A Heuristic for Railway Crew Scheduling with Connectivity of Schedules

Akshat Bansal, Anoop K P* and Narayan Rangaraj
Department of Industrial Engineering and Operations Research
Indian Institute of Technology Bombay, Mumbai, Maharashtra, India
akshat.bansal.c2022@iitbombay.org, kpanoop05@gmail.com, narayan.rangaraj@iitb.ac.in
* Corresponding author

Abstract: This paper addresses the crew scheduling for long-distance passenger train services. A heuristic with bin packing features is developed to generate repeatable crew schedules that satisfy the operational and crew allocation rules. By ensuring the connectivity of crew duties that can be repeated over periodic train schedules, a better estimate of the crew requirement in a region is also obtained. Further, the heuristic ensures a fair division of total workload and creates long duty cycles which also makes the process of cyclic rostering easier. The paper also presents an exact approach for crew scheduling using a combination of constraint programming and set covering formulations. The exact approach is not computationally viable for practical scale problem instances, but the heuristic generates good quality solutions (often very close to optimal) even on large data sets. We illustrate the approach on data from the Mumbai Division in Indian Railways and the computational results show that there is potential to reduce the total number of crew duties in the region by around 12%. The heuristic approach provides an efficient way to generate improved crew schedules every time there is a change in the train timetable.

Keywords: Railway crew scheduling; Indian Railways; Constraint programming; Bin packing; Heuristic algorithm

1 Introduction

Crew scheduling is an important activity in all transportation systems and shares many similarities. However, each transportation mode has its own unique features and challenges associated with it. In rail transportation, the considerations in crew scheduling vary with the type of service being operated. Train services are broadly classified into passenger and freight services. Passenger services are operated based on a pre-determined timetable, while freight services are a combination of scheduled and ad-hoc services. Passenger train services can further be divided into three major categories: long-distance, suburban, and urban services, depending on their operational territory. This work considers the crew planning for long-distance passenger train services.

A generic planning process for passenger train operations involves multiple steps as shown in fig.1 [1]. It starts with line planning that determines various lines, their origin/destination/in-between stations, and types and frequencies of trains on each line to satisfy all the travel demands. Afterwards, train timetabling is carried out by fixing each train's arrival and departure times at each station in the section, ensuring all appropriate safety constraints. The next step is train platforming, i.e., assigning the platforms to the trains at the stations they halt. The following task is rolling stock scheduling wherein rolling stock units (railway vehicles) are assigned to scheduled trains with a predefined timetable. Often, regular maintenance activities are integrated into this planning step. Further, during the night and no rush hour, when the trains are not in use or are in maintenance, they need to be parked in a shunting area near one of the stations or depots, and it is known as a train unit shunting problem. The following two planning tasks concern the crew members, i.e., the crew scheduling and rostering. Crew scheduling consists of creating anonymous duties covering all the trains for a defined period based on a given timetable. Each duty specifies a sequence of tasks (or trips) satisfying operational constraints and labour union rules. Finally, the duties are combined for a larger time and assigned to individual crew members, known as crew rostering. A fair distribution of work, tracks and rolling stock knowledge, vacations, etc., are considered while assigning duties to crew members.

Figure 1: Planning process for passenger train operations

From a hierarchical perspective, railway crew planning problems can be classified as follows [2]:

- Strategic level planning: Decisions such as number of crew regions required, locations of base stations and crew change stations, allocation of duties among regions etc. It addresses system-wide issues to ensure a smooth operation of the system.
- Tactical level planning: Determines the crew capacity required to operate a set of train schedules that have been assigned to a region. The territory of decision-making is generally confined to a region.
- Operational level planning: Manages daily operations with a short planning horizon of usually a week. Crew scheduling and crew rostering are done at this level and adequate adjustments are made locally from time to time to ensure an unhindered operation.

Railway crew scheduling as an operational-level problem has been studied widely over the years. However, as pointed out by Suyabatmaz & Şahin [3], this may not completely help the tactical level planners to assess the exact crew capacity required in a region. This is because the recurrence of crew schedules and rosters over periodic train schedules are overlooked during this level of planning. Common practice is to follow a hierarchical top-to-bottom approach wherein the crew capacities at various regions are decided first, followed by the generation of crew schedules in accordance with it. Adequate adjustments are made at the operational level to ensure the repeatability of the crew schedules/roasters prepared. However, it is more desirable to reflect the concerns at the lower level planning to a higher level since this can lead to better overall utilization of resources. In this paper, we try to determine the regional crew capacity required to operate a given set of passenger train schedules by generating crew schedules which can be repeated over periodic train schedules (connected crew schedules) without any conflict. This makes the problem more interesting and also bridges the gap between tactical-level crew capacity planning and operational-level crew scheduling.

We consider the long-distance passenger train services operated by the Mumbai Division of Indian Railways and incorporate the crew allocation rules and guidelines applicable in the context. We develop a heuristic by extending the idea of the classic bin packing problem to generate a set of repeatable crew schedules for a given passenger train timetable. The long crew duty cycles generated using these schedules also ensure a fair division of work over a period of time since each crew member performs all the train tasks over a duty cycle (or multiple crew duty cycles). We also model the crew scheduling problem using a combination of constraint programming and set covering formulations. The models can be solved in sequence to obtain optimum solutions for small-sized problem instances. This provides a means for validating the crew schedules generated by the heuristic and for obtaining valid lower bounds for larger problem instances. We also compare the heuristic results with the existing crew schedules used by the Mumbai Division generated via a manual and laborious exercise involving multiple stakeholders.

While there are numerous models and approaches reported in the literature for railway crew scheduling, we identified that the existing models and approaches are inadequate to address the complex settings associated with large railway operators which are predominantly passenger carriers, like Indian Railways (Indian Railways operates more than 13,000 passenger trains daily over a route length of more than 68,000 km). The complexities include additional constraints related to home and outstation crew duties, differential rest times, a limit on the time a crew can stay away from the home base and heterogeneity in: (i) train characteristics and (ii) the number of trains operated on different days in a week. A solver-based exact solution approach is not expected to be computationally feasible since the model might have to be solved for a period exceeding a week (incorporating all the trains operated by a region on multiple routes) to ensure connectivity between crew schedules. We develop an efficient heuristic-based approach which can be implemented at a regional level where a single resource pool of crew is to be utilized for multiple routes or sections in an effective manner. The heuristic is scalable (some regions can have more trains to be operated on more routes) and is found to provide solutions which are within an optimality gap of 2.5% for the Mumbai Division data considered in the study. Moreover, the approach can help the regional-level crew planners to generate better crew schedules in a quick time. This can result in significant savings in crew expenses and the man hours expended to generate a new crew schedule every time there is a change in the train timetable.

The paper is organized as follows: Section 2 presents a review of the literature on railway crew planning. Section 3 describes the problem and explains the crew scheduling in Indian Railways. The solution approaches and the model formulations are presented in Section 4. Section 5 reports and discusses the results and findings from the computational study. Section 6 concludes the work with the scope for future research.

2 Literature Review

The origins of the crew scheduling problem in the transportation industry date back to the 1950s and 1960s. Arabeyre et al. [4] survey the different approaches adopted by airlines to optimize the allocation of crews to flights. The area gained greater momentum with the advances in computational power in the 1980s. Wren [5] examines various methodologies that have been applied to vehicle and crew scheduling for the bus transit system. In the 1990s, the railway industry came to the forefront of crew scheduling research realizing the potential savings that could be achieved by the application of various operations research techniques. Further, the deregulation and privatization of the railway industry in Europe also forced the rail operators to achieve cost-efficient utilization of various resources including the crew. The article by Caprara et al. [6] outlines different ways of modelling the railway crew scheduling and rostering problems and the possible solution methods. The research interest in railway crew scheduling has sustained over the last two decades. The recent article by Heil et al. [1] provides a review of railway crew scheduling literature published since 2000.

Crew allocation is subject to various rules and guidelines governing the maximum duty times and mandatory rest periods. The majority of these rules are country or region specific and are to be addressed at a local level. For instance, Abbink et al. [7] consider the crew scheduling problem associated with the Dutch railway operator and Neufeld et al. [8] consider the case of German railway operator DB Regio AG. Further, Khosravi et al. [9] and Khosravi & Tamannaei [10] consider crew scheduling in the Iranian railway network. The work by Kasliwal et al. [11] addresses the crew planning for commuter rail operations considering the case of the Mumbai Suburban railway network in India. The present work is the first to consider crew scheduling for long-distance passenger train services in Indian Railways to the full extent.

