

## TRANSLATION

# Automatic Generation Method for Crew Scheduling After a Large-Scale Natural Disaster

Satoshi Kato<sup>1</sup> | Jun Imaizumi<sup>2</sup> | Taichi Nakahigashi<sup>1</sup>

<sup>1</sup>Railway Technical Research Institute, Kokubunji, Tokyo, Japan | <sup>2</sup>Toyo University, Bunkyo-ku, Tokyo, Japan

**Correspondence:** Satoshi Kato ([kato.satoshi.58@rtri.or.jp](mailto:kato.satoshi.58@rtri.or.jp))

**Keywords:** column generation | crew schedule | disaster | mathematical optimization | temporary duty | temporary timetable

## ABSTRACT

When railway line sections are partially disrupted due to damage from large-scale natural disasters, it is necessary to prepare crew schedule plans for temporary timetables. In such a case, a crew schedule plan ensures that duties once assigned to crew members are not changed. As this task is time-consuming for schedule planners, an automatic crew scheduling method for supporting planners is desired. In this paper, we focus on crew scheduling following a large-scale natural disaster. The proposed algorithm is based on mathematical programming and column generation. In addition, we show the results of computational experiments based on a real disaster case, which indicate that the proposed algorithm can generate an efficient schedule plan in a short time.

## 1 | Introduction

For railway operators, timetabling and planning are essential practical tasks. For example, rolling stock and crews are not only necessary but also finite resources to be planned reasonably and efficiently. In particular, crews (drivers and conductors) are precious resources in scarce supply, and crew scheduling is an important issue.

On the other hand, the determination of daily train timetables often requires not only planning but also a temporary response, such as rescheduling train services when timetables are disrupted. Especially in Japan, a country with geographical conditions that make natural disasters unavoidable, railway lines may be cut off depending on the scale of the disaster and the scope of recovery. When the service recovery takes a considerable amount of time, temporary operations are provided where possible; however, such timetables and plans differ from regular operations in prerequisites and working environment. This involves various administrative tasks; considering the recent increase in the frequency and severity of natural disasters, such

tasks are expected to occur often, and methods and means of support should be available.

## 2 | Research Background and Approach

### 2.1 | Crew Schedules and Actual Operation

To operate a train service, one needs not only timetables but also a transportation plan that includes a rolling stock operation plan and a crew schedule based on the timetables. The transportation plan consists of a “basic plan” created with regard to repeated weekday and weekend operations and an “implementation plan” that reflects extra trains operated on certain days or periods and other changes to the basic plan. The basic plan is created when the timetables are revised; in so doing, the basic plan is modified according to the implementation plan [1].

Crew schedules are first prepared for the basic plan to become the basis for the actual planning of crew duties through the implementation plan. Thus, efficient plans must be prepared

Translated from Volume 145 Number 5, pages 343–353, DOI: [10.1541/ieejias.145.343](https://doi.org/10.1541/ieejias.145.343) of *IEEJ Transactions on Industry Applications* (Denki Gakkai Ronbunshi D).

for timetable revisions to meet various constraints based on labor rules and other regulations, while ensuring effective use of limited resources.

However, crew schedules are currently created manually by experienced personnel. When framing basic plans for timetable revisions, an operator has some time allowance of several days to several weeks per line; however, there are problems, including reducing the burden on the personnel and passing on the skills when the veteran planners retire. In this context, many operators need support based on mathematical methods, and research has been done on automated crew scheduling using mathematical optimization and other information technologies.

Based on such crew schedules, task planners draw up monthly attendance tables with regard to the wishes of every crew member and other factors. At the end of each month, every crew member is notified of their actual work schedule for the following month. While being based on the crew schedule of the basic plan, additional factors (changes reflected in the implementation plan, tasks other than crew duties, etc.) are taken into account in determining the actual work of every crew member.

On the other hand, in the event of a disaster, operation schedules must be revised on the assumption that the crew members have been informed of their attendance tables. This will be explained in more detail in the following subsection.

## 2.2 | Difference Between Emergency Planning and Basic Plan

Due to global warming, Japan has been frequently affected by typhoons and floods recently; in combination with earthquakes, these disasters often hit railway lines. When railway facilities are damaged, train service may be suspended over long sections for long periods of time. In such cases, operations are usually resumed gradually as restoration works proceed, and transportation plans are created accordingly.

As a matter of course, post-disaster scheduling must be completed by the day when a service is resumed (or line sections are extended). However, like a basic plan, a transportation plan includes sequentially drawing up timetables, rolling stock schedules, and crew schedules; the latter often must be done within an extremely limited time frame. This poses even stricter time constraints after disasters. This results in a very heavy workload, including overtime during night hours and bringing in batteries of manpower; moreover, operational sections can be extended repeatedly.

On the other hand, there are some differences between basic plans and emergency plans.

When formulating a basic plan, it is necessary to secure crew members with regard to the staffing plan, seasonal fluctuations of passenger traffic, etc.; as a result, no excessive surplus or shortage of crew members occurs when attendance tables are created based on crew schedules. However, the situation after a disaster is often different from normal assumptions, and scheduling must be done under a surplus or shortage of crew members. In each case, the

content and objectives of planning are different, and automated schedule generation should support such differences.

Naturally, the number of trains in service, their departure and arrival times, and line sections are usually different from those during normal service. The crew's duties are determined accordingly, but not only will these duties differ from normal operations (the basic plan), but they must also be consistent with normal operation schedules. Specifically, the circumstances of crew members have to be taken into account.

For example, when a disaster occurs after all crew members were notified of their attendance tables for the current month, and some sections are closed or some trains are suspended, this results in changes in the crew's schedules and tasks, and days off may be canceled. However, such sudden changes should be avoided as much as possible to prevent a burden on the crew members involved. In addition, it is desirable to avoid substantial changes in work start and end times. For that reason, some measures must be provided in post-disaster operation plans that must be revised with respect to the basic plans.

Therefore, there are the following differences.

- Sustentative requirements may be different from the basic plan on a case-by-case basis.
- Scheduling must be done with respect to the existing basic plan and corresponding work schedules and tasks.

Thus, there is a need for an automatic schedule generation method different from crew scheduling used in many previous studies for basic plans.

## 2.3 | Approach of this Study

In the above context, this paper focuses on automatic generation of crew schedules after a disaster, with the aim of labor-saving and deskilling.

Specifically, we consider drawing up a “temporary plan” with respect to the basic plan, including “temporary duties” corresponding to “temporary timetables” in the event of a disaster; these temporary duties correspond to “basic duties” that express a work unit (departure from and return to home base) of the basic plan.

