Shift Scheduling for Train Dispatchers*

Tomas Lidén, Christiane Schmidt, Rabii Zahir [a]
Communications and Transport Systems, ITN
Linköping University, Norrköping, Sweden
[a] E-mail: rabii.zahir@liu.se, Phone: +46 (0) 11363348

The type of a submission
This submission is of type A.

Abstract
Train dispatchers monitor and control train traffic from a dispatching center, which is responsible for a certain region in the railway network. This region is divided into subareas, where each train dispatcher controls one or several subareas at any time. Given the high safety concerns of their profession, dispatchers' working shifts should fulfill several legal and operational constraints, such as bounds on the length of shifts and on the resting periods between shifts. To construct shift schedules for train dispatchers is a complex and time-consuming process that is currently done manually. In this paper, we present an optimization framework to automate this process, based on a model for single-day shifts. Here, we focus on the objective of minimizing the number of dispatchers as a baseline for future objectives.
We present experimental results for real-world sized data (number of geographical areas and train movements in the order of magnitude as for one dispatching center in Sweden, covering the southern part of the country). We study the impact on the run time for different input parameters, namely: the total number of geographical areas, the maximum number of geographical areas that can be assigned to a dispatcher in any period, changes in adjacency between the geographical areas, and the number of geographical areas that each dispatcher is qualified to monitor. The run time for the instances is between 19 and 305 seconds.

Keywords
Integer Programming, Shift scheduling, Railway dispatching, Area combination

1 Introduction

Thanks to their high speed, large loading capacity, high degree of safety and reduced environmental impact, trains are considered as one of the best choices among the available transportation modes for both passengers and freights. Due to these and other advantages, rail transportation is becoming more and more popular among users, which is confirmed by the increase in the passenger numbers during the last two decades. In Sweden, the number of train passengers increased by 75% between the years 2004 and 2019 (EUROSTAT (2022a)), while the train movements, given in train kilometers, increased by 32% during the same period (EUROSTAT (2022b)).

*The first author of this paper is Rabii Zahir, with the main ideas and help coming from Christiane Schmidt, and minor support, modelling suggestions and reviewing from Tomas Lidén

This increase in train traffic brings with it many challenges of how to manage the infrastructure to accomplish the operations safely and efficiently. The safety here refers to avoiding hazardous situations that involve trains and/or maintenance crew out in the field, while the efficiency refers to minimizing the operational costs. Train delays and cancellations cause significantly high costs for train operating companies as well as for the whole society. The major work—besides guaranteeing safety—to reduce these costs during operations is done in the dispatching center. The dynamicity and complexity of train traffic require quick and continuous decision making, which leads to an increase of the train dispatchers' mental cost, which is commonly called workload (WL). Generally, the WL is correlated with the number and complexity of the executed tasks. Given the finite number of available resources, such as the number of dispatchers and/or working positions, it is important to limit the number of tasks for each dispatcher such that none of them should experience an overwhelming WL. This limitation should already be considered during the shift planning phase, which is also called shift scheduling. Generally, shift scheduling is a tactical planning step that consists in defining when each shift will start and end, and later to assign personnel to their work stations.

The Swedish railway network is controlled by eight different dispatching centers. A major one is in Malmö and it controls the traffic in the southern region of the country. The shift scheduling practices at Malmö are considered representative for most of the Swedish dispatching centers. This main region is divided into smaller geographical areas that should be controlled by one dispatcher. An area could be one or several train stations, one or several track sections or a combination of them. To continuously guarantee traffic safety, each area must always be controlled by some dispatcher.

In the current practice, train dispatchers' schedules are made by hand, which is a difficult and time-consuming task with a considerable space for improvement—e.g., by automating the planning process and by improving the quality of the schedules. The complexity of the scheduling process stems from the number and type of restrictions that a feasible schedule must fulfill. These constraints are both legal ones, such as the allowed shift length or the maximum number of consecutive night shifts, as well as operational constraints, such as the number of areas a dispatchers can handle during each period or how these areas could be combined. To our knowledge, no previous research has tackled the issue of train dispatchers shift scheduling, and therefore we see this as an unexplored area with important possible improvements.

The main goal of this paper is to present an optimization framework that can be used for improving or automating the shift scheduling. The generated shift schedules should fulfill all the legal and operational constraints that are adopted in Malmö dispatching center. In addition, no dispatcher should be assigned a too high WL. Since the WL is a subjective metric that is difficult to estimate, we, in this paper, choose to use task load as a simplification of it. Moreover, we approximate task load by the number of train movements in each geographical area during each time period. In our optimization problem, we aim to obtain a shift schedule with the minimum labor cost, which means with the minimum number of used dispatchers.

As a contribution of this paper, we formulate an optimization framework to automatize the process of train dispatchers' shift scheduling. The framework considers limitations regarding shift lengths, task load, dispatcher competences and feasible area assignments. In this work we present a novel approach for modelling the combination of the areas to assign to each dispatcher. Moreover, we run an experimental study where we change different parameters, such as the maximum number of areas that can be combined and the ratio of areas for which dispatchers hold an endorsement. However, we do not consider ad hoc re-scheduling in case of extreme events such as major weather disturbances.

The paper is organized as follows: in Section 2, we present related work on the dispatching work environment and on shift scheduling in general. In Section 3, we present the problem statement. In Section 4, we describe our IP model. We present the experiments and discuss their results in Section 5. Finally, we conclude and give suggestions for future work in Section 6.

2 Related Work

In this section, we describe research literature related to the train dispatchers' working environment, and to shift scheduling.