While generating connected and repeatable crew schedules for longer time periods, it is important to ensure fair distribution of work across duties. For this, an appropriate mix of train tasks of different duration and working conditions is to be maintained. The article by van Rossuma et al. [12] formulates an operational railway crew scheduling problem with Sharing-Sweet-and-Sour rules, a framework introduced to ensure a fair allocation of work by mixing appropriately the train tasks designated as sweet (attractive tasks) and sour (less attractive tasks). The formulation is applied at an individual level and further, a sequential solution approach is proposed to solve the problem considering the case of Netherlands Railways. The present work also ensures a fair distribution of the total workload by incorporating selective addition of tasks while generating duties. Further, the longer crew duty cycles generated by combining multiple crew duties ensure that each crew member undertakes every train task in the long run.

From a methodological perspective, we find a wide application of column-generation, meta-heuristics and heuristics-based approaches for solving crew scheduling problems of practical size. This is because the standard railway crew scheduling problem is proven to be NP-hard and generating exact solutions using approaches like integer programming is found to be a challenge. Exact solution approaches are thus limited to small or sparse problem instances. Jütte et al. [13] develop a crew-scheduling software based on a column-generation solution technique at DB Schenker. Jütte & Thonemann [14] present a column-generation based decomposition algorithm and tests it on instances of a major European freight railway carrier. Hanafi & Kozan [15] apply a hybrid constructive heuristic with simulated annealing to minimize the number of crew duties while reducing idle transition times. Liu et al. [16] develop a genetic algorithm-based column-generation heuristic to minimize the total cost of payment to passenger rail crew members in the North American context. Further, Nishi et al. [17] propose dual inequalities for the Dantzig-Wolfe decomposition of railway crew scheduling problems to reduce the number of replications in the column-generation procedure and improve its convergence. Hoffmann et al. [18] develop a prototype for a multi-period railway crew scheduling problem with attendance rates and apply a hybrid column-generation approach, which solves the pricing problem by means of a genetic algorithm. Also, Hoffmann & Buscher [19] present an arc flow model for the railway crew scheduling problem with attendance rates. The authors define various inequalities and perform computational tests to estimate the influence of different valid inequalities on computation times and bounds of the linear relaxation. An evolving methodological approach based on machine learning is also to be reported as seen in the recent paper by Gattermann-Itschert et al. [20]. The authors use machine learning to learn and predict planners' preferences in crew scheduling. A random forest classifier is trained on planner feedback to predict the probability that a duty is perceived as bad by the planners.

The article by Han & Li [21] addresses the crew scheduling problem for a Taipei mass rapid transit system. The authors propose a constraint programming-based approach for duty generation and a set covering problem formulation for duty optimization. The objective is to minimize the total number of duties subject to constraints related to crew location, continuous driving time, rest and meal hours in a shift. Also, the recent work by Chen et al. [22] adopt a similar approach combining constraint programming and integer programming (set covering formulation) methodologies. The paper reports the application of the proposed approach to the Kaohsiung depot of the Taiwan Railways Administration. In addition to the heuristic developed, the present work also adopts a similar combination of approaches to model the underlying crew scheduling problem. However, the problem definition and setting have been modified in the context of long-distance passenger services operated by Indian Railways. Further, the paper by Tapkan et al. [23] proposes a constraint programming-based column generation approach for crew scheduling considering the case of Kayseri Railway. However, the problem setting corresponds to a light rail system which is operated within a smaller territory with no complexities related to multiple outstation relief points and rests.

The idea of the bin packing problem has been applied previously in crew scheduling problems. Feng & Rui-hua [24] address the case of China's urban rail transit system and treat crew scheduling as a one-dimensional packing problem. The objective is to find the minimum number of drivers required, and constraints related to continuous driving time, total driving time in a day and rest time are considered. The paper proposes a best-fit algorithm to solve the problem. Further, Qiao et al. [25] apply the bin-packing algorithm to a vehicle and crew scheduling problem for the bus transit system. The objective is to minimize the total costs of runs while covering all the tasks and considering the constraints related to crew location, working hours and meal breaks in a run.

The articles by Şahin & Yüceoğlu [2] and Suyabatmaz & Şahin [3] investigate tactical level crew capacity planning problem in railways. The first article determines the number of crew required in a region to operate a set of train duties satisfying the day-off requirement of the crew. The second article minimizes the number of crew required in a region while both feasibility and connectivity of crew schedules are maintained. However, these works do not consider the complex constraints that exist at the operational level crew scheduling like differential rest periods, a limit on the time a crew can stay away from his home base etc. The present work bridges the gap between operational level crew scheduling problem and tactical level crew capacity planning problem by generating crew schedules that satisfy all the operational rules and which can be repeated over periodic train schedules. This helps in providing an estimate of the actual crew capacity required in a region. In the literature, integrated crew scheduling and rostering has also been attempted. Ernst et al. [26] propose an integrated optimization model to solve crew scheduling and crew rostering problem. However, the authors solve only a relaxed version of the optimization model. Lin & Tsai [27] propose a formulation that integrates crew scheduling and rostering problems and develop a branch-and-price-and-cut algorithm and a depth-first search-based algorithm to solve the composite problem. Further, Feng et al. [28] consider a crew scheduling and rostering problem for metro operations based on the duty path. The paper proposes a model that can solve the crew scheduling and rostering problems successively while considering duty types and practical requirements such as maximum on-duty time constraints, maximum continuous working time constraints, and meal constraints.

In the present work, we consider all the crew allocation rules and guidelines that are applicable in the context of Indian Railways to generate crew schedules that can be repeated over time without any conflict. Existing standard models and methodologies for railway crew scheduling do not address all these considerations and cannot be applied directly for crew scheduling in Indian Railways. This is primarily because the crew allocation rules vary from one country to another. In Indian Railways, there is a differential rest rule for outstation and return journeys, rest rules that depend upon the duration of the trips and a limit on the maximum time a crew can stay away from their home base. Further, in the case of long-distance passenger train services, some tasks span multiple days and the crew may not be able to return to the home base on the same day. Also, there is heterogeneity in the number of long-distance trains being operated on different days of the week, i.e., some trains run weekly, some twice a week, etc. As a result, the problem is to be solved for a longer duration (at least a week) to generate a crew schedule that can be used repeatedly. Currently, a panel consisting of railway officers from the operations department and representatives of the labour union for loco pilots and guards prepare the crew schedules for each route manually. The developed heuristic gives a realistic estimate of the crew requirement in a region by generating repeatable crew schedules and also provides a way to quickly update the crew schedules whenever there is a change in the corresponding train schedules.

3 Problem Description

The rules and regulations governing crew scheduling differ from one country to another. In this section, we describe the rules and guidelines for crew scheduling applicable in the context of Indian Railways.

3.1 Crew Scheduling in Indian Railways

Indian Railways is a statutory body that comes under the ownership of the Ministry of Railways, Government of India. The Indian Railways network is one of the most extensive rail networks in the world with a total route length of more than 68,000 km and more than 7,300 stations [29]. To manage and coordinate the operations, the entire network is divided into 17 zones. These zones are further subdivided into a total of 71 smaller units called divisions for administrative purposes.

In Indian Railways, each crew member is affiliated with a division and is assigned to one of its crew bases (also known as the home base of the crew). Typically, crew bases are the important terminals or stations within a division. In the case of long-distance services, the entire train journey needs to be divided into smaller tasks and are to be taken up by a series of crew members. The stations where the crew change happens are called crew change points (CCP), which are generally the stations with a high frequency of train services or stations at the border of a division's jurisdiction. All CCPs are equipped with basic facilities like resting rooms for the crew members. Note that every crew base is a CCP but not vice versa. Further, the territory within which a crew (stationed at a crew base) is authorized to take up the duty (usually, this will be a section between a pair of CCPs) is known as a crew beat. Lastly, a trip in which a crew member is transported as a passenger to the home base or another CCP is known as a deadheading trip.

At present, in Indian Railways, a committee of experienced railway staff manually prepares the crew duties. They prepare a so-called 'detail book' based on the given train timetable. The book lists the group of tasks to be operated by the assigned crew member on a day and is known as a detail (as shown in Appendix A). Further, these details are arranged in a specific sequence to construct longer duty cycles so that after operating a particular detail, the crew operates the following detail in the cycle. Factors like the type of trains - daily or non-daily, slow or fast, etc., are considered while preparing the detail. Although the cyclic nature of the plan leads to a good balance of the total workload in the long term as everyone operates every task, this way of preparing the duties is a timely and labour-intensive process which can easily lead to solutions which are far from optimum. Moreover, this exercise is to be repeated every time there is an update in the train timetable.

