ISSN 0005-1179, Automation and Remote Control, 2018, Vol. 79, No. 9, pp. 1661–1672. © Pleiades Publishing, Ltd., 2018.
Original Russian Text © M.V. Buyanov, A.V. Naumov, 2018, published in Avtomatika i Telemekhanika, 2018, No. 9, pp. 143–158.

OPTIMIZATION, SYSTEM ANALYSIS, AND OPERATIONS RESEARCH

Optimizing the Operation of Rolling Stock in Organizing Cargo Transportation at a Railway Network Segment

M. V. Buyanov[*,a] **and A. V. Naumov**[*,b]

[*]*Moscow Aviation Institute, Moscow, Russia*
e-mail: [a]*buyanovmikhailv@gmail.com,* [b]*naumovav@mail.ru*

Received March 14, 2017

Abstract—We propose a mathematical model for the assignment of locomotives to transport freight trains. We consider various objective functions. One of the optimization objectives in our model is to minimize the number of locomotives involved in transportation by choosing the routes of trains and locomotives given that the daily transportation plan is fulfilled. The model is capable to account for different types of locomotives as well as different types of their technical maintenance. We propose a new heuristic algorithm for finding an approximate solution for this problem. The main tool of the proposed algorithm is a heuristic utility function that takes into account the topology of the railway network, restrictions imposed on the movement of locomotives, and also the need for technical inspection and repair of locomotives. Results of numerical simulation are presented with the example of real data regarding the movement of freight trains on a section of the Moscow Railway. We pay special attention to performing a qualitative analysis of the resulting solution, in particular, in order to reveal the dependencies between the values of the main qualitative characteristics of the motion and coefficients in front of the variables in the utility function. We assume that it is possible to control the total number of locomotives involved by changing the percentage of admissible idle and auxiliary runs.

Keywords: graph theory, integer optimization, locomotive assignment, utility function, freight transportation

DOI: 10.1134/S0005117918090102

1. INTRODUCTION

Recently, the strategy of scientific and technological development of the Russian Railways has been aimed especially at carrying out interdisciplinary works on the improvement and automation of cargo management systems. Recent works [1–8] consider many different methods and algorithms for the optimization of railroad cargo transportation. The work [5] proposes a way to improve the organization of freight transportation based on considering a single model of the logistics chain of supply of goods, including production, construction, reconstruction, and maintenance of infrastructure, that would ensure that all trains are received without delay even taking into account their uneven movement. In [6], two combinatorial tasks are considered: the scheduling problem for an aperiodic schedule and the problem of assigning platforms to trains. The deterministic combinatorial task of composing trains and constructing a railroad schedule has been formulated in [7, 8]. The work [8] presents various models that arise in scheduling for railway transportation, in particular an operational control model for the movement of a train and the model for the formation of freight flows through sorting stations.

In this works, we consider the mathematical model of assigning locomotives for the transportation of freight trains [1]. Unlike [1], in this work we take into account additional constraints on the passage of technical inspection of locomotives and types of locomotive traction. The improved model is closer to the real situation on the railroad, and the modified problem setting allows to optimize several auxiliary criteria by using a heuristic utility function while maintaining the basic idea of the model, i.e., minimizing the size of working locomotive park. We give a description of the developed algorithm for constructing an approximate solution of the problem.

2. OPTIMIZATION MODEL

We begin with a description of the basic concepts used in this paper. Most definitions given below have been introduced in [1], but some of them are presented with significant changes.

We associate a section of the railway network with a weighted directed graph $G = (V, A)$, where V is the set of vertices and A is the set of arcs. The vertices of graph G are significant stations, that is, stations where cargo trains are constructed (sorting stations) and stations where locomotive traction can change. Some significant stations are depots, and the corresponding subset of vertices V will be denoted by D. Arcs correspond to the runways connecting significant stations.