Temporary duties are designed as closely as possible to the corresponding basic duties in terms of work start and end times. This is to prevent changes in crew attendance tables in the event of a disaster. At the same time, start and end times of every duty become a constraint on duty sequence (“roster”: to be explained later) underlying the content of work specified for crew members in attendance sheets, which provides some benefits, including the following.

- If the content of a roster (duties and their sequence designed in the basic plan) can be changed by simply replacing affected duties, then the other duties (not requiring revision) can be reused as they are.

- Even if all duties need to be revised, work and off days in a roster will remain unchanged after replacing the original duties with temporary duties.

As a result, there is no need to change attendance tables that have already been distributed, which greatly reduces the burden on both planners and crew.

After introducing temporary duties, crew schedules must be basically as efficient as in normal operation, and the primary goal is to create schedules under a limited number of crew members. However, because disaster situations are different from normal operations, it is unclear whether operation schedules created in this way are adequate. For example, if the number of trains on the temporary timetable is considerably fewer than in regular operation and the only goal is to minimize the crew size, then even if temporary duties meet the roster requirements, there is a possibility that the revised roster will be significantly different from normal operation; such situations should be avoided.

In this study, we use mathematical optimization and have practitioners evaluate schedules generated automatically; the evaluation results are used to refine the method. That is, first, evaluation indicators are set, and a schedule is generated using the same approach as for the basic plan; after that, the method is revised based on the evaluation results and opinions of the practitioners.

### 3 | Crew Scheduling and Requirements for Automatic Generation

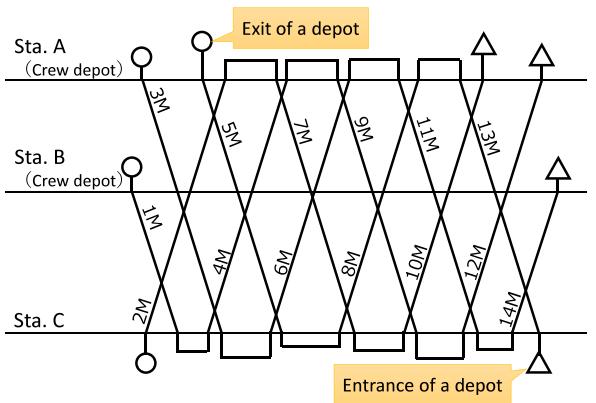
#### 3.1 | Crew Scheduling: Terms and Definitions

Crew scheduling includes “duties” and “rosters”. The following explanations pertain to both basic and temporary plans.

##### 3.1.1 | Duties

A duty is the work of a crew member from clock-in to clock-out. Every crew member is assigned to a crew depot located near several stations on a railway line; duties are designed so that crew members begin their work at the station nearest to their depot and return to the depot after working on several trains. There are “day duties” completed in one calendar day and “overnight duties” that span two calendar days. Usually, most trains are operated during the morning rush hours, and overnight duties outnumber day duties.

Duties are subject to multiple constraints on labor time, crew service time, etc., based on labor rules, crew roster rules, and other regulations; all such constraints must be satisfied when designing duties. Taking into account these various conditions, it is necessary to create several duties for each crew depot so that one or more crew members can be assigned to every section and every task on the timetable.



**FIGURE 1** | An example of a train timetable and rolling stock schedule. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

##### 3.1.2 | Rosters

After the duties have been created, a duty assignment sequence is determined for each crew depot to create a schedule of duties for several days to several weeks. This schedule is called a “roster”. Various conditions must be met when designing rosters, for example:

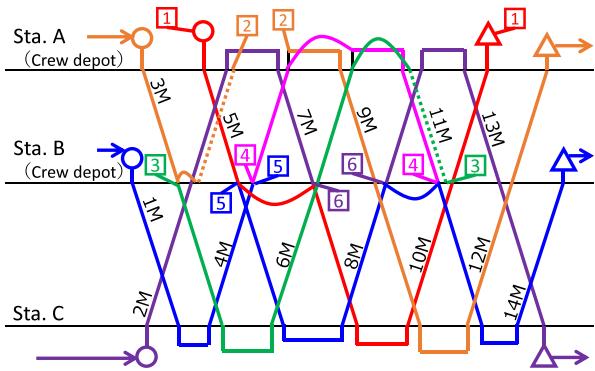
- Adequate rests (below, referred to as “home rests”) must be provided between duties depending on the content of the previous duty.
- Rosters must include appropriate days off.

Actual tasks of each crew member are notified to them in the form of attendance tables created monthly.

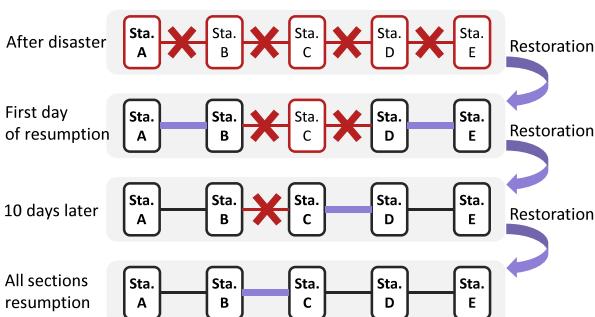
##### 3.1.3 | Examples of Duties

Here, the above explanations are illustrated by examples. Crew schedules are based primarily on timetables. However, rolling stock schedules also serve as reference information required for the convenience of discussing mathematical solutions. Thus, our explanations assume that rolling stock schedules are available as well.

Figure 1 shows an example of a train timetable and rolling stock schedule; Figure 2 shows the corresponding duties. There are a total of 14 trains in this example; in Figure 1, the circles and triangles show, respectively, the exit from and entrance to a depot. Stations A and B are the nearest stations to crew bases; treating the stations as Depots A and B, two duties 1 and 2 are designed for Depot A, and four duties 3–6 are designed for Depot B (Figure 2). Duties 1, 3, and 4 are day duties, while the duties 2, 5, and 6 are overnight duties. These six duties are allocated to all trains and all sections. The dotted lines in duties 2 and 3 pertain to travels not involving crew service (passive travels, or “deadheadings”). Deadheading is considered waste and, therefore, should be minimized.



**FIGURE 2** | An example of crew duties. [Color figure can be viewed at wileyonlinelibrary.com]



**FIGURE 3** | An example of the extension of operational sections. [Color figure can be viewed at wileyonlinelibrary.com]

### 3.2 | Requirements for Crew Scheduling After Natural Disaster

In order to extract the requirements for crew scheduling after a disaster, we conducted a hearing survey with practitioners who had experience in transportation planning after large-scale disasters. In addition, we analyzed actual operation schedules in the event of a disaster. As a result, the following requirements were extracted for post-disaster crew scheduling.

#### 3.2.1 | Plans Must be Created Each Time Operational Sections Are Extended

In the event of a natural disaster, works are carried out to restore the affected areas, and operations are resumed as the restoration works are completed. Thus, when multiple locations are affected, operational sections may be extended gradually as restoration works proceed, and transportation plans must be created at every phase of extension.