3.2 Case Study: Mumbai Division, Central Railway

We consider the case of the Mumbai Division, which comes under the central zone of Indian Railways. The division manages a total route length of more than 575 km. It is important to note that all the major terminals and stations in the Mumbai region (like CSMT, LTT, etc.) act as a single CCP and have been referred to as 'Mumbai' in this paper. Additionally, Mumbai is the only crew base in this division, i.e., all the crew members are based out of here. The crew belonging to the Mumbai division operates long-distance passenger trains in the following three routes: Mumbai - Igatpuri, Mumbai - Pune and Mumbai - Ratnagiri. However, since Mumbai - Ratnagiri is a relatively longer section, crew tasks for low-speed trains are fixed up to Roha (which comes ahead of Ratnagiri), to ensure that no task exceeds the maximum running time limit. This results in a total of four sections and the resulting network is a star-type network as shown in fig.2. The distance and traversal times for all the sections are shown on their respective arcs.
The distribution of train tasks over these sections and different days of the week are summarized in table 1. Here, the UP direction refers journey towards Mumbai, and the DOWN direction refers journey away from

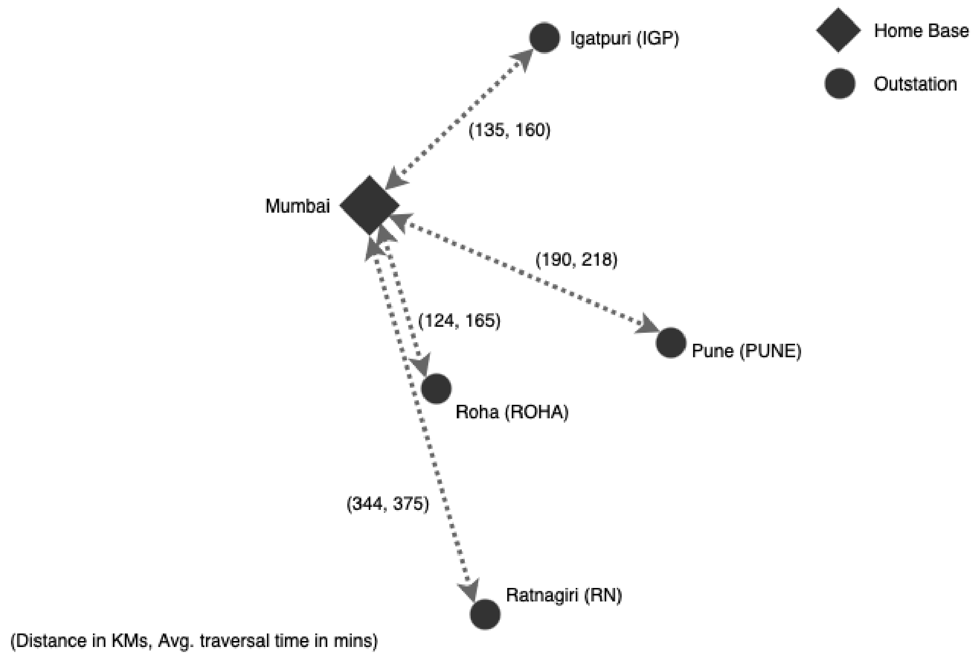


Figure 2: Mumbai Division: Route map

Mumbai. Further, IGP is the station code for Igatpuri, PUNE for Pune, ROHA for Roha and RN for Ratnagiri. In this work, we also compare the impact of the following two scenarios on crew utilization: section-wise planning, where crew schedules are generated for a single section at a time and integrated planning, where crew schedules are prepared for the entire Mumbai division with the crew being allotted tasks in any of the sections in the division.

Table 1: Distribution of train tasks over a week

Sections	**Direction**	**M**	**Tu**	**W**	**Th**	**F**	**Sa**	**Su**	**Total tasks in a week**
Mumbai - IGP	UP	39	38	40	39	37	38	38	269
	DOWN	39	36	42	36	37	42	37	269
Mumbai - PUNE	UP	26	25	27	26	25	27	26	182
	DOWN	25	26	27	26	25	27	26	182
Mumbai - RN	UP	11	11	9	9	12	14	11	77
	DOWN	12	9	12	13	11	11	9	77
Mumbai - ROHA	UP	6	5	5	8	4	7	4	39
	DOWN	4	7	5	5	6	5	7	39
Total tasks	-	162	157	167	162	157	171	158	1134

3.3 Problem Objective and Crew Allocation Rules

The objective of this work is to find the minimum number of crew duties required to cover all the tasks in a defined period (which cover all the recurring set of tasks) based on a given train timetable. The recurring period in this case is a week since the maximum gap between repeating long-distance passenger services operated by Indian Railways is a week. In this paper, a task refers to a trip that a crew can operate in a single stretch and is characterized by starting and ending times and location. Additionally, duty refers to a sequence of tasks, satisfying all the operational constraints and labour union rules, that can be assigned to a crew to operate in the defined period.

In the Indian context, the crew allocation constraints are governed by the Railway Servants (Hours of Work and Period of Rest) Rules, 2005. The rules get amended from time to time, but the important rules applicable at the time of this study are as follows:

- Maximum continuous running (task) time for the crew is limited to nine hours
- Maximum task time for the crew from sign-on to sign-off is limited to 11 hours
- Maximum duty time for the crew in a fortnight is limited to 125 hours
- Crew should have a rest of at least 16 hours on returning to the home base
- Crew should have a rest of at least 8 hours at an outstation if the traversal time of the previous task is more than eight hours. If the traversal time is less than eight hours, the rest duration should be at least equal to the traversal time.
- Crew should return to the home base within 72 hours of leaving the home base

4 Solution Methodology

As an initial step, we model the crew scheduling problem using a combination of constraint programming (CP) and set covering problem (SCP) formulations. This helps in understanding and representing the problem in detail and also to assess its inherent complexity. This two-step approach is capable of solving smaller problem instances, whereas, for larger practical size instances, it fails to provide optimum solutions due to an explosion in combination possibilities. However, the partial solutions generated for such large instances can be used as a benchmark to evaluate the solutions provided by the heuristic.

4.1 Mathematical Models

In this two-step methodology, the CP model is used for duty generation and the SCP model is used for duty optimization. The first step is essentially a constraint satisfaction problem, and the objective is to generate all the feasible duties, adhering to all the constraints related to crew allocation. The second step tries to find the optimal set of duties from the set of all feasible duties, covering all tasks. However, there is a possibility depending on the size of the problem that all the feasible duties cannot be generated in a reasonable time by the CP model in the first phase. In that case, a Column-Generation (CG) procedure based on mathematical programming or constraint programming approach can be applied to solve the problem approximately. On the other hand, if the SCP is too large to solve, it can be relaxed as a large-scale linear programming problem, and conventional CG techniques can be applied further.

4.1.1 CP Model for Duty Generation

The goal here is to enumerate all the feasible duties for a defined period based on a given train timetable. The key advantage of the CP approach is that complex real-life constraints can be incorporated easily. Here, a duty is defined as a sequence of tasks $\{x_1, x_2, ..., x_m\}$, where $x_i \in T$ is the i^{th} task of the duty. The representation of duty is shown in fig.3. In addition to that, each task in a duty is defined by the following five parameters:

- train number (num_t),

- origin station (org_t),
- departure time $(dept_t)$,
- destination station $(dest_t)$, and
- arrival time (arr_t)

For each of the working tasks (W), these parameters are derived from the existing train timetable of the Mumbai division.

Figure 3: Representation of a duty

The parameter m limits the maximum number of tasks in a feasible duty and can be estimated by dividing the maximum permissible duty time in the planning horizon with the traversal time of the shortest working task. However, most feasible duties cannot include as many as m train tasks, and when this happens, dummy tasks are assigned to the duty. For a dummy task (d), a null value is assigned to its train number, a large number to its departure and arrival times and a dummy code for its origin and destination stations. The input data and parameters with their notations are summarized in table 2. The parameters, sign-on time (sn_t) and sign-off time (sf_t), are determined by providing some buffer (pw) before the scheduled departure time and after the scheduled arrival time of a task (30 minutes in our case). The crew utilizes this time for completing the paperwork and other checks before and after operating the task. The parameter, rh, defines the minimum amount of rest to be provided to a crew after operating a task, whereas, dh, puts a limit to the maximum permissible duty hours in a week. The parameter, rt, sets the upper limit on the time the crew can remain away from the home base at a stretch (72 hours in our case).

The constraints in this model are as follows:

$$sn_t = dept_t - pw, \ \forall\, t \in W \tag{1}$$

$$sf_t = arr_t + pw, \ \forall\, t \in W \tag{2}$$

Constraints (1) and (2) calculate sign-on and sign-off time for every working task by providing additional time before and after the scheduled departure and arrival time of the task, respectively.