For each of the arcs $a \in A$ we define the electrification parameter $\mathcal{E}_a$ of the corresponding railroad train; it takes the following values: 0 if the run is not electrified, 1 if direct current is used, and 2 if alternating current is used.

Each locomotive has a characteristic T_l, the type of thrust that limits its movement depending on the type of runway electrification. Locomotives are classified according to the type of traction as follows: 0—locomotive (allowed to run through all tracks), 1—DC electric locomotive (only tracks with electrification parameter 1), 2—AC electric locomotive (only 2), 3—multi-system electric locomotive (tracks with electrification parameter 1 or 2). The relation $T_l \doteq \mathcal{E}_a$ determines the arcs with electrification parameter suitable for locomotive l. Locomotives can move only along certain routes (so-called legs), and hence we need to introduce the following definition.

Definition 1. A leg P is a sequence of arcs $a_1, \ldots, a_{I_P}$ of the graph G satisfying the following conditions:

1) all arcs, $a_i = (v_{i-1}, v_i)$, are different: $a_i \neq a_j$, $i, i \in \{1, \ldots, I_P\}$;

2) the first vertex of the first arc in the sequence coincides with the last vertex of the last arc of the sequence, represents a depot station, and is different from all intermediate vertices in the sequence: $v_1 = v_{I_P} \in D$, $v_i \neq v_1$ for $i = \overline{2, I_P - 1}$.

We will also consider sublegs and simple sublegs, defined as follows.

Definition 2. Any subsequence of adjacent arcs $a_i, a_{i+1}, \ldots, a_j$ $(1 \leqslant i < j \leqslant I_P)$, that form a leg is called a subleg of this leg. Any arc $a_i = (v_{i-1}, v_i)$ that occurs in some leg P is called a simple subleg of P.

Let L be the set of all locomotives assigned to the depot stations in question. For each locomotive $l \in L$, we are given a set of admissible legs $\overline{\mathcal{P}}_l$ through which it can move. We associate to each set $\overline{\mathcal{P}}_l$, $l \in L$, the set $\mathcal{P}_l$ composed of all simple sublegs that occur in the legs from the set $\overline{\mathcal{P}}_l$, and for each simple subleg $p \in \mathcal{P}_l$ the relation $T_l \doteq \mathcal{E}_p$ is satisfied. We will assume that for each locomotive $l \in L$ we know a weight norm function $w_l(\cdot)\colon \mathcal{P}_l \to \mathbb{R}$ that relates simple sublegs $\overline{\mathcal{P}}_l$ with the maximum admissible weight for the carriage.

Let S be the set of freight trains. Each train $s \in S$ is characterized by its mass w^s, the origin station v_0^s, the destination station v_f^s, formation time t_0^s, and the time τ_f^s before which the train has to arrive to its destination station, i.e., each train corresponds to a five-tuple $(w^s, v_0^s, t_0^s, v_f^s, \tau_f^s)$. In essence, these characteristics determine the transportation plan.

Movement of locomotives and trains along a given route can be carried out only at certain intervals. The combination of a route and time is called a thread.

Definition 3. A thread N is an ordered set of quadruples (v_1, t_1, v_2, τ_2), (v_2, t_2, v_3, τ_3), $\ldots$, $(v_{I_N-1}, t_{I_N-1}, v_{I_N}, \tau_{I_N})$ that satisfies the following conditions:

1) $v_i \in V$, $i = \overline{1, I_N}$, $t_i \in \mathbb{R}$, $i = \overline{1, I_N - 1}$, $\tau_i \in \mathbb{R}$, $i = \overline{2, I_N}$;

2) $(v_i, v_{i+1}) \in A$, $i = \overline{1, I_N - 1}$;

3) $t_i < \tau_{i+1}$, $i = \overline{1, I_N - 1}$;

4) $\tau_i \leqslant t_i$, $i = \overline{2, I_N}$.