2.1 The Working Environment at Train Dispatching Centers

Railway operations are executed according to tactically predefined plans, which consist in timetables for train movements and maintenance activities. In an ideal world, all the operations would be executed according to the original plans. But in reality, the operations are occasionally disturbed, generating delays, which makes replanning necessary (Marinov et al. (2013)). The causes of delays are many, examples of these are different technical failures in the rolling stocks or the infrastructure, such as stop signals and switches. Severe weather conditions, such as too low temperatures and heavy snow, are also a reason for delays (Mücke (2002)). The recovery from the disturbances is within the dispatchers' responsibility, who should guarantee safe and efficient train movements. To do that, train dispatchers constantly monitor train movements so that they can forecast possible conflicts and then eventually reschedule some trains (Marinov et al. (2013)). Possible rescheduling interventions could be delaying, rerouting or even cancelling some trains (Cacchiani and Toth (2018)). Generally, train dispatchers follow some prioritization rules, i.e., which trains should run first and which could be delayed or cancelled, for making the rescheduling decision. As an effectiveness measure of the taken decision, dispatchers may aim for minimizing the total (weighted) delay or for reducing the deviation from the original timetable as much as possible (Marinov et al. (2013)). Apart from monitoring and controlling the train movements and maintenance works, the dispatchers have a large number of other responsibilities and administrative tasks. Reinach (2006) identified 67 tasks that are executed by train dispatchers, and they grouped them into six categories:

- Actuate signals, switches, blocking devices and bridge controls
- Issue/void dispatcher-authorized mandatory directives
- Grant other track-related permissions, protections and clearances
- Carry out non-authority or non-permission/protection/clearance communications
- Perform general record-keeping tasks
- Review reference materials

Huang et al. (2018) made observations inside a dispatching center and listed a set of tools that dispatchers are provided with. The tool set consists of paper and pencil, communication consoles (radiocommunication system, cellphone and land phone), PC monitors, digital forms (for standard information exchange between dispatchers and other entities) and a Computer Aided Dispatch system, which provides a visual presentation of the trains in the network. With the introduction of Centralized Traffic Control, where train dispatchers have a direct control over the signals and switches in the network, the number of communication calls and warrants decreased but at the same time the dispatchers' responsibility became heavier (Dima et al. (2017)). Due to the complexity of and the limited available time for rescheduling, many models and algorithms were designed to support dispatchers to take a quick and efficient decision but very few of them have been implemented (Törnquist (2006)). Consequently, most of the rescheduling tasks are done manually, especially for freight trains since their timetable is more flexible than that of passenger trains (Palmqvist et al. (2022)). This puts more weight on the dispatchers' innate characteristics, such as stress tolerance, good short- and long-term memory, a short reaction time, convergent thinking. These and other characteristics are among the highest ranked requirements for selecting train dispatcher candidates (Gertler (2003)). Over time, dispatchers sharpen their skills by receiving formal training and by accumulating experience, which make them qualified for controlling more complex and busier geographical areas (Huang et al. (2018)).

In an environment with such a huge responsibility, a lack of proper WL policies and their accurate implementation may result in very serious consequences (NTSB (1998)). Thus, WL level is an important aspect for train dispatchers and must be taken in consideration when designing control areas and constructing shift schedules.

2.2 Shift Scheduling

As labor cost has a major weight on the companies' economy, many employers are very concerned of making their staffing process as effective as possible. This means that the staffing level should neither be too low, which would reduce the customers' service level, nor too high, which would increase the total labor costs. This interest from the companies resulted in an extensive exploration of the area of personnel scheduling during the last decades. As a limitation in our literature review, we focus on shift scheduling at fixed work places where the task load is continuous and not given by predefined and timed specific tasks, which is the case in some professions like airline and train crew.

Baker (1976) categorized personnel scheduling into three categories: shift scheduling over one day; scheduling days-off patterns over longer time horizons; and the combination of the first two. The shift scheduling problem was first presented by Dantzig (1954) as a set-cover problem where the objective function was to minimize the total labor cost. This approach becomes impractical for solving instances with a larger number of shift patterns, especially in cases with temporal fluctuation of demand. To reduce the memory usage and computation time, Bechtold and Jacobs (1990) implemented an implicit modeling approach, which yielded an equivalent formulation to the one by Dantzig, where breaks are assigned flexibly.

Apart from the labor cost, researchers pointed out different aspects to consider during the construction process. Moen and Yu (2000) examined the relationship between the number of working hours and the life quality in terms of stress, overload and work/family conflicts. Colligan and Rosa (1990) studied the effect of shiftwork type on the employees' social and family life, where the shiftwork types are fixed days, fixed afternoons, fixed nights and rotating shifts. Barnett (2006) noted how the employees' life quality is impacted not only by the number of working hours but also by the distribution of these hours, especially in today's society where nonstandard working schedules are becoming more popular. Ahasan (2002) highlighted the importance of an effective design of shift schedules that takes in consideration job satisfaction and the social life of the employees. Since the effectiveness of a design is affected by the level of automation and the number of dispatchers, Zeilstra et al. (2017) developed the WASCAL-Tool to predict the staffing level for train dispatching given a certain level of automation. Cunha et al. (2020) evaluated the benefits of 12-h working shifts and denoted its acceptance among the industrial workers despite its negative impact in terms of fatigue, muscle pain and sleep deprivation.