The constraints related to maximum running time and time from sign-on to sign-off in a single stretch are taken care by appropriately dividing the entire train journey into smaller tasks using CCPs.

$$x_{i+1} \in W \implies org_{x_{i+1}} = dest_{x_i} \ \forall\, i \in \{1, 2, \ldots, m-1\} \tag{3}$$

Constraint (3) is a location fit constraint that ensures that the following working task in the duty originates from the same station where the previous task has ended.

$$\sum_{i=1:x_i \in W}^{m} sf_{x_i} - sn_{x_i} \leq dh \tag{4}$$

Table 2: CP model: Input data and parameters

Notation	Description
T	Set of all tasks (working and dummy)
W	Set of working tasks
d	A dummy task
$home$	Home base station
num_t	Train number of task t
org_t	Originating station of task t
$dest_t$	Destination station of task t
$dept_t$	Departure time of task t
arr_t	Arrival time of task t
sn_t	Sign on time for task t
sf_t	Sign off time for task t
pw	Time allotted for paper work and other checks
rh	Minimum rest to be provided after operating a task
dh	Maximum duty hours permissible in a week
rt	Maximum time a crew can stay away from their home base
m	Maximum number of permissible tasks in a duty

Constraint (4) ensures that the total duty time in a week doesn't exceed the maximum permissible limit.

$$x_{i+1} \in W \implies sn_{x_{i+1}} - sf_{x_i} \geq rh, \; \forall\, i \in \{1, 2, ..., m-1\} \tag{5}$$

Constraint (5) is the mandatory rest hour constraint that ensures sufficient rest is provided to a crew between two consecutive working tasks.

In the Indian context, the mandatory rest, rh, is defined based on the station where the rest is to be provided and the traversal time of the previous task undertaken by the crew.

$$rh = \begin{cases} 16 \text{ hours, if } dest_{x_i} = home \\ 8 \text{ hours, if } dest_{x_i} \neq home \;\; \& \;\; sf_{x_i} - sn_{x_i} \geq 8 \text{ hours} \\ sf_{x_i} - sn_{x_i}, \text{if } dest_{x_i} \neq home \;\; \& \;\; sf_{x_i} - sn_{x_i} < 8 \text{ hours} \end{cases} \tag{6}$$

Constraint (6) caters to the different possible cases as follows:

- if the previous task has ended at the home base, then the next working task in the duty can start after a rest of at least 16 hours.
- else if the previous task has ended at the outstation with a traversal time of 8 hours or more, then the next working task in the duty can start after a rest of at least 8 hours.
- else if the previous task has ended at the outstation with a traversal time of less than 8 hours, then the next train task in the duty can start after a rest equal to the traversal time of the previous task.

$$\begin{gathered} org_{x_i} = home \;\; \& \;\; sf_{x_j} - sn_{x_i} \geq rt \;\; \& \;\; x_i, x_j \in W \implies \\ \sum_{k=i}^{j-1} (dest_{x_k} = home) \geq 1, \; \forall\, i, \; j \; \in \{1, 2, ..., m\} \end{gathered} \tag{7}$$

Constraint (7) ensures the return of a crew within a specific time limit after leaving the home base.

The feasible duties generated from this CP model serve as an input to the SCP model, explained in the next section.

4.1.2 SCP Model for Duty Optimization

Set covering problem (SCP) is a well-known problem in combinatorial optimization and is proven to be NP-hard. Given a collection of elements, the SCP aims to find the minimum number of sets required to cover all these elements at least once. The other related problem is the set partitioning problem (SPP), where the objective is the same but covers all elements exactly once. As far as crew planning is concerned, SCP formulation is commonly used in the case of bus or railway modes of transportation where the deadheading of a crew is not that costly. However, SPP is widely used in the airline industry, where deadheading a crew is very expensive.

In this work, we use the SCP model and the input parameters used in the formulation are summarized in table 3.

Table 3: SCP model: Input data and parameters

Notation	Description
F	Set of feasible duties
W	Set of working tasks
a_{ij}	binary parameter, which takes a value 1 if the feasible duty i contains train task j, and 0 otherwise

The model is as follows:

$$\text{minimize} \sum_{i \in F} y_i \tag{8}$$

subject to,

$$\sum_{i \in F} a_{ij} \times y_i \geq 1 \;\forall\; j \in W \tag{9}$$

$$a_{ij} = \begin{cases} 1, \text{if duty } i \text{ contains task } j \;\forall\; i \in F,\; j \in W \\ 0, \;\text{otherwise} \end{cases} \tag{10}$$

$$y_i = \begin{cases} 1, \;\text{if duty } i \text{ is selected} \;\forall\; i \in F \\ 0, \;\text{otherwise} \end{cases} \tag{11}$$

The objective (8) is to select the minimum number of duties covering all the working tasks. Constraint (9) ensures that each working task is a part of at least one of the selected duties. Constraints (10) and (11) define the binary input data (a_{ij}) and decision variable (y_i) to the problem, respectively. Note that if any task appears in more than one selected duty, it would be a deadheading trip in all those duties except the one for which it would be a working task.

4.2 Heuristic Approach

The heuristic developed in this work extends the idea of the classic bin-packing problem to a crew scheduling setting with appropriate modifications. The bin packing problem is an optimization problem in which items of varying sizes must be packed into a finite number of bins, each of a fixed given capacity, by minimizing the number of bins used. Here, the crew tasks represent the items to be packed and crew duties represent the bins. The difference from the classic bin-packing problem is that here all the bins (duties) are not active all the time because of the location-fit constraints for the crew. Additionally, each new task that is being added to a duty depends on the previous task due to the differential rest rules applicable between them. Therefore, the selection of an appropriate bin (duty) for an item (task) is more complex in this setting.

In the heuristic also, a crew duty is defined in the same way it is described in the exact approach. The input data and parameters used in the heuristic are summarized in table 4.

Table 4: Heuristic: Input data and parameters

Notation	Description
W	The set of train tasks
$home$	Home base station
num_t	Train number of task t
org_t	Originating station of task t
$dest_t$	Destination station of task t
$dept_t$	Departure time of task t
arr_t	Arrival time of task t
sn_t	Sign on time for task t
sf_t	Sign off time for task t
pw	Time required for paper work
rh	Minimum rest to be provided after operating a task
dh	Maximum duty hours permissible in a week

The heuristic has been designed considering the rail network of the Mumbai division, where there is a central home base with all the outstation CCPs in the division connected to it. This characteristic ensures that an outstation-to-home trip follows a home-to-outstation trip. The heuristic can be customized for other divisions with minimal modifications. Further, the heuristic does not consider deadheading during duty construction. Deadheading may not be needed because the passenger timetable is symmetric in numbers in the two directions of travel and the services are quite frequent and are available throughout the day. Deadheading may be required for freight services, which could very well be unbalanced in requirements, as well as for certain extreme examples of passenger services. The current heuristic approach does not permit an easy extension to cover deadheading options. Scheduling without deadheading trips can help reduce the unproductive duty time in the system.

4.2.1 Heuristic for Duty Generation

The steps to constructively generate the crew duties are shown in the flowchart given in fig.4.

The idea behind the task allocation policy is to balance the total workload. Therefore, if the task originates from an outstation CCP, it is assigned to the first available duty. This ensures that no crew waits at an outstation for a longer duration. On the other hand, if the task originates from the home base CCP, it is assigned to a least loaded duty. This way, when the crew is at home base and has undertaken relatively longer