An example of the extension of operational sections is shown in Figure 3. In this example, all sections between Stations A and E are affected. First, operations are resumed between Stations A and B and between Stations D and E; then, 10 days later, operations are resumed between Stations C and D. In this example, sections are extended in three phases, but the number of phases can be greater depending on the recovery conditions. Moreover, if the crew depot is near Station A, the crew members must be sent from

Station A to Stations D and E on the first day of service resumption and Stations C and E ten days later. Such “isolated sections” may occur in real life.

#### 3.2.2 | Post-Disaster Plans Are Created in Line with Basic Plan

As mentioned in Subsection 2.1, crew attendance tables for the following month are distributed at the end of each month; that is, when a disaster occurs, the work of every crew member (work and off days, clock-in and clock-out times) for that month has already been determined. Therefore, if both duties and rosters are created in the same way as in the basic plan, the attendance tables must be redesigned. This not only puts a burden on the planners but also creates another burden on the crew members who have to come to work on designated off days, which should be avoided.

Thus, temporary duties are designed in correspondence with the basic duties. With the method proposed in this study, temporary duties are designed to prevent disadvantages for the crew, even though clock-in and clock-out times are changed from the original schedule; as a result, there is no need to redesign rosters, and the work and off days remain unchanged. Specifically:

- If a basic duty is an overnight duty, the temporary duty is created as an overnight duty or a day duty. In the latter case, the second day of the crew in charge is treated as “duty suspension”, and actual work assignments are determined in each case (off day, replacement reserve, etc.).
- If a basic duty is a day duty, the temporary duty is created as a day duty; alternatively, no temporary duty may be created (all trains are covered by other temporary duties). In the latter case, that day is treated as “duty suspension” for the crew assigned to that duty on some day.

In so doing, there is no need to redesign rosters, and the work on off days is avoided.

#### 3.2.3 | Some Work Conditions May be Relaxed

When preparing crew schedules, it is necessary to take into account many work conditions based on various regulations. However, after a disaster occurs, the number of trains often decreases compared to the basic plan because operations are suspended in some sections. In so doing, some work conditions may be relaxed. Based on the results of the hearing survey, we concluded that there is no need to secure an average labor time per day.

#### 3.2.4 | Duties Must be Designed to Achieve a Certain Level of Efficiency and Workload Within Limited Labor Time

Crew scheduling is performed after temporary timetables and rolling stock schedules have been fixed; therefore, there is often a limited amount of time before the temporary timetables are implemented, and the crew schedules must be prepared in a

very short time. In addition, the same schedule is implemented for several days to several weeks until the next phase of section extension; thus, duties must be designed with a certain level of efficiency and workload.

### 3.2.5 | Special Deadheadings Might be Required

When isolated sections occur, connection to a crew depot is disrupted, and the crew cannot make it by train; therefore, another means of transportation is needed.

Specifically, crew members have to use a bullet train if available, taxis, or other specific transport, which is hardly included in basic plans. When using such means of transportation, it is necessary to allow sufficient time to avoid being late to work.

## 3.3 | Evaluation Indicators and Constraints

There are a variety of evaluation indicators for crew scheduling, but the metrics considered the most important, for both basic plans and post-disaster plans, are “efficiency” (total number of crew members) and “number of deadheadings”; the smaller these indicators, the better. Even with disaster plans, both indicators are standard metrics of the planning quality. On the other hand, when scheduling is treated as a mathematical optimization problem, it is necessary to define an “objective function”. When there are multiple metrics, one can think of defining an objective function as a weighted sum of these metrics; however, such an approach involves some problems, such as the determination of weight factors, and it is easier to deal with an objective function based on a single metric or to add another metric if there is no other way.

In this study, we define an objective function using “number of working days” (day duty: 1, overnight duty: 2) and “number of deadheadings”; the latter is provided with an appropriate weight factor.

Tentatively, only these metrics are considered in optimization. However, it is unclear how special circumstances following a disaster will affect the proposed method and the generated set of duties; thus, we decided to have practitioners evaluate the results for further consideration.

The constraints on scheduling are presented below.

- a. Temporary duties are assigned to all trains, all sections, and all tasks on the temporary timetable.
- b. Each temporary duty begins at a crew depot (actually, at the adjacent station) and ends at the same crew depot.
- c. Each temporary duty satisfies an upper limit on the duty time, being different between day duties and overnight duties. The duty time is the time from the start to the end of a duty.
- d. Each temporary duty satisfies an upper limit on the labor time. The labor time includes preparation time in addition to the operation time.

- e. Each temporary duty satisfies a lower limit on the rest time. The rest time is defined as the duty time less the labor time. This limit varies depending on the labor time (6 to 8 h or more than 8 h). This constraint applies only to day duties. In the case of overnight duties, this lower limit on the rest time is satisfied automatically during sleep.
- f. Necessary meal times (breakfast, lunch, dinner) must be provided in each temporary duty. This is formulated as “providing sufficient periods of time during certain hours”. For example, at least 50 min must be provided between 11:00 and 14:00, after the arrival of a train and the departure of the next train to be serviced. The necessity of each meal time is determined separately for day duties and overnight duties.
- g. There is a crew service zone for each crew depot.
- h. In the case of overnight duties, the necessary amount of sleep must be ensured between the arrival of the last train on the first day and the departure of the first train on the second day.
- i. An upper limit on the continuous crew service time must be satisfied. The continuous crew service time is defined as the time of continuous crew service without a break.
- j. Time intervals between successive crew services must satisfy a lower limit.

Furthermore, the following constraints must be taken into account, based on the planning approach described in Subsection 2.3.

- k. In the case that a basic duty is a day duty, the corresponding temporary duty shall be a day duty with the same start and end locations.
- l. If a basic duty is an overnight duty, the corresponding temporary duty shall be an overnight duty with the same start and end locations (day duties are not allowed because it is difficult to take into account home rest time).
- m. Clock-in time of each temporary duty must be no later than the clock-in time of the corresponding basic duty.
- n. Clock-out time of each temporary duty must be within a fixed interval after clock-in time of the corresponding basic duty.

The constraint (m) on clock-in time is imposed because when a temporary duty begins earlier than the basic duty, there is a risk of being late to work. Besides, clock-out time is also constrained (constraint (n)); this is to avoid violating home rest regulations when a roster of the basic plan is implemented as is.

Clock-out times in rosters of actual basic plans are set with a certain margin against prescriptions; hence, the above constraint stems from the fact that no problem occurs in practice even if a certain period has elapsed.