During the last decades, several approaches have been developed to model and solve shift scheduling problems in many different sectors. Alfares (2007) formulated an Integer Programming (IP) model for staffing and scheduling of operators in an IT-help center. In the first step, the author estimated the hourly labor demand based on the number of received calls, and then used the estimated demand to determine the weekly schemes with the lowest labor cost. Bard, Binici, et al. (2003) also used an IP approach to schedule postal service operators and evaluated different scenarios, such as the usage of part-time workers, restricting two days off to be consecutive, and other parametric manipulations. Wright, Bretthauer, et al. (2006) presented a multi-objective nonlinear integer model to schedule nurses, where the first objective function is to minimize the total regular time and overtime wage, and the second objective is to increase the job satisfaction by minimizing the total number of undesirable shifts per nurse. That model was improved by Wright and Mahar (2013), who argued the benefits, in terms of lower labor costs and higher job satisfaction, of having a centralized nurse scheduling for all hospital departments. Breugem et al. (2022) introduced the fairness-oriented crew rostering problem, where they optimize cyclic rosters considering explicitly the trade-off between the fairness and attractiveness of the assignments as perceived by the employees. Zhong et al. (2017) proposed a two-stage heuristic algorithm for nurse scheduling and rostering (i.e., assigning shifts to individual workers) taking into consideration a fair distribution of working weekends among the caregivers. Another two-stage approach was proposed by Kim and Mehrotra (2015) who integrated nurse staffing and scheduling given a stochastic patient demand. Hur et al. (2019) formulated a multiple-stage optimization model for airport baggage handlers under uncertain demand and argued the benefits of using flexible breaks in the resulting shifts. Lapègue et al. (2013) presented what they call *the Shift-Design Personnel Task Scheduling Problem with an Equity Criterion*, where they optimized shift schedules for pharmaceutical personnel with the goal of having the fairest possible distribution of task load among the employees. Other optimization methods have been implemented to solve scheduling problems, such as column generation (Bard and Purnomo (2005); Al-Yakoob and Sherali (2008)), simulating annealing (Akbari et al. (2013); Cordeau et al. (2010)) and genetic algorithms (Frey et al. (2009). Recently, Josefsson et al. (2017) formulated a Mixed Integer Linear Program (MILP) to construct rosters for air traffic controllers operating from a remote tower center. Their model minimizes the number of used controllers allowing some of them to control several airports simultaneously based on the actual number of movements (departures/arrivals) in each airport. We will formulate an IP problem using a similar modelling approach, adapted to the railway dispatching case.

3 Problem Statement

Given: A set of geographical areas covering a railway network and a set of subsets representing different combinations of these areas; a set of train dispatchers; a list for each dispatcher describing the areas for which they hold an endorsement; a set of time periods; task load values representing the train movements in each area and time period; an upper bound on the task load per dispatcher and period; upper and lower bounds on the shift length and the resting period between two shifts; and an upper bound for the number of controllable areas per dispatcher per period.
Find: A one-day shift schedule that uses the fewest possible dispatchers while respecting the following legal and operational constraints, that:

- Each area is assigned a dispatcher in every time period
- Dispatchers are assigned one, or several combinable areas, for which they hold an endorsement
- Each dispatcher shift complies with the upper and lower bounds on the shift length and the resting periods between shifts
- The task load for each dispatcher does not exceed $TL^{\max}$ in any time period

4 The Optimization Model

In this section, we present the optimization model for shift scheduling for train dispatchers. We start by describing the model notation in Subsection 4.1, then we present its mathematical formulation in Subsection 4.2. To formulate the legal and operational constraints, we use information obtained from the union agreement (Trafikverket (2019)) and from interviews with experts from Trafikverket.

4.1 Model Notation

Table 1: Model Parameters

Parameters	**Description**
D	set of train dispatchers, indexed by i
A	set of geographical areas, indexed by j
P	set of time periods, indexed by k
C	set of area combinations, indexed by ℓ
$TL_{j,k}$	task load in area j during period k
$TL^{\max}$	maximum allowed task load
$A^{\max}$	maximum number of assigned areas to a dispatcher per period
$e_{i,j} \in \{0,1\}$	=1 if dispatcher i holds an endorsement for area j
$T^{\min}$	minimum shift length (in time periods)
$T^{\max}$	maximum shift length (in time periods)
$R^{\min}$	minimum number of rest periods between two shifts
$p = \|P\|$	number of time periods in the time horizon

The input to the model is a set A of geographical areas of the railway network. Each area j accommodates train traffic that should be controlled around the clock. During a time period k each area should be controlled by exactly one dispatcher i from the set of dispatchers D. A dispatcher i can be assigned at most one area combination ℓ with the condition that i holds an endorsement for all the areas in ℓ. A dispatcher can simultaneously be assigned an area combination with more than one area if the sum of task load $TL_{j,k}$ over the assigned areas does not exceed $TL^{\max}$. As a simplification, we assume that the task load is given by the number of train movements in each area. The set C of area combinations contains subsets of areas that are combinable, where each subset in C consists of areas that form a connected component in the network's graph. Given the limitation on the number of controllable areas by a single dispatcher during each period, the cardinality of each assigned area combination ℓ should not exceed $A^{\max}$. In our model, we take care of this constraint when we generate the set C. Legal restrictions determine the allowed shift length, which should not be below or above the values $T^{\min}$ and $T^{\max}$, respectively. Moreover, a minimum number of rest periods, $R^{\max}$, is required between two consecutive shifts. The constraints on the shift length and resting periods are modelled based on the strong formulation for min/max on/off sequences presented in Pochet and Wolsey (2006), including a crew usage variable in a similar way as done by Lidén et al. (2018). For the summary of the parameters used in the model, see Table 1.

4.2 Mathematical Formulation

In this section, we present the mathematical formulation of the optimization problem, while in Table 2 we describe the model's variables. The constraints are formulated in Constraints (1)-(15).

Table 2: Model Variables

Variables	Description
$x_{i,j,k} \in \{0,1\}$	=1 if dispatcher i is assigned area j during period k
$c_{i,\ell,k} \in \{0,1\}$	=1 if dispatcher i is assigned area combination ℓ during period k
$y_{i,k} \in \{0,1\}$	=1 if dispatcher i is at work during period k
$v_{i,k} \in \{0,1\}$	=1 if dispatcher i starts a shift at the beginning of period k
$q_i \in \{0,1\}$	=1 if dispatcher i is used during some period

$$\sum_{j \in A} x_{i,j,k} \cdot TL_{j,k} \leq TL^{\max} \qquad \forall i \in D, \forall k \in P \tag{1}$$

$$x_{i,j,k} \leq e_{i,j} \qquad \forall i \in D, \forall j \in A, \forall k \in P \tag{2}$$

$$v_{i,k} \geq y_{i,k} - y_{i,(k-1) \pmod p} \qquad \forall i \in D, \forall k \in P \tag{3}$$

$$v_{i,k} \leq y_{i,k} \qquad \forall i \in D, \forall k \in P \tag{4}$$

$$\sum_{\mu=k+1-T^{\min}}^{k} v_{i,\mu \pmod p} \leq y_{i,k} \qquad \forall i \in D, \forall k \in P \tag{5}$$