In the definition above, the value t_i corresponds to the departure time from a station v_i, and τ_{i+1} is the arrival time to a station v_{i+1}. These conditions express the natural properties of train traffic, that is, the fact that the movement can be carried out only over tracks (conditions 1, 2), the departure time from a station cannot be later than the time of arrival at the next station (condition 3), and the time of arrival to a station cannot be later than the departure time from the same station (condition 4).

By analogy with legs and simple legs, we introduce subthreads and simple subthreads.

Definition 4. Each subsequence of adjacent quadruples that forms a thread N is called a subthread. Each quadruple $(v_i, t_i, v_{i+1}, \tau_{i+1})$, $i = \overline{1, I_N - 1}$, that comprises a string N, is called a simple subthread.

Consider a set $\overline{\mathcal{N}}$ of threads. We assign to each element N in this set a set $\mathcal{F}(N)$, which is an unordered set of simple subthreads that comprise the thread N. The set of all simple subthreads obtained from the set of threads $\overline{\mathcal{N}}$ is denoted by $\mathcal{N}$, i.e.,

$$\mathcal{N} = \bigcup_{N \in \overline{\mathcal{N}}} \mathcal{F}(N). \tag{1}$$

It is important to note that each simple subthread passes through only one of the arcs of the graph.

On the set $2^L \times \mathcal{N}$, which is the Cartesian product of all possible combinations of locomotives and the set of simple subthreads, we define a function $W(\pi_n)$ that specifies the maximum mass of a train that the corresponding combination of locomotives $\pi_n \subset L$ can transport along a given simple subthread $n \in \mathcal{N}$. Obviously, if $\pi_n = l \in L$, where $n = (v, t, v', \tau)$, and $(v, v') \in \mathcal{P}_l$, then $W(\pi_n) = w_l((v, v'))$. The combination of locomotives π_n is called a composite locomotive and is used to transport the train through their joint operation.

Since the movement of locomotives is carried out only along threads and legs, we introduce the definition of an admissible route of a locomotive's rotation relative to the set of legs. In this definition, we also take into account that the locomotive must pass $I^{\mathrm{MR}} = 5$ types of maintenance and repair (MR) in time intervals $T^{\mathrm{MR}} = (2, 30, 90, 365, 1095)$ days and with duration $t^{\mathrm{MR}} = (8, 8, 12, 600, 1080)$ hours respectively, at specially equipped stations V^{MR}, where $V^{\mathrm{MR}} \subset V$. Depending on the locomotive type, values in the vectors T^{MR}, t^{MR} may be different. We will assume that each locomotive $l \in L$ at the initial moment of time is characterized by elements of the vector τ_l^{MR}, $|\tau_l^{\mathrm{MR}}| = I^{\mathrm{MR}}$, which are equal to the times elapsed since the corresponding maintenance. If the locomotive is undergoing maintenance at the initial moment of time, then the corresponding element of the set τ_l^{MR} takes a negative value with absolute value equal to the time before the end of the work.

Definition 5. An admissible turnaround route M_l of locomotive l with respect to the set of legs $\overline{\mathcal{P}}_l$ is a sequence of simple subthreads (v_1, t_1, v_2, τ_2), (v_2, t_2, v_3, τ_3), $\ldots$, $(v_{I_l-1}, t_{I_l-1}, v_{I_l}, \tau_{I_l})$ that satisfies the following conditions:

1) $\tau_i \leqslant t_i$, $i = \overline{2, I_l - 1}$;

2) $(v_i, v_{i+1}) \in \mathcal{P}_l$, $i = \overline{1, I_l - 1}$;

3) there exist $m = \overline{1, I^{\mathrm{MR}}}$ increasing sequences $i_1^m, \ldots, i_{f_l}^m$ of numbers chosen from the set $\{2, 3, \ldots, I_l\}$;

4) $(\tau_l^{\mathrm{MR}})_m + \tau_{i_1^m} \leqslant (T^{\mathrm{MR}})_m$;