## 4 | Development of Automatic Generation Method

### 4.1 | Past Research

There are many studies on the automatic generation of crew schedules dealing with timetable changes or disruptions. Both studies, conducted mostly in Europe and the US, are summarized by subject and method in Heil et al. [2]. However, it should be noted that work conditions—the most important constraints in crew scheduling—are partly regulated by law in each country, and there are often differences in details.

In Japan, too, there are some examples of such research: Imaizumi et al. [3], Nishi et al. [4], and Kato et al. [5] examined timetable changes, while Takahashi et al. [6] and Imaizumi et al. [7] addressed timetable disruptions. Both in Japan and abroad, most of the research on automated crew scheduling has applied mathematical optimization.

On the other hand, few studies have considered natural disaster situations, as assumed in this paper. Thus, Kunimatsu focused on train timetables [8], while Kato et al. [9] addressed the problem of rolling stock scheduling; however, to the best of our knowledge, no research has been done on crew scheduling, either in Japan or abroad. Response to natural disasters is similar, in part, to response to timetable disruptions (in both cases, the basic plan has to be revised); however, in the case of a disaster, time constraints are extremely strict, while some work conditions are relaxed. In other words, the requirements for temporary crew schedules are completely different from both cases of timetable changes and disruptions; therefore, existing methods of automated scheduling cannot be reused.

### 4.2 | Strategy

A crew schedule consists of multiple duties, and a natural approach is to find a good combination of duties. However, when modeling in terms of the classical set covering problem [10] or its derivatives, well known in the field of mathematical programming, one can make a natural connection by associating duties with columns (to be explained later). In so doing, it is unclear how to create unavailable duties, but the column generation algorithm [11] can be used to resolve this issue. That is, in the process of solving a problem, unavailable columns can be generated within a certain theoretical framework.

The column generation algorithm is generally an effective method for crew scheduling, but it does not go so far as to obtain the final schedule, and a different approach is needed for this purpose. Moreover, in the case of disasters considered in this study, there are some special circumstances, such as:

- Schedules with a reasonable level of accuracy should be created in a shorter time than usual.
- Specific conditions of a natural disaster cannot be ignored.

In this context, based on the methods by Nishi [4] or Kato [5], we propose a modified solution to reduce the computational time when addressing the problem dealt with in this study [12].

### 4.3 | Column Generation

For an integer programming problem (master problem), the column generation algorithm:

- Considers a linear programming master problem and assumes that a problem including part of the columns (restricted linear programming master problem) is solved by the simplex method.
- If the master problem is a minimization problem, columns with negative reduced cost (that may improve optimal values of the current restricted linear programming master problem) are found by solving a column generation sub-problem.
- These successive columns are added repeatedly until no more columns are generated to improve the objective function value of the optimal solution of the restricted linear programming master problem, that is, until an optimal solution is obtained for the linear programming master problem.
- Thus, a lower bound of the master problem is obtained.

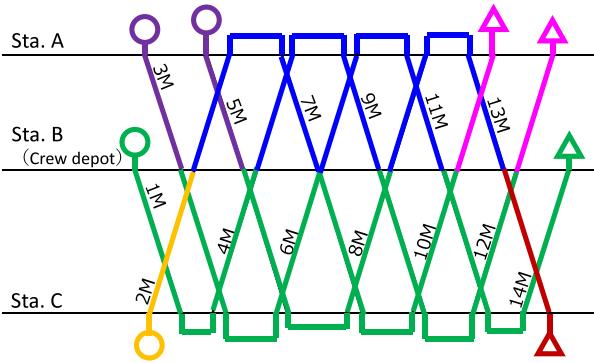
Furthermore, the branch-and-price method [13] combines the column generation algorithm with the branch-and-bound algorithm to find lower bounds for sub-problems, while generating columns even at terminal vertices. Hence, it is clear that even though an optimal solution for a linear programming master problem is obtained, a separate mechanism is required to obtain integer feasible solutions.

### 4.4 | Trip Planning

When modeling based on the set covering problem, one has to decide what the rows and columns correspond to. Typically, the columns are assigned to duties, and the rows are assigned to the smallest units of crew service. Here, these smallest units are “trips” generated based on rolling stock operation.

With reference to rolling stock operation information, trips are extracted by splitting rolling stock operation at transfer points, train stops, and locations where the stopping time is sufficient for transfer. On the other hand, splitting is not performed where the crew is engaged in shuttle service. This reduces the number of trips and, therefore, the number of rows (constraint expressions) in the master problem, that is, the problem scale. The locations where motormen and conductors are engaged in shuttle service are specified in the input data.

The motormen and conductors differ in their tasks (access to depots, shunting, car inspection, etc.), but these tasks are specified together with rolling stock operation and are therefore included in trips. Thus, the start time, end time, and duration of the tasks are defined, and sufficient breaks are ensured.



**FIGURE 4** | An example of trips. [Color figure can be viewed at wileyonlinelibrary.com]

An example of motormen's trips is shown in Figure 4. In this example, the train timetable and rolling stock schedules are the same as in Figure 1, but only Station B has a crew depot where crew changes are possible. In this case, trips are obtained by splitting every rolling stock schedule at Station B. In Figure 4, color changes correspond to trip changes, but trips with the same departure, intermediate stations, and arrival are shown in the same color. For example, Trip 1 is 1M (Station B to Station C) and 4M (Station C to Station B), while Trip 2 is 3M (Station B to Station C) and 6M (Station C to Station B); thus, trips do not split at Station C. In addition, the trips including only exit from a depot (circles) and entrance to a depot (triangles), such as shown in purple in Figure 4, are not allowed. Besides, when turning back at Stations A and C, the time between the arrival of a train and the departure of the next train is a lookout time, which is also included in trips. However, the concept of lookout may differ between motormen and conductors, and the trip's labor time (to be used below in the column generation sub-problem) is counted appropriately, based on the need for lookout, etc. The trips are defined as described above to ensure that sufficient crew is secured for tasks other than servicing main line trains, reduce the number of trips, and lessen the problem scale.

#### 4.5 | Master Problem

The master problem is formulated below. Notations are given in Table 1.

$$\min \sum_{k \in K} \sum_{j \in J^k} c_j^k x_j^k + \sum_{i \in I} d_i y_i \quad (1)$$

$$\text{s.t. } \sum_{k \in K} \sum_{j \in J^k} a_{ij}^k x_j^k - y_i = 1, \quad \forall i \in I \quad (2)$$

$$\sum_{j \in J^k} x_j^k \leq 1, \quad \forall k \in K \quad (3)$$

$$x_j^k \in \{0, 1\}, \quad \forall j \in J^k, \forall k \in K \quad (4)$$

$$y_i \geq 0, \quad \forall i \in I \quad (5)$$

Equation (1) defines the objective function. The 1<sup>st</sup> and 2<sup>nd</sup> terms pertain to working days and deadheading cost. Equations (2) to (5) define constraints. Equation (2) means that all trips are covered by one or another duty. Equation (3) means that up to

**TABLE 1** | Notations for the formulation of the master problem.

Notation	Description
$K$	Set of basic duties
$I$	Set of trips
$J^k$	Set of candidate duties corresponding to basic duty $k$
$a_{ij}^k$	1 if candidate duty $j$ corresponding to basic duty $k$ assigns trip $i$ , 0 otherwise
$c_j^k$	Workdays of candidate duty $j$ corresponding to basic duty $k$
$d_i$	Cost of deadheading of trip $i$
$x_j^k$	1 if candidate duty $j$ corresponding to basic duty $k$ is selected, 0 otherwise
$y_i$	Number of deadheadings of trip $i$

one temporary duty is selected for every basic duty. Equations (4) and (5) determine the possible ranges of variables.

