speed for high frequency railway lines, *Proceedings of the 1st International Seminar on Railway Operations Modelling and Analysis*, Eds. I.A. Hansen, F.M. Dekking, R.M.P. Goverde, B. Hindergott, L.E. Meester, The Netherlands, 2005.

[4] S. Low, F. Paganini, and J. Doyle, "Internet congestion control," *IEEE Control Systems Magazine*, vol. 22, no. 1, pages 28-43, 2004.

[5] D. Miglani, *Simulation and Analysis of Mixed Traffic in Railway Junctions,* M.Tech and B.Tech. dissertation, Department of Mechanical Engineering, IIT Bombay, 2017.
[Available upon request by email to Prof. Narayan Rangaraj or Prof. Madhu Belur, IIT Bombay]

[6] D. Miglani, M.N. Belur and N. Rangaraj, Railway junction simulation and analysis for mixed rail traffic, *Proceedings of the Symposium on Advanced Train Control and Safety Systems for Indian Railways (ATCSSIR)*, Kharagpur, May 26-27, 2017.

[7] A.R. Odoni, The flow management problem in air traffic control, *Flow Control of Congested Networks,* Springer, pages 269-288, 1987,

[8] Operations Document, Indian Railways, National Academy of Indian Railways, (Last accessed: Oct 2017) www.nair.indianrailways.gov.in/uploads/files/1430369423822-operating.pdf

[9] N. Rangaraj and B. N. Vishnu, Node capacity and terminal management on Indian Railways, In: *Vision 2025,* Vadodara Railway Staff College, India, 2002.

[10] S. Salsingikar, N. Rangaraj and L. Sathishkumar, Analysis and planning of train movements at a railway junction, *Proceedings of the International Conference on Railway Transport Technology,* Lille, France, April 4-7, 2017.

[11] UIC leaflet 406, Capacity, UIC International Union of Railways, France, 2004.

[12] R. Vidyadhar, S. Dutta, N. Rangaraj and M.N. Belur, Simulation and analysis of mixed traffic on railway sections, *Proceedings of the Symposium on Advanced Train Control and Safety Systems for Indian Railways (ATCSSIR)*, Kharagpur, May 26-27, 2017.

Data sources, performance measures and techniques used in this study

The study provided summary statistics of passenger train operations on different sections of Indian Railways. Section-wise details were provided on 42 sections on the Howrah - New Delhi route (distance, number of passenger trains, number of freight trains and section utilization as a percentage of available capacity). These numbers need to be verified with actual data on several important sections, but more importantly, it is imperative to update the framework for capacity computations using more relevant and contemporary measures of capacity. To our knowledge, Indian Railways has the largest extent of heterogeneous traffic that any rail network carries and it needs to evolve its own measures of performance to suit its needs.

In addition to the summary data provided by Niti Aayog, we have used data from Indian Railways, especially North Central Railways, and public sources of data such as http://indiarailinfo.com

Some of the analysis is through a rail traffic simulator that has been developed over a number of years at IIT Bombay - this has been used in other studies and analysis in the past [12]. The junction analysis is through a stand alone computational framework that has been developed by students and faculty at IIT Bombay [5, 6, 10]. Extending this to a full fledged software tool is a challenging task and is ongoing work. Feedback and suggestions about more features in the simulation package are welcome.