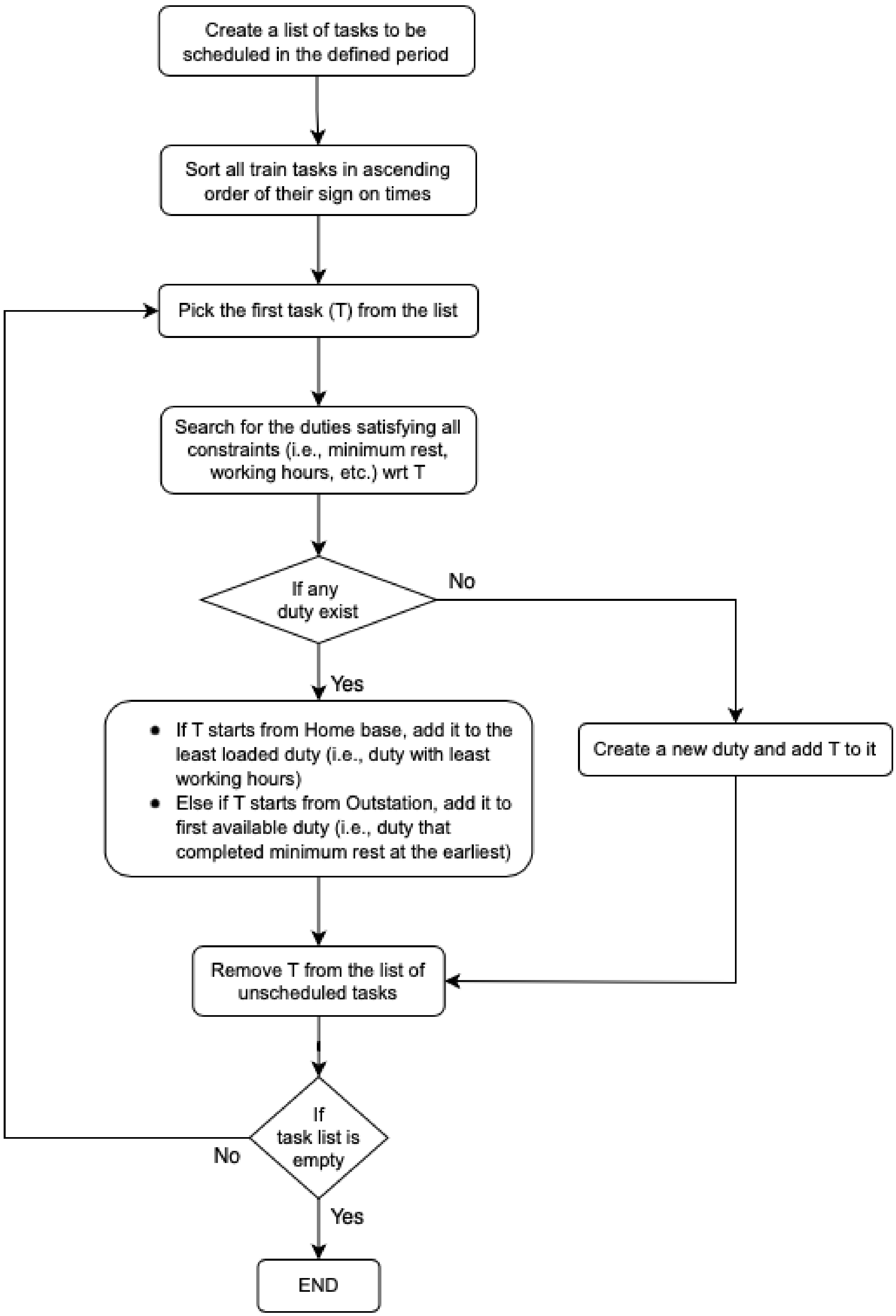


Figure 4: Flowchart of the heuristic

duty hours, they are given extra rest to balance the total workload. The tie-breaker is used if required. The task allocation policy and the tie-breaker rule are given in table 5.

Table 5: Heuristic policy for constructive task allocation and for connecting duties

Task originating station	Primary rule	Tie breaker
Home base CCP	Least loaded duty	First available duty
Outstation CCP	First available duty	Least loaded duty

Phases in heuristic solution generation: The heuristic generates a complete solution in three phases. In the first phase, the total number of duties required to cover all the given train tasks is estimated by linking feasible train tasks following the proposed task allocation policy given in table 5. However, these duties are not retained further since they do not ensure a uniform distribution of the total workload as the duties are created as and when required. This estimate is used to generate the final crew duties (for a week) in the second phase of the heuristic by ensuring uniform workload distribution to the extent possible. In the third phase, an additional set of weekly crew duties are generated following the same procedure with the condition that it should be feasible to connect the last and first tasks of each duty generated in the second and third phases, respectively. This condition ensures the connection of duties over successive periods (weeks) and enables the construction of longer duty cycles (one or more cycles) by sequencing the crew duties in a compatible manner. Duty cycles ensure the connectivity of crew duties subject to time, location, and operational constraints. From a crew duty cycle, a crew member can easily identify the crew duties to perform in succession. The same allocation policy as given in table 5 is used to establish a connection between compatible duties. Here, the total duty hours in the preceding week is considered to identify the least loaded duty.

The solution generated during all three phases of the heuristic for a sample problem instance (Mumbai-ROHA section) is given in table 6. Here, the numbers represent train running tasks (details given in Appendix B) and the alphabets represent duties, which consist of multiple tasks to be performed by a crew in sequence, during a week. Further, in this example, four duty cycles are derived in phase three of the heuristic from the 13 crew duties generated in phase two, respecting the relevant constraints, as shown in table 7. Duty cycle is a cyclic sequence of the compatible crew duties to be undertaken over a longer duration, maybe over some weeks/months. Each of the 13 duties generated in phase two of the heuristic will be a part of one of the four duty cycles and the 13 crew members are required to undertake these duties in the order as given in the corresponding duty cycle. Subsequently, each crew member will shift from one duty cycle to another and, in the long run, will complete all the duty cycles (all the train tasks in the region), thus balancing the total workload.

5 Results and Discussion

In this section, we discuss the computational results from the case study of the Mumbai Division.

5.1 Computational Results: Exact and Heuristic Approaches

The computational results obtained using the Exact and Heuristic approaches are presented in table 8. Here, both the section-wise and integrated crew scheduling have been tried separately and the results are reported.

In the exact approach, both the CP and SCP formulations have been modelled using IBM ILOG CPLEX Optimization Studio 22.1.0 (academic), and the experiments have been carried out on an Intel Xenon 3.5 GHz 64-bit personal computer with 16 GB RAM. Even though the recurring planning period for crew scheduling

Table 6: Heuristic solution: Mumbai - ROHA section

S. No.	Phase one - Estimation	Phase two - Week n		Phase three - Week (n+1)	
		Duty	Duty hours	Duty	Duty hours
1.	1 8 13 28 35 51 68 77	13 28 34 45 57 72 (A)	22.52	4 18 25 41 56 67 (H)	22.12
2.	7 11 18 34 45 58 74	6 21 39 53 61 76 (B)	22.08	12 22 27 43 48 65 (F)	21.25
3.	2 9 15 30 40 55 69	7 24 32 46 63 71 (C)	22.23	9 23 31 47 64 70 (K)	21.8
4.	3 10 14 29 37 52 60 75	11 20 35 51 68 77 (D)	22.62	13 28 34 45 57 72 (A)	22.52
5.	4 16 23 31 39 53 62 78	14 29 37 52 60 75 (E)	21.3	11 20 35 51 68 77 (D)	22.62
6.	5 17 24 32 38 54 61 76	12 22 27 43 48 65 (F)	21.25	1 10 33 44 58 74 (G)	22.55
7.	6 19 25 41 46 63	1 10 33 44 58 74 (G)	22.55	6 21 39 53 61 76 (B)	22.08
8.	12 20 26 42 57 72	4 18 25 41 56 67 (H)	22.12	3 17 38 54 62 78 (I)	22.92
9.	21 33 44 49 66	3 17 38 54 62 78 (I)	22.92	14 29 37 52 60 75 (E)	21.3
10.	22 27 43 48 65	2 16 36 50 59 73 (J)	23.43	5 19 26 42 49 66 (M)	21.02
11.	36 50 56 67 71	9 23 31 47 64 70 (K)	21.8	8 15 30 40 55 69 (L)	22.38
12.	47 64 70	8 15 30 40 55 69 (L)	22.38	7 24 32 46 63 71 (C)	22.23
13.	59 73	5 19 26 42 49 66 (M)	21.02	2 16 36 50 59 73 (J)	23.43

Table 7: Heuristic solution: Duty cycles for Mumbai - ROHA section derived from Table 6

S. No.	Duty Cycles
1.	A → H → I → E → D → A
2.	B → F → G → B
3.	C → K → L → C
4.	J → M → J

(to cover all the train tasks) is a week, it was not possible to solve the models for a week in any of the sections due to the larger size of the problem. Hence, for all the sections, the experiments have been run for the maximum time period possible (two - six days depending on the number of tasks). From table 8, it can be observed that with an increase in the number of working tasks and/or the planning horizon, the number of feasible duties explode due to combination possibilities and the CP model fails to generate all the feasible duties in a reasonable time. Therefore, the sections with comparatively fewer tasks (like Mumbai - ROHA) can be solved for a longer horizon than those with a larger number of tasks (like Mumbai- IGP). The partial solutions (solutions for a shorter planning period) obtained from this approach serve as a reasonable lower bound to compare the heuristic solutions.