$$\sum_{\mu=k+1-T^{\max}}^{k} v_{i,\mu \pmod p} \geq y_{i,k} \qquad \forall i \in D, \forall k \in P \tag{6}$$

$$
\begin{aligned}
v_{i,k} &\leq q_i && \forall i \in D, \forall k \in P && (7)\\
\sum_{k\in P} v_{i,k} &\geq q_i && \forall i \in D && (8)\\
y_{i,k} &\leq \sum_{j\in A} x_{i,j,k} && \forall i \in D, \forall k \in P && (9)\\
y_{i,k} &\geq x_{i,j,k} && \forall i \in D, \forall j \in A, \forall k \in P && (10)\\
\sum_{\mu=k+1}^{k+R^{\min}} v_{i,\mu (\mathrm{mod}\ p)} &\leq q_i - y_{i,k} && \forall i \in D, \forall k \in P && (11)\\
\sum_{i\in D} x_{i,j,k} &= 1 && \forall j \in A, k \in P && (12)\\
x_{i,j,k} &\geq c_{i,\ell,k} && \forall i \in D, \forall k \in P, &&\\
& && \forall \ell \in C, \forall j \in A : j \in \ell && (13)\\
\sum_{\ell\in C} c_{i,\ell,k} &= y_{i,k} && \forall i \in D, \forall k \in P && (14)\\
x_{i,j,k} &\leq 1 - c_{i,\ell,k} && \forall i \in D, \forall \ell \in C, &&\\
& && \forall j \in A \setminus \{\ell\}, \forall k \in P && (15)
\end{aligned}
$$

Constraint (1) prohibits the task load for each dispatcher during each period to not exceed TL^{max}. Constraint (2) allows the assignment of an area only to a dispatcher that holds the corresponding endorsement. Constraints (3) and (4) connect the variables $y_{i,k}$ and $v_{i,k}$. Constraint (5) guarantees the minimum shift length, by assuring that whenever a dispatcher starts a shift at some period, this shift can end only after at least $T^{\min}$ working periods. Similarly, Constraint (6) imposes a limit on the maximum allowed shift length. Constraints (7) and (8) connect the variables $v_{i,k}$ and q_i, by assuring that once a dispatcher i starts a shift at any period then the correspondent q_i is set to 1. Constraints (9) and (10) connect the variables $x_{i,j,k}$ and $y_{i,k}$, such that whenever a dispatcher is at work in a given period then at least one area should be assigned to that dispatcher during that period. At the same time if $y_{i,k}$ is set to 0, i.e., a dispatcher is not working in a given period, then no areas are assigned to that dispatcher during that period. Constraint (11) assures the minimum rest periods between two consecutive shifts by prohibiting any dispatcher from starting a new shift within R^{min} periods after finishing a previous one. Equation (12) makes sure that each area is assigned to exactly one dispatcher during each period. Constraint (13) forces the assignment of all areas within a given subset of areas ℓ at a given period to the same dispatcher if this one is assigned that area combination ℓ during that period, while equation (14) imposes the assignment of exactly one area combination to each working dispatcher at each working period. Finally, Constraint (15) makes sure that each dispatcher during each period is not assigned areas that are not in the subset of the assigned area combination of that dispatcher during that period.

The objective function minimizes the total number of used dispatchers and is given by (16).

$$
\min. \quad \sum_{i\in D} q_i \tag{16}
$$

5 Experimental Study

In this section, we test our IP model using artificial data that has a real-world size as in Malmö dispatching center and its corresponding covered geographical areas. The reason behind using artificial data are confidentiality constraints for accessing and publishing the real data. We run all our experiments on an HP laptop with a processor AMD Ryzen PRO and a 16 GB RAM. We implemented the model using Python 3, and we solved it using Gurobi Optimizer version 9.5.1.

5.1 Base Scenario

We start by generating a set of 15 geographical areas representing the region controlled by a dispatching center. Figure 1 illustrates the layout of these geographical areas.

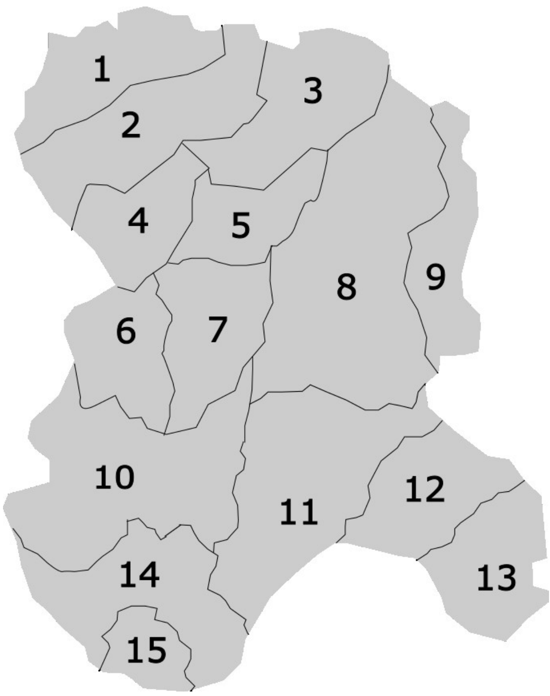

Figure 1: The geographical areas that cover the railway network

An area, depending on its size and complexity, could be a single-track, a double-track or a junction, which we denote as a "complex area". We estimate the task load in the single- and double-tracks as equal to the number of movements per time period present in these areas. In an actual working situation, train movements affect task load in an area depending on its type, i.e., a task load generated by x train movements in single-track area may correspond to the same task load generated by $\alpha \cdot x$ movements in a complex area. This is because the latter one usually is smaller and trains will transit through it in a shorter time compared to a single- and double-track, which gives a higher dispatching task load. In our experiments, we assume the factor α equal to 3. During the night (00-06 a.m), train movements in a single track are at most one per period, yet the task load is still higher because of the infrastructure maintenance that takes place during these periods. To account for these, we estimate two train movements per period as a contribution of maintenance to the task load. During rush hour, the number of train movements is higher than during the rest of the day, which contributes to a higher task load. Based on an interview with experienced train dispatchers from Malmö dispatching center, we generated an estimation of the task load in each area and time period. In Table 3, we present the areas, which type they are, and the sets from which we randomly draw the numerical value of train movements in these areas.