Please note that in this formulation, as compared to the set covering problem, the original inequality constraint ( $\geq 1$ ) is changed to an equality constraint to count deadheadings; besides, an additional constraint is introduced so that up to one temporary duty candidate is assigned to every basic duty.

The linear programming master problem is obtained when the 0–1 condition for  $x_j^k$  is changed to  $x_j^k \geq 0$ , and the restricted linear programming master problem is obtained when  $J^k$  is replaced with its subset  $\bar{J}^k$ .

#### 4.6 | Column Generation Sub-problem

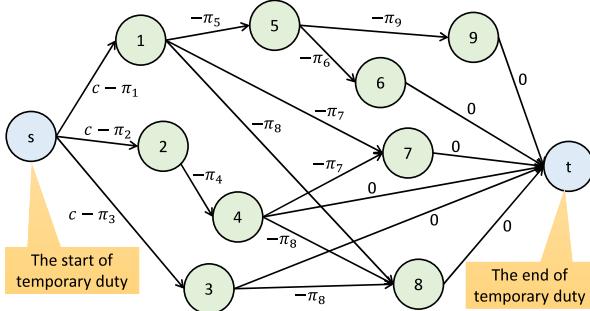
The column generation sub-problem is to derive new columns (duties) that can improve the objective function value, based on the information about solutions to the current restricted linear programming master problem. Generally, the current value of the objective function can be improved if the indicator called “reduced cost” is negative for columns unavailable in the current restricted linear programming master problem.

In this problem, to generate columns (temporary duties) corresponding to each basic duty, a column generation sub-problem is set for each basic duty, and all sub-problems are solved at every iteration.

Reduced cost of column  $j$  corresponding to basic duty  $k$  is defined as shown in Equation (6) below.

$$c_j^k - \sum_{i \in I} a_{ij}^k \pi_i + \theta^k \quad (6)$$

here  $\pi_i$  and  $\theta^k$  are dual variables for Equations (2) and (3), respectively. As explained above, the columns mean temporary duties; therefore, the generated columns must meet the constraints on temporary duties presented in Subsection 3.3. The problem of finding columns with the smallest reduced cost reduces to a constrained shortest path problem.



**FIGURE 5** | An example of a network of column generation subproblem. [Color figure can be viewed at wileyonlinelibrary.com]

In order to use the model of the constrained shortest path problem, the generation of temporary duties is expressed as a network. An example of the network is shown in Figure 5. Each node (point) corresponds to a trip. Each arc (branch) can handle connections between trips. The symbols over the arcs show arc-traversing costs set based on the dual variables. Here,  $s$  and  $t$  denote the start and end points, respectively; the goal is to find the shortest path on the network from  $s$  to  $t$ .

This network is defined for each basic duty; arcs from  $s$  and arcs to  $t$  are extended so as to satisfy constraints (b) and (k)-(n) in Subsection 3.3. Thus, these constraints can be satisfied by setting arcs on the networks and do not need to be taken into account in path search. Now, when it is necessary to send a crew to an isolated section, the required time must be taken into account when drawing arcs between nodes.

A pull-based labeling algorithm [3, 14] is applied to the shortest path problem. Each label is given multiple attributes that provide information about the path. The attributes of  $n$ -th label of node  $v$  are defined as shown below.

$$LABEL_n^v = \{f_n^v, g_n^v, h_n^v, p_n^v, q_n^v\} \quad (7)$$

Here the meaning of the symbols is as follows.

- $f_n^v$ : cost of  $n$ -th label of node  $v$  (reduced cost).
- $g_n^v$ : duty time of  $n$ -th label of node  $v$ .
- $h_n^v$ : labor time of  $n$ -th label of node  $v$ .
- $p_n^v$ : continuous crew service time of  $n$ -th label of node  $v$ .
- $q_n^v$ : meal time flag of  $n$ -th label of node  $v$  (flags are needed for all necessary items [breakfast, lunch, dinner]).

During the search, the label of a node is generated based on the label of the previous node; this is called “traversal”.

When traversing labels, a check is made to see whether the duty conditions (constraints (c) to (j) in Subsection 3.3) would be violated when that node (trip) were passed. No labels are generated in the case of a violation.

In addition, after traversal, a newly generated label is compared to all previous labels associated with the node. Let  $Label_1^v$  denote

the newly traversed node and  $Label_2^v$  denote the node to be compared;  $Label_2^v$  is deleted if the following condition is satisfied.

$$f_1^v \leq f_2^v \wedge g_1^v \leq g_2^v \wedge h_1^v \leq h_2^v \wedge p_1^v \leq p_2^v \wedge q_1^v \leq q_2^v \quad (8)$$

Alternatively,  $Label_1^v$  is deleted if the following condition is satisfied.

$$f_1^v \geq f_2^v \wedge g_1^v \geq g_2^v \wedge h_1^v \geq h_2^v \wedge p_1^v \geq p_2^v \wedge q_1^v \geq q_2^v \quad (9)$$

Thus, two labels are compared, and if all attribute values of one label are equal to or greater than those of the other, the latter is deleted. As a result, an increase in labels can be suppressed without deleting the shortest paths that meet the constraints, thus enabling efficient search. If comparison cannot determine the better choice, then deletion does not occur, and multiple labels exist on the node.

#### 4.7 | Calculation of Integer Solutions

Once an optimal solution to the linear programming master problem has been obtained, an integer feasible solution can be found in a number of ways, such as:

- Using an algorithm to generate integer solutions for generated columns.
- Using branch-and-price method.

In the former case, many columns (duty candidates) are generated at the starting point, and good combinations are found by selecting among those columns. However, the duty candidates are limited to those currently available, and obtaining an inherent optimal solution is not guaranteed, while some issues remain with the selection method. To guarantee an inherent optimal solution, one has to rely on the latter method that includes potential duties; however, this is a derivative of the branch-and-bound algorithm, and it is difficult to estimate the computational time.