5) $t_{i_j^m} - \tau_{i_j^m} \geqslant (t_l^{\mathrm{MR}})_m, \quad j = \overline{1, f_l - 1}$;

6) $\tau_{i_j^m} - t_{i_{j-1}^m} \leqslant (T^{\mathrm{MR}})_m, \quad j = \overline{2, f_l}$;

7) $\tau_{I_l} - t_{i_{f_l}^m} \leqslant (T^{\mathrm{MR}})_m, \quad \text{if } f_l \neq I_l$;

8) $v_{i_j^m} \in V^{\mathrm{MR}}, \quad j = \overline{1, f_l}$.

Condition 1 requires that the arrival time to the station should not be earlier than the departure time from the same station. Condition 2 limits possible movements of a locomotive by moving only along legs. Condition 3 requires the locomotive to pass maintenance at specified intervals. Sequences of time moments $t_{i_1^m}, \ldots, t_{i_{f_l}^m}$ correspond to the moments when the corresponding series of maintenance begins. Condition 4 requires that the departure time for the first maintenance in the series m does not exceed $(T^{\mathrm{MR}})_m$ from the previous maintenance. According to condition 5, transit time of the corresponding maintenance cannot be less than $(t_l^{\mathrm{MR}})_m$. Condition 6 implies that the time between the beginning of the movement after maintenance and departure to the next maintenance cannot be greater than $(T^{\mathrm{MR}})_m$. According to condition 7, the time when movement begins after the last maintenance must be no later than the time $(T^{\mathrm{MR}})_m$ before the end of the considered period of traffic planning. Condition 8 guarantees that maintenance is done at specially equipped stations V^{MR}.

Note that a turnover route is a spatio-temporal concept. We denote the set of admissible turnover routes for a locomotive l by $\mathcal{M}_l$. The initial and final stations of a turnover route M_l will be denoted by $v_0(M_l)$ and $v_f(M_l)$ respectively, the start time of the first thread of this traffic route will be denoted by $t_0(M_l)$, the time of arrival at the destination station, by $\tau_f(M_l)$.

We introduce the definition of an admissible train run, which is also a spatio-temporal characteristic, just like a locomotive's turnover route.

Definition 6. A valid run R_s of a train $s \in S$ is a sequence of simple subthreads (v_1, t_1, v_2, τ_2), $(v_2, t_2, v_3, \tau_3), \ldots, (v_{I_s-1}, t_{I_s-1}, v_{I_s}, \tau_{I_s})$ that satisfies the following conditions:

1) $v_1 = v_0^s$;

2) $v_{I_s} = v_f^s$;

3) $\tau_i \leqslant t_i$, $i = \overline{2, I_s - 1}$;

4) $t_0^s \leqslant t_1$;

5) $\tau_f^s \geqslant \tau_{I_s}$.

Conditions 1 and 2 determine the initial and final stations of the route, condition 3 specifies natural constraints on departure and arrival times, conditions 4 and 5 require that transportation is fulfilled according to the plan.

The set of admissible routes for a train s will be denoted by $\mathcal{R}_s$.

Similar to the threads, we define the set $\mathcal{F}(M_l)$ of all simple subthreads that comprise a turnover route M_l of a locomotive l, $l \in L$, and the set $\mathcal{F}(R_s)$, $s \in S$, of all simple subthreads that comprise a run R_s of a train s.

For each simple subthread $n \in \mathcal{N}$ and each set of turnover routes for the locomotives $M = \{M_l\}_{l \in L}$, we define the set $\pi_n(M)$ composed of all locomotives moving along a simple subthread n in the set of locomotive turnover routes M:

$$l \in \pi_n(M) \Leftrightarrow n \in \mathcal{F}(M_l). \tag{2}$$

3. PROBLEM SETTING

Consider a section of the railway network with graph $G = (V, A)$ defined above. Suppose that the set of locomotives L, the set of trains S, the set of threads $\overline{\mathcal{N}}$ and the corresponding simple subthreads $\mathcal{N}$, and the weight function $W(\cdot)$ for composite locomotives are known. For each locomotive $l \in L$ the corresponding set of legs $\overline{\mathcal{P}}_l$ and simple legs $\mathcal{P}_l$ are defined.