The heuristic is coded in c++ and is found computationally efficient as it generates the complete solutions for all the problem instances (sections) quickly (less than a second). A comparison between the partial solutions (optimal for the corresponding shorter planning horizon) and the heuristic solutions can be made from table 8. The solution corresponding to the largest planning horizon that could be solved optimally in each scenario is used for this comparison. The average optimality gap (in percentage) between the heuristic and partial solutions, defined here as ((heuristic solution value - partial solution value)/partial solution value)x100, is found to be 10.625% for the section-wise scenarios. For the integrated crew planning scenario, the optimality gap is 2.5%. These optimality gap values indicate that the heuristic is able to generate reasonably good solutions in a very quick time. Further, the results from the heuristic suggest that integrated planning covers all the given train tasks with a 4% lesser number of crew duties than section-wise planning. Table 8 shows that for section-wise planning, the total number of crew duties adds up to 141 (Mumbai - IGP: 57, Mumbai - PUNE: 44, Mumbai - RN: 27 and Mumbai - ROHA: 13), while for integrated planning, the total number of crew duties is 135. It is because the combination possibilities increase in integrated planning due to which the waiting times at stations decrease, resulting in better crew utilization. This is further established in fig. 5 where statistics related to the total duty time in a week and waiting time after a task for different sections are shown on their respective arcs. The statistics for integrated planning are shown outside the network. Note that in the case of section-wise planning, the average duty hours (or the crew utilization) is significantly low in the Mumbai - ROHA section due to the fact that there are fewer tasks in this section and they are of shorter duration, which results in longer waiting times for the crew.

Table 8: Exact and Heuristic Results

Section	**Planning horizon**	**Working tasks**	**Exact solution**			**Heuristic solution**
			Feasible duty solutions	Optimal duties	CPU Time (CP, SCP) (sec)	No. of duties
Mumbai - IGP	1 day	78	314	54	1, <1	54
	2 days	152	16,254	**54**	85, 2	54
	3 days	234	>400,000	-	-	57
	4 days	309		-	-	57
	5 days	383	-	-	-	57
	6 days	463	-	-	-	57
	7 days	538	-	-	-	57
Mumbai - PUNE	1 day	51	103	40	<1, <1	40
	2 days	102	2,788	42	15, <1	42
	3 days	156	77,903	**43**	414, 11	43
	4 days	208	>400,000	-	-	44
	5 days	258	-	-	-	44
	6 days	312	-	-	-	44
	7 days	364	-	-	-	44
Mumbai - RN	1 day	23	38	19	<1, <1	19
	2 days	43	280	22	1, <1	22
	3 days	64	2,640	22	14, <1	22
	4 days	86	20,631	24	113, 2	25
	5 days	109	201,987	**24**	564, 57	27
	6 days	134	>400,000	-	-	27
	7 days	154	-	-	-	27
Mumbai - ROHA	1 day	10	17	7	<1, <1	7
	2 days	22	139	9	<1, <1	10
	3 days	32	930	9	5, < 1	10
	4 days	45	6,499	10	34, < 1	11
	5 days	55	44,936	10	239, 3	12
	6 days	67	280,199	**10**	985, 53	13
	7 days	78	>400,000	-	-	13
Mumbai Division (Integrated)	1 day	162	584	117	3, <1	118
	2 days	319	38,103	**120**	204, 10	123
	3 days	486	>400,000	-	-	126
	4 days	648	-	-	-	130
	5 days	805	-	-	-	132
	6 days	976	-	-	-	135
	7 days	1134	-	-	-	135

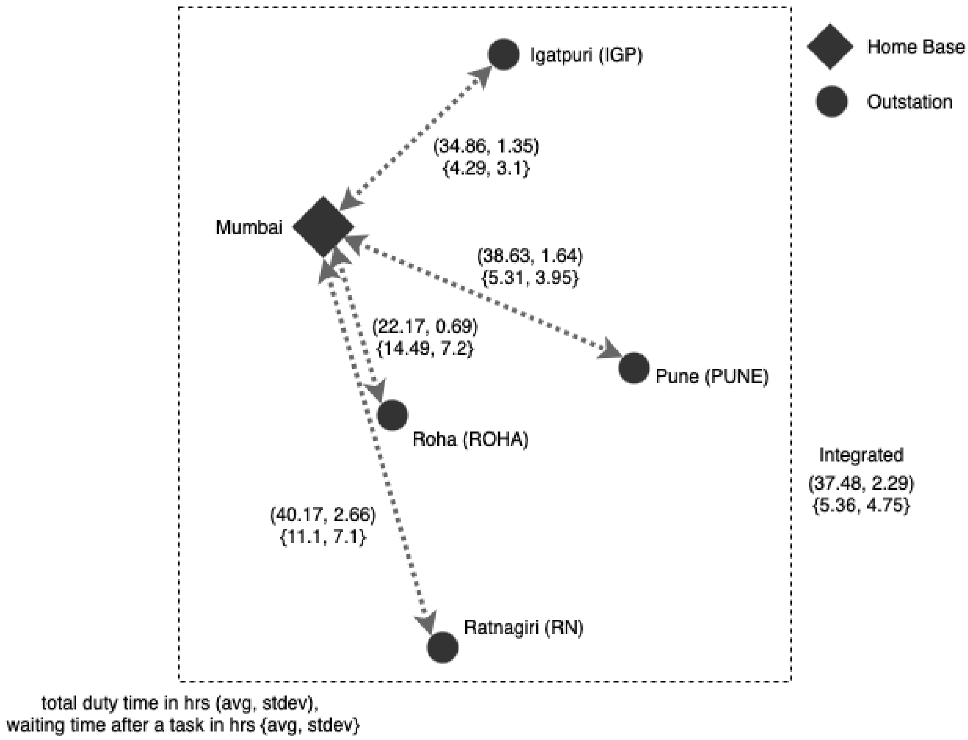


Figure 5: Heuristic solution: Duty hours and waiting hours

5.2 Comparison with the Existing Crew Links used by Mumbai Division

The existing crew links (crew link is the term used internally by Indian Railways to represent a duty cycle) used by the Mumbai division cover all the given train tasks in a week with 170 duties, which includes around 180 deadheading tasks. The solution generated by the heuristic shows that the same set of tasks can be covered with 135 duties. An additional 10% is usually reserved for periodical rest and possible fluctuations in traversal times. Thus, the final rounded-off number comes to 150, which is almost 12% less than the existing number of duties. This shows that there is potential for significant savings in terms of crew costs. The crew links generated by the heuristic are found to be better because the existing method of generating crew links is not fully integrated (crew links for daily and non-daily trains are prepared separately) and depend highly on the skills of the personnel involved. In this work, we don't distinguish based on the train type or crew type and assume that any task can be assigned to any crew. We also find that there is scope for reducing the deadheading tasks, which are unproductive and costly. However, in each case, the final crew requirement would be more than the number of duties to compensate for the crew members on leave, training, etc., as shown in table 9. These concerns are mostly handled at the crew rostering stage.

Table 9: Estimate of total crew requirement in Mumbai Division

Components	**Existing crew links**	**Crew links generated by the heuristic**
Crew duties [A]	170	150
(Deadheading trips)	(180)	(0)
Leave (10%) [B]	17	15
Other contingencies (5%) [C]	8	7
Total crew requirement [A+B+C]	195	172

6 Conclusion

Railway crew scheduling problems are complex, and finding an optimal solution for large-size, real-life instances is a challenge due to an explosion in combination possibilities. This paper proposes a heuristic which extends the idea of the classic bin-packing problem to generate crew schedules that can be repeated over periodic train schedules (connected crew schedules) by incorporating the rules related to crew allocation. By ensuring the connectivity of crew schedules, a more realistic estimate of the crew requirement in a region is also obtained, thus linking both an operational-level crew scheduling problem and a tactical-level crew capacity planning problem. Further, the crew duty cycles generated by the heuristic (by linking various crew duties) also ensure a fair distribution of workload among the crew and in the long run, each crew takes up all the train tasks associated with that crew region. These cyclic crew links also help in preparing cyclic crew rosters in the subsequent planning stage.

In this work, we first implement an exact approach based on a combination of constraint programming and set covering formulations to generate optimum crew schedules. We consider the additional constraints like differential rest rules for outstation and return journeys, rest rules that depend upon the duration of the trips, and a limit on the maximum time a crew can stay away from their home base, which are applicable in the Indian context. Also, the heterogeneity in the number of long-distance passenger train services operated on different days in a week extends the planning horizon to a week, thus making the problem more complex. The models can be solved in sequence to obtain optimum solutions for small-sized problem instances. These results are also used for validating the crew schedules generated by the heuristic and for obtaining valid lower bounds for larger problem instances. A comparison of the heuristic results with the existing crew schedules used by the Mumbai Division shows a possibility of reducing the total duty requirement by around 12%. Based on the results, we also recommend integrated crew planning over section-wise planning as the former results in better crew utilization due to more combination possibilities.