On the other hand, with the problem considered in this study, obtaining a solution with a certain degree of accuracy within a practical computational time is more important than achieving optimality in the strict sense. Thus, we use heuristics to use some variable values obtained by multiple iterations of the simplex method in column generation as variable values of integer solutions. In so doing, we aim to find feasible solutions within a practical amount of time that is easy to estimate. This approach has already demonstrated a certain level of performance and reliability.

The basic idea is that when an optimal solution to the master problem is obtained through conventional column generation, the value of variable  $x_j^k$  is fixed at 1 if available; otherwise, the value of the largest variable is fixed at 1. Then column generation is applied again to the linear programming problem composed of undetermined variables. Such variable fixing is repeated until all trips are covered.

Heuristics for integer programming are described below.

**Step 1:** Let  $J_+ \leftarrow \emptyset$ , and  $I_+ \leftarrow \emptyset$ . LP optimal solution  $x$  for set of trips  $I$  is found by column generation.

**Step 2:** Let  $J_+ \leftarrow J_+ \cup \{(j,k) | x_{j,k}^k = 1\}$ . Besides, let  $I_{j,k}^k$  be set of trips covered by temporary duty  $(j,k)$  with  $x_{j,k}^k = 1$ , and  $I_+ \leftarrow I_+ \cup I_{j,k}^k$ . If one or more temporary duties were added to  $J_+$ , then

**Step 3;** otherwise, **Step 4**.

**Step 3:** Temporary duty  $(j,k)$  with the maximum fractional value is added to  $J_+$ . That is,  $J_+ \leftarrow J_+ \cup (j,k)$ . Besides,  $I_+ \leftarrow I_+ \cup I_{j,k}^k$ .

**Step 4:** If  $I = I_+$ , then all trips have been covered, and processing is terminated.

**Step 5:** LP optimal solution  $x'$  is found by column generation, and the algorithm returns to **Step 2** with  $x := x'$ .

The column generation in Step 1 above is a procedure to obtain an optimal solution to the linear programming master problem using a conventional column generation approach.

However, the original algorithm [4] addresses the set covering problem, while its derivative [5] is also different from the problem of this study in terms of constraints; thus, there were unclear points in the solution behavior.

## 4.8 | Measures to Reduce Computational Time

Preliminary experiments showed the need to reduce the computational time. Thus, the following measures were taken.

### 4.8.1 | Approximate Optimization of Column Generation Sub-Problem

When the column generation sub-problem is optimized, labels are not deleted easily if the conditions for label deletion are considered strictly. Preliminary experiments showed that this would lead to the generation of an enormous number of labels, thus requiring enormous computational time to find optimal paths. This phenomenon is likely to occur in problems with a large amount of label attributes.

On the other hand, the simplex method works for columns with negative reduced cost even if the values are not minimal (optimal); therefore, one can think of a way to easily find columns with negative reduced cost regardless of optimality. To obtain an optimal value of the linear programming master problem, in the strict sense, it is necessary to ensure that there are no columns with negative reduced cost. However, in this study, the calculation of optimal values of the linear programming master problem is relatively unimportant.

Thus, in this study:

- The column generation sub-problem is solved approximately, without strict comparison of labels (without focusing on columns with minimum reduced cost).

- Hence, the convergence of reduced cost is checked without strictly confirming the absence of columns with negative reduced cost.

When deleting labels, only  $g_{j,k}^v$  and  $h_{j,k}^v$  are compared in addition to  $f_{j,k}^v$ . The attributes other than reduced cost are generally non-decreasing in values and are considered to be dominant attributes in the labels.

### 4.8.2 | Addition of Multiple Columns to Restricted Linear Programming Master Problem

It was reported [3] that optimization can be completed faster when multiple rather than a single column with minimum reduced cost is added to the restricted master problem. Anticipating this effect, in this study, multiple columns with minimum reduced cost were added at every iteration of column generation. Specifically, for each basic duty, up to  $N_{\max}^1$  day duties or  $N_{\max}^2$  overnight duties were added as temporary duty candidates. These  $N_{\max}^1$  and  $N_{\max}^2$  are preset parameters.

## 5 | Verification

### 5.1 | Target Area and Verification Method

To evaluate the performance of the proposed method, we conduct verification in the case of a past large-scale natural disaster.

#### 5.1.1 | Problem Overview

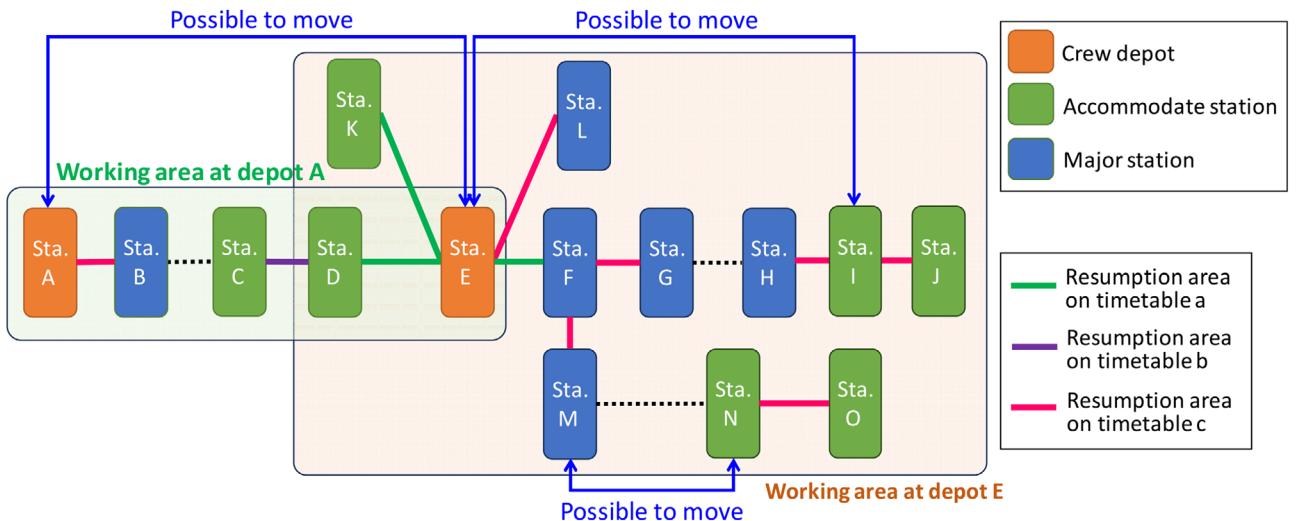
Routes of the target area are outlined in Figure 6. The target area includes a regional core city, the central station being Station E. The routes include the main line between Stations A and J, two branch lines stemming from Station E, and one branch stemming from Station F. The trains are operated with conductors; there are two crew depots near Stations A and E. The working areas of each crew depot are shown in the diagram.