At the initial time moment, a locomotive can be in one of the following states: in motion, at a station, or in a depot. For each locomotive $l \in L$, we are given a station v_0^l that, depending on the state of the locomotive, is either a locomotive station if the locomotive is at a station or located in the depot, or its destination station if the locomotive is in motion, and time t_0^l when the locomotive will complete the current work, i.e., will arrive at the destination station.

Suppose that for each locomotive $l \in L$ we know the set of times τ_l^{MR} that have elapsed since the last maintenance, as well as the sets T^{MR}, t^{MR}, V^{MR} that define the conditions for passing maintenance for all locomotives. Taking into account the times τ_l^{MR} for each locomotive $l \in L$, we obtain a set of valid turnover routes $\mathcal{M}_l$ according to definition 5. Similarly, for each train $s \in S$ we know a set of valid runs $\mathcal{R}_s$. Let $||L||$ be the number of locomotives in the set L that have a non-empty turnover route.

Suppose that for each train $s \in S$ we specify a set of threads $\overline{\mathcal{N}_s} \subset \overline{\mathcal{N}}$ along which it can be transported. We denote by $\mathcal{N}_s$ the set of corresponding simple subthreads. These constraints are due to the fact that some of the threads can be used to transport trains only of a certain kind.

Let $M = \{M_l\}_{l \in L}$ be a selectable set of turnover routes for all locomotives, $R = \{R_s\}_{s \in S}$, a selectable set of runs for all trains, $\mathcal{M} = \{\mathcal{M}_l\}_{l \in L}$, the set of valid routes for the turnover of all locomotives, $\mathcal{R} = \{\mathcal{R}_s\}_{s \in S}$, the set of admissible runs of all trains.

The problem is to find such a set M of turnover locomotive routes and such a set R of runs for the trains for which the total number of locomotives used for transportation of trains will be minimal, with all runs of trains covered by locomotive routes.

In [1], the following statement of the problem was proposed:

$$||L|| \to \min_{M \in \mathcal{M}, R \in \mathcal{R}} \tag{3}$$

under constraints

$$M_l \in \mathcal{M}_l, \quad l \in L, \tag{4}$$

$$R_s \in \mathcal{R}_s, \quad s \in S, \tag{5}$$

$$\bigcup_{s \in S} \mathcal{F}(R_s) \subset \bigcup_{l \in L} \mathcal{F}(M_l), \tag{6}$$

$$\mathcal{F}(R_s) \cap \mathcal{F}(R_{s'}) = \varnothing, \quad s \neq s', \quad s, s' \in S, \tag{7}$$

$$W(\pi_n) \geqslant w^s, \quad n \in \mathcal{F}(R_s), \quad s \in S, \tag{8}$$

$$\mathcal{F}(M_l) \subset \mathcal{N}, \quad l \in L, \tag{9}$$

$$\mathcal{F}(R_s) \subset \mathcal{N}_s, \quad s \in S, \tag{10}$$

$$v_0(M_l) = v_0^l, \tag{11}$$

$$t_0(M_l) \geqslant t_0^l. \tag{12}$$

Conditions (4), (5) mean that only admissible locomotive routes and runs of trains are considered, in particular those for which valid legs exist. We also note that the admissibility of train runs requires the transportation plan to be completed within a specified period.