Table 3: Type of areas, and train movements value (randomly chosen from the given set) for different time periods

Type and (list of areas)	**Night time (0-6)**	**Morning rush (6-9)**	**Evening rush (15-20)**	**Day time (9-15 & 20-24)**
Single-track (1,10,11,12,13)	$\{2,3\}$	$\{14,15,16\}$	$\{14,15,16\}$	$\{9,10,11\}$
Double-track (2,5,6,8,9,14,15)	$\{9,10,11\}$	$\{9,10,11\}$	$\{19,20,21\}$	$\{9,10,11\}$
Complex (3,4,7)	$\{10,11,12,13,14\}$	$\{13,14,15,16\}$	$\{14,15,16,17\}$	$\{10,11,12,13,14\}$

Table 4: Parameter values in the base scenario

Parameter	**Value**
$\|A\|$ (number of areas)	15
$\|D\|$ (number of available dispatchers)	22
$T^{\min}$	4
$T^{\max}$	11
$R^{\min}$	11
$TL^{\max}$	30
$\|\ell\| \in C$ (max. size of area combinations)	$<= 3$

For all the experiments, we set the lower and upper bounds for the shift length, $T^{\min}$ and $T^{\max}$, to 4 and 11 periods, respectively, and we set the minimum resting period $R^{\min}$ to 11. For simplicity, we assume that the maximum manageable number of train movements $TL^{\max}$ per each dispatcher is 30 movements per time period, but this parameter could be customized according to each dispatcher's ability and task load tolerance. These parameter values are based on the information from the union agreement and the experts from Trafikverket. Moreover, we consider a time horizon of 24 hours where each unit period is equal to one hour. For the set of available dispatchers, we start with a low value and gradually increase it until obtaining a feasible solution. In the base scenario, we have 22 available dispatchers, where all of them have endorsements for all the areas, but each dispatcher can not control more than three areas at any time period, i.e., the cardinality of any area combination ℓ is at most three. In Table 4, we present the values of the parameters used in the base scenario.

	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23
D1												5 8 9	6 10	11 12 13	2 3	1 2	12 13	9	15	3 4	6	11 12 13		
D2	15	10	4 6	3	11 14 15	5 8 9	6 7	6 7	11 12	11 12 13	8 9													
D3														4	4 5	8	14	4 7	9	12 13	12 13	3	3 5	13
D4		3 5	1 2 3	11 12 13	12 13	1 2 3	1 2	1 2	3 4	1 2	1 2 3													
D5				10 14 15	1 2 3	4 6	10 11	10 11	14 15	6	11 12 13	10 14	1 2	15										
D6	5	4														7	5	14	5	5	7 10	6	7	1 2 3
D7	3	1	9	5	5 8 9	7	12 13	3 4														10	6 8 10	12
D8						11 12 13	14 15	12 13	1 2	14 15	14	1 2	8 9	1	1									
D9						10 14	5 8 9	14 15	5 8 9	7 8 9	6 7	3 4	4 7	8 9	6 8 10	12 13								
D10									6	10	10	6 7	12 13	2 3	11 14 15	14	4 7	8	11					
D11	6 10 14	6 7	11 12 13	1 2 4															12 13	14	3 4	1 2	14 15	10
D12	12	14 15	5 7 10														3	10 11	6	15	11	4	13	8 11
D13																15	2	15	3	9	5 8	8	1 2	7
D14	1 2 4	2														5	6	1	10	8	9	7	11 12	4 5 6
D15									13	4 5	4 5	11 12 13	5	5	7	9	15	6	14					
D16	13															3 4	1	5	2	10	15	14 15	9	9
D17	7	11 12 13	8	6 7	4 6	15	3 4	5 8 9	7 10	3	15													
D18												15	11 14 15	7	12 13	10 11	8	3	8	1 2	14	9		
D19																	9	12 13	7	11				
D20													3	6 10 14	9	6	10 11	2	4	7				
D21	8 9 11	8 9	14 15	8 9	7 10														1	6	1 2	5	4	14 15

Figure 2: An optimum one-day shift schedule of the base scenario, where each row represents a shift for a dispatcher and the columns represent the time periods (from 00 to 23). The numbers in the cells show the assigned areas per dispatcher and working period

The result from this run is a shift schedule with the minimum number of dispatchers, that determines the shift's start and end for each one of the used dispatchers. We show the resulting schedule in Figure 2, where 21 dispatchers were used to control just up to three areas each. All dispatchers but 3, 13, 16 and 19 were assigned three areas during some period. As expected, the three complex areas 3, 4 and 7 are never simultaneously assigned to any dispatcher during rush hour since the total task load in that case would be greater than $TL^{\max}$. Most of the shifts has the maximum allowed length, while the shortest one is four hours long. In our output schedule, dispatchers often work with one set of areas in one period and with a complete different set in the consecutive period—they *switch* the areas assigned to them. As we do not steer towards fewer area switches, this in an expected behavior. However, we will integrate constraints on more practicable shifts in future work.

To evaluate our results and to be able to compare the different experiments, we looked at the following metrics: total number of used dispatchers; minimum, maximum and average shift length; average number of assigned areas per working period for all dispatchers; and the run time. The values of these metrics for the base scenario are presented in Table 5.

Table 5: Results for the base scenario

Metric	**Units**	**Notation**	**Value**
Total number of used dispatchers		$\boldsymbol{T}$	21
Minimum shift length	Hours	$\boldsymbol{m}$	4
Maximum shift length	Hours	$\boldsymbol{M}$	11
Average shift length	Hours	$\overline{\boldsymbol{L}}$	10.23
Average number of assigned areas		$\overline{\boldsymbol{A}}$	1.67
Run time	Seconds	$\boldsymbol{R}$	57

5.2 Changing the Endorsement Ratio

In this experiment, we change the ratio of areas for which a dispatcher holds an endorsement. We assume the same ratio for all dispatchers, and we ensure different subset of areas for each dispatcher (such that the endorsements of all dispatchers cover all areas). In the first instance, we set the endorsement ratio to 1/2 and we call this $E_{1/2}$, while in the second instance we decrease the ratio to 1/3 and we call this $E_{1/3}$.