#### 5.1.2 | Affected Sections, Service Disruption, Extension of Operational Sections

Many sections in the area were affected by a natural disaster, and the situation changed as follows.

- In a few days after the resumption of train service, only sections D-F and E-K were operational (temporary timetable a).
- A few days later, operations were resumed on section C-D (temporary timetable b).
- Then, in a few weeks, operations were resumed on sections A-B, F-G, H-J, E-L, F-M, and N-O (temporary timetable c).

In this study, verification is applied to these three temporary timetables.



**FIGURE 6** | Outline of target area for computational experiments. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

**TABLE 2** | Features of each instance using experiments.

Timetable	No. of trains	No. of trips
Basic plan	619	—
a	271	388
b	333	408
c	681	728

At the time of timetable *a*, only crew Depot E was operational, while both Depots A and E were operational when timetables *b* and *c* were implemented. However, even at the stage of timetable *c*, there were many disruptions along the way, and crew members had to be sent to isolated sections; bullet trains and taxis were used for that purpose. Specifically, on section F-J, crew members were sent from Depot A to Station I; on section N-O, trains were used from Station E to Station M, and then crew members were sent to Station N. Besides, though not isolated, section C-D was beyond the working area of Depot E; thus, crew members were first sent from Depot A to Station E and then traveled between Stations D and E.

The numbers of trains and trips at each stage are given in Table 2.

### 5.1.3 | Experimental Parameter Setting

Deadheading cost was set to  $d_i = 0.1(\forall i \in I)$ , and the upper limit for additional columns was set to  $N^1_{\max} = 10$ ,  $N^2_{\max} = 30$ . Now, crew members were assumed to leave a depot at the start of a duty and return only at the end of a duty, depending on the operating situation.

### 5.1.4 | Experimental Environment

Computation was conducted on a PC (OS: Windows 10 Pro, CPU: Core i7-8700K, RAM: 64 GB). Optimization was performed using

a mathematical optimization solver (Gurobi Optimizer 10.0.1) [15].

## 5.2 | Verification Results and Discussion

Results obtained by applying the proposed method to the temporary timetables are presented in Table 3. The table shows working days, deadheadings, average labor time (per day), standard deviation of labor time (s), and computational time (s) for every schedule. In addition, actual schedules created by the operator are also shown for comparison.

### 5.2.1 | Numerical Results

In either case, the number of working days was significantly smaller than in actual operation, thus confirming that efficient schedules were created by the proposed method. Also, the deadheadings were reduced for temporary timetables *a* and *b*. The computational time was also sufficiently short, even considering the requirements of disaster response: less than 10 min for timetables *a* and *b*, and less than 1 h for timetable *c* involving many trains. As shown in Subsection 4.7, feasible solutions were obtained through repeatedly solving an optimization problem not including integer variables, which is supposed to speed up the calculation.

Multiple columns were added during column generation, but it is difficult to determine how much it has contributed to the reduction of computational time. It should also be noted that approximate optimization of the column generation problem results in a dramatic reduction of the total computational time.

### 5.2.2 | Evaluation by Practitioners

On the other hand, when, in addition to these numerical results, the created schedules were shown to the practitioners in the

**TABLE 3** | Results of each timetable (\* refers to actual schedule).

Timetable	No. of working days	No. of deadheadings	Average labor time	Standard deviation of labor time (s)	Computational time (s)
a	59	6	5:42	3000.0	287
a*	72	14	4:50	4013.3	—
b	70	20	6:03	2772.1	413
b*	87	36	4:52	4334.9	—
c	126	39	5:59	4279.2	2112
c*	137	40	5:23	6047.8	—

conventional format, they pointed out that though the temporary duties did not contradict regulations, labor intensity was high.

In fact, the average labor time for any temporary timetable was longer than in actual operation. This can be interpreted in the following way: while efficiency improved as intended, it resulted in a heavier workload for each duty. One can assume that in the actual operation, schedules were created manually with regard to workload as well as efficiency, and workload should be taken into account in automated scheduling. We also found that the standard deviation of labor time was somewhat smaller as compared to actual operation, which was not an immediate problem.

Thus, we decided that a revision is needed to reduce and equalize the labor time as far as possible, even by using crew members at the expense of efficiency, and

### 5.3 | Revised Formulation

Based on the verification results of the previous subsection, we put an upper limit constraint on the working days. Specifically, the following constraint was added to the master problem in Subsection 4.5.

$$\sum_{k \in K^l} \sum_{j \in J^k} c_j^k x_j^k \leq w^l, \forall l \in L \quad (10)$$

here  $L$  is set of crew depots,  $K^l$  is set of basic duties  $k$  assigned to depot  $l$ , and  $w^l$  is the upper limit of working days for depot  $l$ .

Here, the upper limit was set for each crew depot individually, rather than simply on total working days; this is because in the latter case, duties might be concentrated in certain depots. Furthermore, when this constraint is imposed, working days are omitted from the objective function (1<sup>st</sup> term in Equation (1) is deleted), and only deadheadings are subject to minimization unless some other indicators are introduced.

Ideally, it would be desirable to be able to derive results within a specified number of working days. However, we used an inequality constraint in Equation (10) rather than an equality constraint; this is because we assumed that it would be difficult to calculate feasible solutions under an equality constraint in the context of the heuristics described in Subsection 4.7. An inequality constraint can significantly reduce the risk of not obtaining feasible solutions, while there is no guarantee that the obtained solutions reach the upper limit.

When the constraint of Equation (10) is imposed, the reduced cost in Equation (6) is revised as follows.

$$c_j^k - \sum_{i \in I} a_{ij}^k \pi_i + \theta^k + c_j^k \mu^l \quad (11)$$

here  $\mu^l$  is the dual variable for Equation (10). This variable depends on the crew depot and does not change with the duty content; therefore, it has no effect on the column generation network;  $\mu^l$  can be used to determine whether the final reduced cost will be negative.

The experiments were carried out again based on this revision. Besides, taking into account the practitioners' evaluation, clock-in time was introduced to the evaluation function, in addition to the deadheading minimization. Specifically, assuming that commuting to work early in the morning is a burden for the crew members, zero cost is assigned when clock-in time is  $u$  or later; otherwise, a cost is assigned proportionally to the earliness.

### 5.4 | Results of Formulation Revision

Table 4 presents results obtained for temporary duty  $c$  with the upper limit imposed on the working days (the actual schedule  $c^*$  is repeated from Table 3). These results pertain to the solution obtained with the total of constant  $w^l$  on the right side of the constraint set to 133.