The heuristic approach provides an efficient way to generate improved crew schedules in the region whenever there is any modification in the corresponding train schedules. Currently, a long and time-consuming manual process is followed for crew duty generation which involves multiple stakeholders. The heuristic is developed considering the Mumbai division network and it can be customized as required to be applied elsewhere. Also, the heuristic can be improved further to make the subsequent crew rostering stage easier by incorporating the periodical rest allotted to crew members in a systematic way. In the proposed integrated model, we assume that all crew can operate any given task in a region. Even though this is found feasible with respect to the familiarity of routes (route knowledge) in a region, there can be issues when a certain crew is not trained to operate a particular type of locomotive or train set. In that case, the train tasks are to be split into multiple sets and corresponding sets of crew cycles are to be generated by repeated execution of the proposed heuristic. This ensures that each crew cycle contains only those tasks that can be operated by a particular set of crew. In Indian Railways, a section of the senior crew members who are not trained for operating electric locomotives are allocated tasks on trains with diesel locomotives on special request. Multiple sets of duty cycles are generated in this case. Further, the effect of probable service delays and disruptions on crew duties is handled in the following ways: (i) defining the train tasks by maintaining adequate buffer time between the scheduled duration of the task and the maximum continuous running time of the crew, and (ii) adding an additional 10% of the actual crew requirement to the crew strength as per industry standards. Future research can focus on the following related aspects: (i) the addition of different types of train tasks and crew sets into the model, and (ii) incorporating resilience parameters in the modelling framework to enhance the robustness of the crew schedules.

Author Contributions: The authors confirm contribution to the paper as follows: Study conception and

design: Akshat Bansal, Anoop K P, Narayan Rangaraj; Data collection: Akshat Bansal, Anoop K P, Narayan Rangaraj; Analysis and interpretation of results: Akshat Bansal, Anoop K P, Narayan Rangaraj; Draft manuscript preparation: Akshat Bansal, Anoop K P, Narayan Rangaraj. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgements: The authors thank Mr. K. N. Singh (Former Chief Freight Transportation Manager), Mr. P. N. Ramachandran (Senior Divisional Operations Manager, Coaching) and other railway personnel at Central Railway (Indian Railways) for supporting this study by providing insights about the actual operations.

Funding: The authors thank the Ministry of Education, Government of India and Indian Institute of Technology Bombay for supporting this work through Post-graduate and Post-doctoral fellowships. The funding agency has not been involved in the design and execution of the study, nor in the collection, analysis and interpretation of data, the writing of the report, and the decision to submit the paper for publication.

Declaration of Conflicting Interests: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

[1] J. Heil, K. Hoffmann, and U. Buscher. Railway crew scheduling: Models, methods and applications. *European Journal of Operational Research*, 283(2):405–425, 2020.

[2] G. Şahin and B. Yüceoğlu. Tactical crew planning in railways. *Transportation Research Part E: Logistics and Transportation Review*, 47(6):1221–1243, 2011.

[3] A. Ç. Suyabatmaz and G. Şahin. Railway crew capacity planning problem with connectivity of schedules. *Transportation Research Part E: Logistics and Transportation Review*, 84:88–100, 2015.

[4] J. P. Arabeyre, J. Fearnley, F. C. Steiger, and W. Teather. The airline crew scheduling problem: A survey. *Transportation Science*, 3(2):140–163, 1969.

[5] A. Wren. A general review of the use of computers in scheduling buses and their crews. *Publication of: Leeds University, England*, 1980.

[6] A. Caprara, M. Fischetti, P. Toth, D. Vigo, and P. L. Guida. Algorithms for railway crew management. *Mathematical programming*, 79(1):125–141, 1997.

[7] E. J. W. Abbink, L. Albino, T. Dollevoet, D. Huisman, J. Roussado, and R. L. Saldanha. Solving large scale crew scheduling problems in practice. *Public Transport*, 3(2):149–164, 2011.

[8] J. S. Neufeld, M. Scheffler, F. Tamke, K. Hoffmann, and U. Buscher. An efficient column generation approach for practical railway crew scheduling with attendance rates. *European Journal of Operational Research*, 293(3): 1113–1130, 2021.

[9] A. Khosravi, M. Tamannaei, and M. Reisi-Nafchi. A comprehensive approach for railway crew scheduling problem (case study: Iranian railway network). *International Journal of Transportation Engineering*, 4(3):197–210, 2017.

[10] A. Khosravi and M. Tamannaei. Two mathematical models for railway crew scheduling problem. *International Journal of Railway Research*, 4(2):11–22, 2017.

[11] N. Kasliwal, S. Pulapadi, M. N. Belur, N. Rangaraj, S. Mishra, S. Monga, A. Singh, S. G. Sagar, P. K. Majumdar, and M. K. Jagesh. Crew planning for commuter rail operations, a case study on mumbai, india. In *Operations Research Proceedings 2019: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR), Dresden, Germany, September 4-6, 2019*, pages 813–819. Springer, 2020.

[12] B.T.C. van Rossuma, T. Dollevoeta, and D. Huismana. Operational railway crew scheduling with equity concerns. 2022.

[13] S. Jütte, M. Albers, U. W. Thonemann, and K. Haase. Optimizing railway crew scheduling at db schenker. *Interfaces*, 41(2):109–122, 2011.

[14] S. Jütte and U. W. Thonemann. Divide-and-price: A decomposition algorithm for solving large railway crew scheduling problems. *European Journal of Operational Research*, 219(2):214–223, 2012.

[15] R. Hanafi and E. Kozan. A hybrid constructive heuristic and simulated annealing for railway crew scheduling. *Computers & Industrial Engineering*, 70:11–19, 2014.

[16] M. Liu, A. Haghani, and S. Toobaie. Genetic algorithm-based column generation approach to passenger rail crew scheduling. *Transportation research record*, 2159(1):36–43, 2010.

[17] T. Nishi, Y. Muroi, and M. Inuiguchi. Column generation with dual inequalities for railway crew scheduling problems. *Public transport*, 3(1):25–42, 2011.

[18] K. Hoffmann, U. Buscher, J. S. Neufeld, and F. Tamke. Solving practical railway crew scheduling problems with attendance rates. *Business & Information Systems Engineering*, 59(3):147–159, 2017.

[19] K. Hoffmann and U. Buscher. Valid inequalities for the arc flow formulation of the railway crew scheduling problem with attendance rates. *Computers & Industrial Engineering*, 127:1143–1152, 2019.

[20] T. Gattermann-Itschert, L. M. Poreschack, and U. W. Thonemann. Using machine learning to include planners' preferences in railway crew scheduling optimization. *Transportation Science*, 57(3):796–812, 2023.

[21] A. F. Han and E. C. Li. A constraint programming-based approach to the crew scheduling problem of the taipei mass rapid transit system. *Annals of Operations Research*, 223(1):173–193, 2014.

[22] G. H. Chen, J. C. Jong, and A. F. W. Han. Applying constraint programming and integer programming to solve the crew scheduling problem for railroad systems: Model formulation and a case study. *Transportation Research Record*, 2676(1):408–420, 2022.

[23] P. Tapkan, S. Kulluk, L. Özbakır, F. Bahar, and B. Gülmez. A constraint programming based column generation approach for crew scheduling: A case study for the kayseri railway. *Journal of the Operational Research Society*, pages 1–15, 2022.

[24] Z. Feng and X. Ruihua. Study on model and algorithm for urban rail transit crew scheduling system. In *2010 International Conference On Computer Design and Applications*, volume 5, pages 213–216. IEEE, 2010.

[25] W. Qiao, M. Hamedi, and A. Haghani. Algorithm for crew-scheduling problem with bin-packing features. *Transportation research record*, 2197(1):80–88, 2010.

[26] A. T. Ernst, H. Jiang, M. Krishnamoorthy, H. Nott, and D. Sier. An integrated optimization model for train crew management. *Annals of Operations Research*, 108(1):211–224, 2001.