In Table 6, we present the results of the instances $E_{1/2}$ and $E_{1/3}$, together with results from other experiments. To compare the results from different instances we use the same metrics as in Table 5.

In instance $E_{1/2}$, decreasing the endorsement ratio by one-half does not affect the objective function value, which is 21 used dispatchers, while in $E_{1/3}$ 22 dispatchers are needed. The minimum and average shift length increases in both instances, while the run time decreases by around 65%. This decrease could be explained by the search space reduction due to the endorsement limitation. For the same reason the average number of assigned areas decreases to 1.49 in $E_{1/3}$, while it remains unchanged in $E_{1/2}$.

5.3 Changing the Maximum Cardinality of Area Combinations

This experiment is a set of three instances, $\{M_4, M_2, M_1\}$, where in each instance we set the maximum cardinality of area combinations to four, two and one, respectively. The corresponding results in Table 6 show that the total number of used dispatchers does not change when we allow area combinations of cardinality four (M_4). A reason for this is that usually most of the generated area combinations of cardinality four would have a total task load greater than $TL^{\max}$, thus, they cannot be part of a feasible solution. When we restrict the cardinality to two (M_2), the number of used dispatchers increases only by one, compared to the base scenario. This small effect could be attributed to the fact that some few area combinations became infeasible while initially they were feasible. The effect becomes clearly observable when we set the maximum cardinality to one (M_1), which results in 33 used dispatchers.

Comparing the run time within this experiment's instances, including base scenario, shows that higher values for the cardinality of area combinations give an increase in run time. This observation is reasonable since decreasing the number of generated area combinations would decrease the number of variables reducing the search space and, thus, reducing the run time. We also observe that the average number of assigned areas increases when the maximum cardinality of area combinations is higher. This relation is not linear since when we increment the cardinality from one to two we get highest increase in the average of assigned areas, while the increase is lower when we further increment the number of allowed area combinations. A reason for this is that furtherly increasing the cardinality does not necessarily mean that the high-cardinality area combinations could be exploited, since the task load of these combinations would more likely exceed $TL^{\max}$, and, therefore, they would not be used in a feasible solution.

We observe that the average shift length decreases when the cardinality of area combinations increases. The highest average shift length, as expected, is obtained when dispatchers could not be assigned more than one area at a time. This is because each working dispatcher is used as long as possible before using an extra one, which can be seen from the value of the minimum shift length in instance M_1. For the maximum shift length, we observe that it was 11 for all the experiments presented in Table 6, which is equal to the upper bound $T^{\max}$. This also makes sense, given that the objective function is to minimize the number of used dispatchers, and once a dispatcher is used then they are exploited as much as possible, leading to long shifts, before needing more dispatchers. For the minimum shift length, we could not see any connection with the maximum cardinality of area combinations, which was also expected since the minimum shift length cannot be controlled, even indirectly, by the cardinality of area combinations. We can confirm this by considering the shift schedule, presented in Figure 2, where the minimum shift length was four hours (see dispatcher D19). Having the same area combinations as in this solution, we can easily manipulate the shifts and get another solution with the same objective value. An example of this manipulation

Table 6: Results for changing the endorsement ratio, and the maximum cardinality of area combinations

Instance	T	m	M	$\bar{L}$	$\bar{A}$	R
Base	21	4	11	10.23	1.67	57
$E_{1/2}$	21	9	11	10.86	1.67	20
$E_{1/3}$	22	10	11	10.95	1.49	19
M_4	21	5	11	10.09	1.7	97
M_2	22	4	11	10.54	1.55	39
M_1	33	10	11	10.91	1	29

is to move parts of the shift assigned to dispatcher D18 to D19, which would make D18 having the minimum shift length of six hours. Figure 3 shows how this shift manipulation could be done.

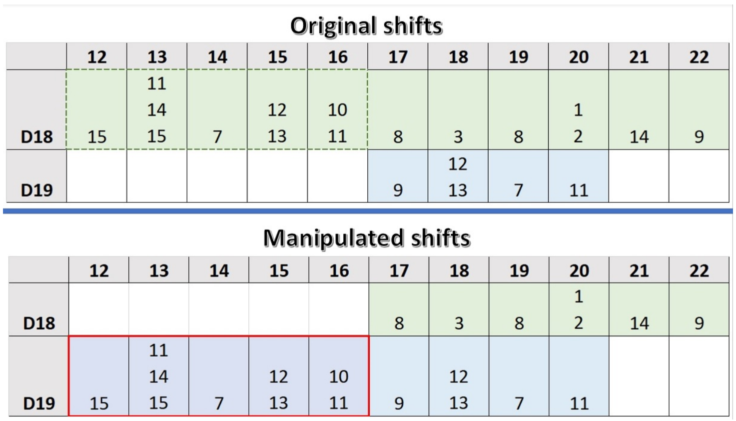

Figure 3: An example of two manipulated shifts with different minimum shift length but with the same objective value (number of used dispatchers)

5.4 Changing the Total Number of Areas

In this subsection, we consider two experiments with three instances each. In each of the experiments, we gradually decrease the total number of areas, starting from the base scenario with 15 areas, by removing an area in each instance. In the first experiment, we target areas with a high degree in the area adjacency graph, where we consecutively remove areas 10, 7, and 5, to obtain the instances A_{10}, $A_{7,10}$, and $A_{5,7,10}$, respectively. In the second experiment, we target areas with a low degree in the same adjacency graph, and we consecutively remove areas 13, 15 and 14. We call these instances A_{13}, $A_{13,15}$ and $A_{13,14,15}$, respectively.