The number of working days increased, but the average and standard deviation of labor time decreased. That is, these results are closer to the practical requirements. However, in this solution, the number of working days is 130 against the upper limit of 133, and the workload may be higher than expected. Thus, the constant  $w^l$  in Equation (10) must be set with care. For example, one can first ascertain the minimum value of the objective function in Equation (1) and then set the upper limit with some margin.

### 5.5 | Supplementary Note

Since the problem addressed in this study adds several constraints to the set covering problem, the behavior of the heuristics of integer solutions was initially unclear. Thus, comparison with the set covering problem is impossible, and there is no detailed information; however, there were no practical issues in terms of the solution accuracy and computational time of the proposed method.

**TABLE 4** | Results of timetable c set to maximum number of working days (\* refers to actual schedule).

Timetable	No. of working days	No. of deadheadings	Average labor time	Standard deviation of labor time (s)	Computational time (s)
c	130	33	5:49	3333.8	1617
c*	137	40	5:23	6047.8	—

The only concern is that feasible solutions were not obtained, depending on the value of the constant in the additional constraint in Equation (10). However, this issue can be resolved as explained in Subsection 5.4.

## 6 | Conclusion

In this study, we assumed a situation where a large-scale natural disaster occurred and focused on scheduling requirements whenever operational sections of railway lines are extended; aiming at reducing workload and deskilling, we developed a method for automatic generation of crew schedules.

An important consideration to keep in mind after a disaster is that crew members have already been notified of their attendance tables. In other words, changing the assigned off days to work days must be avoided. Thus, to prevent changing attendance tables, we took the approach of creating a temporary duty for each basic duty. Based on this approach, we developed a method using the column generation algorithm (a mathematical optimization technique) to create multiple candidates for temporary duties and select the most efficient duties.

To verify the effectiveness of the developed method, we used data from actual railway lines and data from past natural disasters. The verification results obtained for the data involving more than 500 trains demonstrated that automatic schedules generated within several 10 min required fewer working days than the schedules created manually at the time of disaster. On the other hand, the average and standard deviation of labor time increased, which suggested a potential necessity to consider workload in addition to efficiency. Thus, aiming at more practical scheduling, we proposed a method to put an upper limit on working days.

In the future, we will expand verification to more railway sections and other disaster cases; we will also explore an interface with existing railway systems toward early practical realization.

## References

- RTRI Transport Operation Systems Lab: Railway Scheduling Algorithm, NTS (2005) [in Japanese].
- J. Heil, K. Hoffmann, and U. Buscher, “Railway Crew Scheduling: Models, Methods and Applications,” *European Journal of Operational Research* 283, no. 2, 405–425 (2020).
- J. Imaizumi, T. Ueda, and S. Morito, “Application of Column Generation Algorithm to Railway Crew Scheduling Problem,” *Communications of the Operations Research Society of Japan* 56, no. 2 (2011): 104–109.
- T. Nishi, Y. Muroi, M. Inuiguchi, S. Takahashi, and K. Kataoka, “Application of Column Generation for Railway Crew Scheduling Problems With Practical Constraints,” *IEEJ Transactions on Electronics, Information and Systems* 131, no. 6 (2011): 1199–1208.
- S. Kato, T. Nakahigashi, and T. Kokubo, “Crew Scheduling Algorithm for Minimizing the Required Number of Crew Members,” *IEEJ Transactions on Industry Applications* 144, no. 5 (2024): 342–352.
- S. Takahashi, K. Kataoka, T. Kojima, and M. Asami, “An Algorithm for Automatically Modifying Train Crew Schedule,” *IEEJ Transactions on Industry Applications* 128, no. 11 (2008): 1291–1297.
- J. Imaizumi, R. Miura, E. Shigeta, and S. Morito, “A Column Generation Approach to Railway Crew Recovery Problem in Disruption,” *IEEJ Transactions on Electronics, Information and Systems* 134, no. 3 (2014): 418–424.
- T. Kunimatsu, “Automatic Generation of Provisional Timetables With Regard to Available Rolling Stock and Crew, and Fluctuation of Running Time” [in Japanese], in *29th Joint Symposium on Railway Technology (J-RAIL 2022)*, 173–176, (2022).
- S. Kato, J. Imaizumi, T. Nakahigashi, and T. Kokubo, “Optimization Approach to Rolling Stock Scheduling for Timetable Temporarily Changed due to Disaster” [in Japanese] *IEEJ Transactions on Industry Applications* 143, no. 5 (2023): 405–416.
- S. Umetani, From Mathematical Optimization Models to Algorithms, Kodansha (2021) [in Japanese].
- M. Kubo, A. Tamura, and Ch. Matsui, *Handbook of Applied Mathematical Programming* [in Japanese] (Asakura Publ, 2002).
- S. Kato, T. Nakahigashi, and J. Imaizumi, “Application of Mathematical Programming Method to Crew Scheduling After a Large-Scale Natural Disaster,” [in Japanese] (IEE Japan, 2024; TER-24-015/LD-24-015).
- C. Barnhart, E. L. Johnson, G. L. Nemhauser, M. W. P. Savelsbergh, and P. H. Vance, “Branch-and-price: Column Generation for Solving Huge Integer Programs,” *Operations Research* 46, no. 3 (1998): 316–329.
- J. Desrosiers, Y. Dumas, M. M. Solomon, and F. Soumis, “Time Constrained Routing and Scheduling”, in *Network Routing*, vol. 8 of *Handbooks in Operations Research and Management Science*, ed. M. O. Ball, T. L. Magnanti, C. L. Monma, and G. L. Nemhauser (Chapter 2, Elsevier, 1995).
- Gurobi Optimizer, <https://www.gurobi.com/jp/products/gurobi/>.

## Biographies



**Satoshi Kato**, member. In 2008 completed post-graduate studies at Waseda University (Grad. School of Fundamental Sci. and Eng.) and was employed by Railway Technical Research Institute. 2024 started doctorate at Waseda University (Grad. School of Creative Sci. and Eng.). Now senior researcher at Transport Operation Systems Lab, Signaling and Operation Systems Technology Division, Railway Technical Research Institute. R&D related to algorithms for transport planning and rescheduling support.



**Jun Imaizumi**, member. In 1989 graduated from Waseda University (Fac. of Fundamental Sci. and Eng., Dept. of Industrial & Management Systems Eng.), 1993 was employed as assistant by Waseda University, 1997 lecturer at Toyo University (Fac. of Business Adm.), 2002 assistant professor, then associate professor, since 2009 professor. D.Eng. Research related to various scheduling problems.



**Taichi Nakahigashi**, member. In 2015 completed postgraduate studies at Kyoto University (Grad. School of Sci.) was employed by West Japan Railway Company, 2021–2024 transferred to Railway Technical Research Institute. Now chief investigator at Rail System Transformation Section, Transport Department, West Japan Railway Company. Engaged in technology planning for transportation sector.