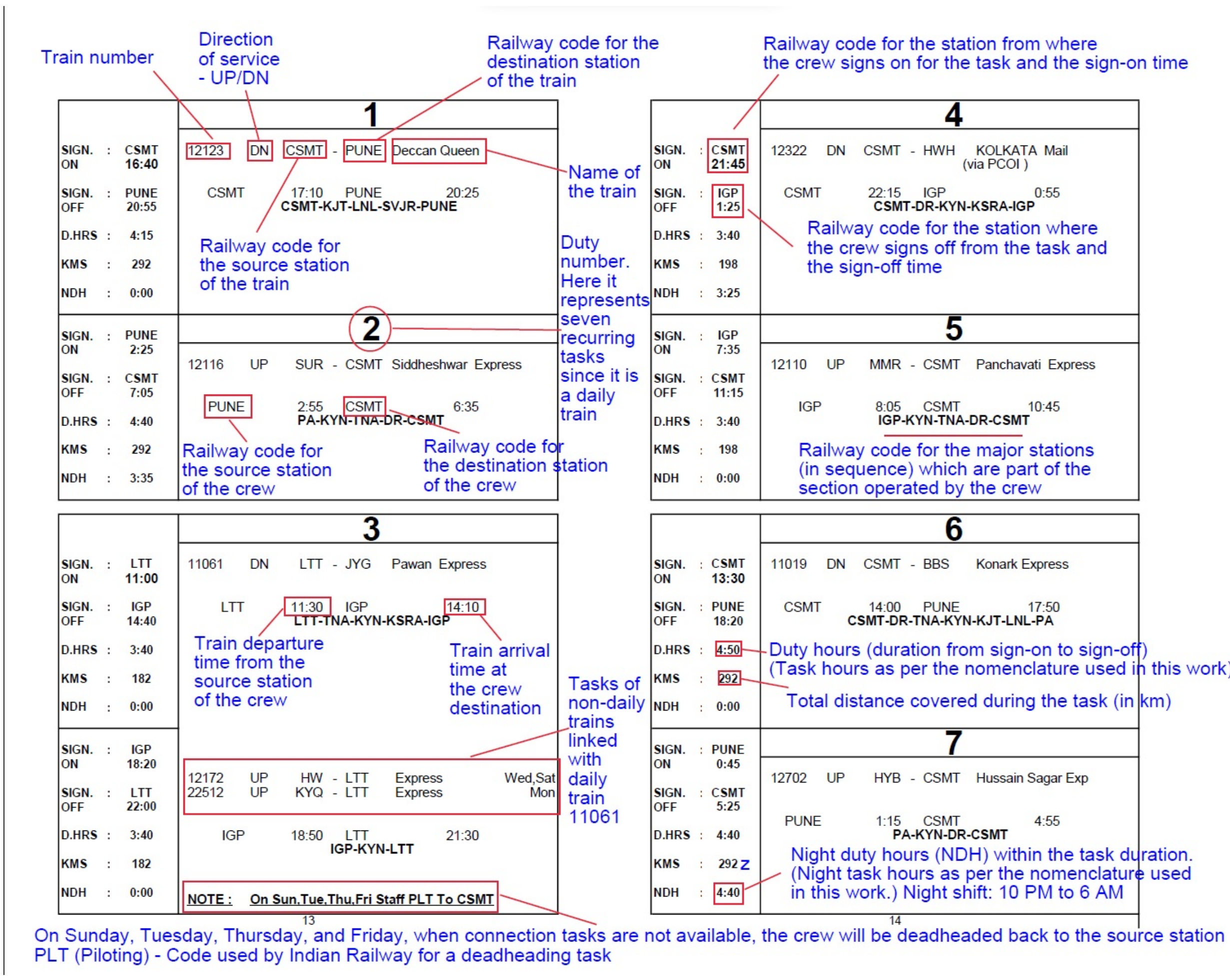
[27] D. Y. Lin and M. R. Tsai. Integrated crew scheduling and roster problem for trainmasters of passenger railway transportation. *IEEE Access*, 7:27362–27375, 2019.

[28] T. Feng, R. M. Lusby, Y. Zhang, Q. Peng, P. Shang, and S. Tao. An admm-based dual decomposition mechanism for integrating crew scheduling and rostering in an urban rail transit line. *Transportation Research Part C: Emerging Technologies*, 149:104081, 2023.

[29] Government of India. Ministry of Railways. Indian railways annual report and accounts 2021-22, 2023. Available from https://iritm.indianrailways.gov.in/instt//uploads/files/1680252554005-Indian%20Railways%20Annual%20Report%20&%20Accounts%20English%202021-22_web_Final.pdf. Accessed July 20, 2023.

Appendix A

Snapshot of crew detail book used by Mumbai Division, Indian Railways (annotations given in blue colour)



Appendix B

Details of Mumbai-ROHA tasks

Task	Day	Train	Origin station	Departure time	Destination station	Arrival time	Task	Day	Train	Origin station	Departure time	Destination station	Arrival time
1	M	11003	MUM	00:05	ROHA	02:50	40	Th	22629	MUM	20:40	ROHA	23:25
2	M	11085	MUM	05:33	ROHA	08:15	41	Th	11004	ROHA	03:33	MUM	06:40
3	M	10105	MUM	06:25	ROHA	09:05	42	Th	50104	ROHA	11:00	MUM	13:25
4	M	16345	MUM	11:40	ROHA	14:10	43	Th	22630	ROHA	12:20	MUM	15:00
5	M	50103	MUM	15:20	ROHA	17:30	44	Th	16346	ROHA	13:52	MUM	16:45
6	M	12450	MUM	23:10	ROHA	02:30	45	Th	10106	ROHA	17:35	MUM	20:10
7	M	11004	ROHA	03:33	MUM	06:40	46	F	11003	MUM	00:05	ROHA	02:50
8	M	50104	ROHA	11:00	MUM	13:25	47	F	10105	MUM	06:25	ROHA	09:05
9	M	16346	ROHA	13:52	MUM	16:45	48	F	16345	MUM	11:40	ROHA	14:10
10	M	10106	ROHA	17:35	MUM	20:10	49	F	50103	MUM	15:20	ROHA	17:30
11	Tu	11003	MUM	00:05	ROHA	02:50	50	F	11004	ROHA	03:33	MUM	06:40
12	Tu	10105	MUM	06:25	ROHA	09:05	51	F	50104	ROHA	11:00	MUM	13:25
13	Tu	16345	MUM	11:40	ROHA	14:10	52	F	16346	ROHA	13:52	MUM	16:45
14	Tu	50103	MUM	15:20	ROHA	17:30	53	F	11086	ROHA	14:13	MUM	17:10
15	Tu	19260	MUM	22:50	ROHA	01:30	54	F	10106	ROHA	17:35	MUM	20:10
16	Tu	11004	ROHA	03:33	MUM	06:40	55	F	19259	ROHA	20:10	MUM	23:10
17	Tu	12217	ROHA	09:35	MUM	12:45	56	Sa	11003	MUM	00:05	ROHA	02:50
18	Tu	50104	ROHA	11:00	MUM	13:25	57	Sa	11099	MUM	00:45	ROHA	03:15
19	Tu	16346	ROHA	13:52	MUM	16:45	58	Sa	12218	MUM	06:10	ROHA	09:20
20	Tu	11086	ROHA	14:13	MUM	17:10	59	Sa	10105	MUM	06:25	ROHA	09:05
21	Tu	10106	ROHA	17:35	MUM	20:10	60	Sa	16345	MUM	11:40	ROHA	14:10
22	Tu	12449	ROHA	18:40	MUM	21:05	61	Sa	50103	MUM	15:20	ROHA	17:30
23	W	11003	MUM	00:05	ROHA	02:50	62	Sa	12450	MUM	23:10	ROHA	02:30
24	W	10105	MUM	06:25	ROHA	09:05	63	Sa	11004	ROHA	03:33	MUM	06:40
25	W	20932	MUM	10:10	ROHA	12:55	64	Sa	50104	ROHA	11:00	MUM	13:25
26	W	16345	MUM	11:40	ROHA	14:10	65	Sa	20931	ROHA	12:20	MUM	15:10
27	W	50103	MUM	15:20	ROHA	17:30	66	Sa	16346	ROHA	13:52	MUM	16:45
28	W	11004	ROHA	03:33	MUM	06:40	67	Sa	10106	ROHA	17:35	MUM	20:10
29	W	50104	ROHA	11:00	MUM	13:25	68	Su	11003	MUM	00:05	ROHA	02:50
30	W	16346	ROHA	13:52	MUM	16:45	69	Su	10105	MUM	06:25	ROHA	09:05
31	W	10106	ROHA	17:35	MUM	20:10	70	Su	16345	MUM	11:40	ROHA	14:10
32	W	12449	ROHA	18:40	MUM	21:05	71	Su	50103	MUM	15:20	ROHA	17:30
33	Th	11003	MUM	00:05	ROHA	02:50	72	Su	11004	ROHA	03:33	MUM	06:40
34	Th	11085	MUM	05:33	ROHA	08:15	73	Su	12217	ROHA	09:35	MUM	12:45
35	Th	12218	MUM	06:10	ROHA	09:20	74	Su	50104	ROHA	11:00	MUM	13:25
36	Th	10105	MUM	06:25	ROHA	09:05	75	Su	22476	ROHA	12:20	MUM	15:10
37	Th	16345	MUM	11:40	ROHA	14:10	76	Su	16346	ROHA	13:52	MUM	16:45
38	Th	22475	MUM	14:15	ROHA	17:00	77	Su	10106	ROHA	17:35	MUM	20:10
39	Th	50103	MUM	15:20	ROHA	17:30	78	Su	11100	ROHA	21:20	MUM	23:45