In both experiments (see Table 7), we observe a decrease in the number of used dispatchers when we reduce the total number of areas, which is expected since the more areas we have to cover, the more likely we will need additional dispatchers. While we can see an increase in the minimum shift length in the first experiment as the number of areas decreases, this trend does not hold in the second experiment. Thus, this confirms our conclusion that the number of areas has no impact on the minimum shift length. We draw the same conclusion for both the maximum and the average shift length.

The average number of assigned areas decreases when the total number of areas is reduced in the first experiment, but this was not the case in the second experiment, which means that we cannot observe any connection between the average number of assigned areas and the number of areas.

Examining the computation performance, we can see that in the first experiment the run time decreases as the total number of areas gets smaller, which seems reasonable. However, in the second experiment, this trend surprisingly does not hold. Indeed, the run time increases notably in instance A_{13} compared to the base scenario, while another smaller increase in run time occurs between instances $A_{13,15}$ and $A_{13,14,15}$ when we furtherly remove area 14. A plausible explanation for the increase in run time in A_{13} could be that area 13 can be assigned to a dispatcher either alone, together with 12 or as a triple with 11 and 12, where the last two cases would usually appear in an optimum solution due to their efficiency. This remark is confirmed by the optimum solution presented in Figure 2, where area 13 was assigned alone to a dispatcher during only 17% of the time periods, and that for the remaining of the time, 13 was only combined with 12 or with 12 and 11 together. These combinations dominated in the optimum solution, while alternative combinations that include areas 11, 12 or 13, were not interesting for further evaluations. When we removed area 13, many more new possible solutions needed to be evaluated, which could have led to a higher run time. The reason behind designing this last experiment was to show that the properties of the areas, in this case the degree in the area adjacency graph, may affect the run time more than the total number of areas.

Table 7: Results for changing the total number of areas

Instance	T	m	M	$\overline{L}$	$\overline{A}$	R
Base	21	4	11	10.23	1.67	57
A_{10}	20	8	11	10.05	1.6	39
$A_{7,10}$	19	10	11	10.95	1.5	25
$A_{5,7,10}$	18	11	11	11.00	1.45	21
A_{13}	20	8	11	10.55	1.59	305
$A_{13,15}$	18	5	11	10.05	1.72	245
$A_{13,14,15}$	17	5	11	10.58	1.6	252

6 Conclusions and Future Work

In this paper, we formulated an IP for automating shift scheduling for train dispatchers. Apart from legal and operational constraints that are similar to shift scheduling problems for many other professions, we integrated coverage of connected geometric areas. In this first work on train-dispatchers shift scheduling, we minimize the number of dispatchers to create a baseline for future optimization using our model. Moreover, we focus on one-day shifts.

We ran experiments on real-world sized instances. The run time for different instances was between 19 and 305 seconds, which is fast enough for tactical planning. Moreover, we tested the model on real-world instances provided by Trafikverket and we got similar run times. Unfortunately, these results could not be published due to the mentioned confidentiality constraints.

To gauge the behavior of our model, we ran three sets of experiments: varying the ratio of areas for which each dispatcher is qualified, the upper bound on the number of areas that can be simultaneously monitored by one dispatcher, and the total number of areas in the instance. Maybe surprisingly, decreasing the number of areas does not always reduce the problem complexity. We highlight a dependence on the degree of the removed areas in the area-adjacency graph: deleting high-degree areas decreases the complexity, deleting low-degree areas may not do so.

The average and minimum shift length are not clearly connected to the parameters of the input—the model only enforces an upper and a lower bound. Hence, if some fairness between shifts is desirable, this needs to be integrated explicitly.

In future work, we aim to investigate further objectives (e.g., minimizing the number of areas assigned to a single dispatcher, minimizing the number of switches between areas in a shift). We also plan to expand the time frame to several weeks—which includes new constraints on the maximum working time during a week depending on the time of day at which the work is performed. Additionally, we intend to adapt our model for ad hoc rescheduling in case of some expected but stochastic events, such as weather or unplanned track repair.

Acknowledgements

This work is part of the Trafikverket-funded project Capacity Modeling and Shift Optimization for Train Dispatchers (CAPMO-Train) (TRV 2020/99267). We would like to thank Magnus C. Johansson, Carl-Johan Wallentin, and Mariella Arcelay from Trafikverket, TC Malmö, for their time and expertise in mapping the constraints on train-dispatcher shifts.

References

Ahasan, R. (2002). "Human adaptation to shift work in improving health, safety and productivity–some recommendations". In: *Work study* 51.1, pp. 9–16.

Akbari, M., Zandieh, M., and Dorri, B. (2013). "Scheduling part-time and mixed-skilled workers to maximize employee satisfaction". In: *The International Journal of Advanced Manufacturing Technology* 64.5, pp. 1017–1027.

Alfares, H. K. (2007). "Operator staffing and scheduling for an IT-help call centre". In: *European Journal of Industrial Engineering* 1.4, pp. 414–430.

Baker, K. R. (1976). "Workforce allocation in cyclical scheduling problems: A survey". In: *Journal of the Operational Research Society* 27.1, pp. 155–167.

Bard, J. F., Binici, C., and Desilva, A. H. (2003). "Staff scheduling at the United States postal service". In: *Computers & Operations Research* 30.5, pp. 745–771.

Bard, J. F. and Purnomo, H. W. (2005). "Preference scheduling for nurses using column generation". In: *European Journal of Operational Research* 164.2, pp. 510–534.

Barnett, R. C. (2006). "Relationship of the number and distribution of work hours to health and quality-of-life (QOL) outcomes". In: *Employee health, coping and methodologies*. Emerald Group Publishing Limited.

Bechtold, S. E. and Jacobs, L. W. (1990). "Implicit modeling of flexible break assignments in optimal shift scheduling". In: *Management Science* 36.11, pp. 1339–1351.

Breugem, T., Dollevoet, T., and Huisman, D. (2022). "Is equality always desirable? Analyzing the trade-off between fairness and attractiveness in crew rostering". In: *Management Science* 68.4, pp. 2619–2641.

Cacchiani, V. and Toth, P. (2018). "Robust train timetabling". In: *Handbook of Optimization in the Railway Industry*. Springer, pp. 93–115.

Colligan, M. J. and Rosa, R. R. (1990). "Shiftwork effects on social and family life." In: *Occupational medicine (Philadelphia, Pa.)* 5.2, pp. 315–322.

Cordeau, J.-F. et al. (2010). "Scheduling technicians and tasks in a telecommunications company". In: *Journal of Scheduling* 13.4, pp. 393–409.

Cunha, L. et al. (2020). “Do we want to keep working in 12-h shifts? The follow-up of the work schedule change in a Portuguese industrial company”. In: *International Journal of Industrial Ergonomics* 77, p. 102958.

Dantzig, G. B. (1954). “A comment on Edie’s “Traffic delays at toll booths””. In: *Journal of the Operations Research Society of America* 2.3, pp. 339–341.

Dima, M., Chihaia, I. A., and Surugiu, M. C. (2017). “Communications in railway centralized traffic control systems”. In: *IEEE Proc. Of 8th International Conference on Electronics, Computers and Artificial Intelligence*.

EUROSTAT (2022a). *Passengers transported in Sweden 2004-2019*. URL: https://ec.europa.eu/eurostat/databrowser/view/RAIL_PA_QUARTAL__custom_2820826/default/line?lang=en (visited on 06/21/2022).

— (2022b). *Train movements in Sweden 2004-2019*. URL: https://ec.europa.eu/eurostat/databrowser/view/RAIL_TF_TRAINMV__custom_2823628/default/line?lang=en (visited on 06/21/2022).

Frey, L., Hanne, T., and Dornberger, R. (2009). “Optimizing staff rosters for emergency shifts for doctors”. In: *2009 IEEE Congress on Evolutionary Computation*. IEEE, pp. 2540–2546.

Gertler, J. (2003). *Selection of railroad dispatcher candidates*. Tech. rep.

Huang, L., Cummings, M., and Nneji, V. C. (2018). “Preliminary analysis and simulation of railroad dispatcher workload”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 62. 1. SAGE Publications Sage CA: Los Angeles, CA, pp. 691–695.

Hur, Y. et al. (2019). “A stochastic optimization approach to shift scheduling with breaks adjustments”. In: *Computers & Operations Research* 107, pp. 127–139.

Josefsson, B. et al. (2017). “Scheduling air traffic controllers at the remote tower center”. In: *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*. IEEE, pp. 1–10.

Kim, K. and Mehrotra, S. (2015). “A two-stage stochastic integer programming approach to integrated staffing and scheduling with application to nurse management”. In: *Operations Research* 63.6, pp. 1431–1451.

Lapègue, T., Bellenguez-Morineau, O., and Prot, D. (2013). “A constraint-based approach for the shift design personnel task scheduling problem with equity”. In: *Computers & Operations Research* 40.10, pp. 2450–2465.

Lidén, T., Kalinowski, T., and Waterer, H. (2018). “Resource considerations for integrated planning of railway traffic and maintenance windows”. In: *Journal of rail transport planning & management* 8.1, pp. 1–15.

Marinov, M. et al. (2013). “Railway operations, time-tabling and control”. In: *Research in Transportation Economics* 41.1, pp. 59–75.

Moen, P. and Yu, Y. (2000). “Effective work/life strategies: Working couples, work conditions, gender, and life quality”. In: *Social problems* 47.3, pp. 291–326.

Mücke, W. (2002). *Operations control systems in public transport*. Tetzlaff.

NTSB (1998). *National Transportation Safety Board: Railroad Accident Report*. Tech. rep. URL: https://static1.squarespace.com/static/505b96a8c4aa40a37a143c49/t/5a3664ad652dea7bc2028431/1513514160179/1997+-+jun+22+-+Devine+TX+-+RAR9802.pdf (visited on 08/22/2022).

Palmqvist, C.-W., Lind, A., and Ahlqvist, V. (2022). “How and Why Freight Trains Deviate from the Timetable: Evidence from Sweden”. In: *IEEE Open Journal of Intelligent Transportation Systems* 3, pp. 210–221.

Pochet, Y. and Wolsey, L. A. (2006). *Production planning by mixed integer programming*. Vol. 149. 2. Springer, pp. 341–343.

Reinach, S. J. (2006). “Toward the development of a performance model of railroad dispatching”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 50. 17. SAGE Publications Sage CA: Los Angeles, CA, pp. 2042–2046.

Törnquist, J. (2006). “Computer-based decision support for railway traffic scheduling and dispatching: A review of models and algorithms”. In: *5th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS’05)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

Trafikverket (2019). *Kollektivavtal: Trafikverkets affärsverksavtal mellan Trafikverket och Saco-S, OFR-S och Seko*. URL: https://www.seko.se/siteassets/forhandlings-branschorganisationer/seko-klubb-trafikverket/pdf-er/test/trafikverkets-affarsverksavtal_andringar_tillagg_till_och_med_2019_10_01_20191025.pdf (visited on 06/21/2022).

Wright, D., Bretthauer, K., and Cote, M. (2006). “Reexamining the nurse scheduling problem: Staffing ratios and nursing shortages”. In: *Decision Sciences* 37.1, pp. 39–70.

Wright, D. and Mahar, S. (2013). “Centralized nurse scheduling to simultaneously improve schedule cost and nurse satisfaction”. In: *Omega* 41.6, pp. 1042–1052.

Al-Yakoob, S. M. and Sherali, H. D. (2008). “A column generation approach for an employee scheduling problem with multiple shifts and work locations”. In: *Journal of the Operational Research Society* 59.1, pp. 34–43.

Zeilstra, M., Wincoop, A. v., and Rypkema, J. (2017). “The WASCAL-Tool: Prediction of Staffing for Train Dispatching as Part of the Design Process of Track Yards”. In: *International Symposium on Human Mental Workload: Models and Applications*. Springer, pp. 143–160.

Zhong, X., Zhang, J., and Zhang, X. (2017). “A two-stage heuristic algorithm for the nurse scheduling problem with fairness objective on weekend workload under different shift designs”. In: *IISE transactions on healthcare systems engineering* 7.4, pp. 224–235.