Condition (9) requires that locomotive turnover routes are composed only of simple subthreads since the set $\bigcup\limits_{l\in L} \mathcal{F}(M_l) \subset \mathcal{N}$ consists of simple subthreads that are part of any locomotive turnover route. Condition (10) specifies a similar requirement for train runs, and in addition it restricts the choice of admissible threads for transporting the train. Condition (6) means that all simple subthreads that comprise a route of some train are used for the motion of some locomotive, i.e., all trains are transported by locomotives. It also follows from this condition that locomotives can move along simple subthreads over which the trains do not move. Thus, each thread used in the solution corresponds to either a train with a locomotive (possibly with several locomotives) or a locomotive moving empty.

Condition (7) means that runs of the trains cannot intersect, i.e., one simple subthread cannot be used for the movement of two trains. Since locomotives can travel in a raft or with a train (so-called auxiliary runs), there is no similar condition for locomotives.

Condition (8) requires that the weight norms of composite locomotives are fulfilled by the trains, i.e., the composite locomotive π_n, used on a simple subthread $n \in \mathcal{F}(R_s)$ by which train s is transported, should be able to carry a train of weight $W(\pi_n)$, no less than the mass w^s of the train s.

Conditions (11), (12) specify the initial state of locomotives.

Note also that the set of trains S and the set of threads $\overline{\mathcal{N}}$ is determined by the daily transportation plan and the number of days for which planning is carried out.

The formulated mathematical model has a general character and assumes optimization both along locomotive turnover routes and along the runs of the trains. In what follows we consider a special case of problem (3)–(12), where we do not optimize train runs and solve only the task of assigning locomotives. We will assume that the set of admissible runs $\mathcal{R}_s$ of train $s \in S$ consists of a single run. Thus, the task is reduced to assigning locomotives to transport trains with specified runs, i.e., to the search for a set of routes M.

Problem (3)–(12) is a complex discrete optimization task whose dimension, given the size of the railway network, does not allow to obtain an exact solution by a complete enumeration of all possible options. In addition, the problem statement shown above assumes optimization by a single criterion, the number of locomotives used, but in practice this may not be enough. The management of locomotive park operation, in addition to fulfilling the transportation plan, also implies considering various characteristics of locomotive use when they are assigned for the transportation of trains, including the number and duration of runs, proportion of auxiliary and idle runs, and so on.

The purpose of this work is to construct a heuristic algorithm for finding an admissible solution to problem (3)–(12) that produces the value of the criterion no worse than the solution currently used by the Russian Railways. We will call the solution we find suboptimal. At the same time, the algorithm should have a number of tuning parameters, and by changing these parameters one should be able to control different characteristics of the use of the locomotive park mentioned above. The main idea of the algorithm is to assign locomotives sequentially to simple subthreads of all trains sorted by the time of departure on the basis of solving an auxiliary problem with an objective function in the form of a special utility function for assigning a locomotive to a simple subthread.

Let us define this utility function and the auxiliary problem. We define the set $\mathcal{F}^*(M_l, R) = \mathcal{F}(M_l) \cap \mathcal{F}(R)$ consisting of simple subthreads of all train runs to which locomotive l is assigned according to route M_l, ordered by the time of departure of simple subthreads, where $\mathcal{F}(R)$ is the set of simple subthreads of runs R of all trains S. We also define the unordered set $\mathcal{F}(M_l, R_s) = \mathcal{F}(M_l) \cap \mathcal{F}(R_s)$ that consists of simple subthreads of run R_s of train s to which locomotive l is assigned according to route M_l. Next we introduce the objective function, on the basis of which each locomotive $l \in L$ will be assigned to a specific route M_l from the set of admissible turnover routes $\mathcal{M}_l$:

$$U(M_l) = u(l, t_0^l, v_0^l, n_1) + \sum_{i=2}^{|\mathcal{F}^*(M_l,R)|} u(l, \tau(n_{i-1}), v_f(n_{i-1}), n_i), \quad n_i \in \mathcal{F}(M_l, R), \tag{13}$$

where

$$u(l, \tau, v_f, n) = \omega_1 K_1 + \omega_2 K_2 + \omega_3 K_3, \tag{14}$$

$$M_l \in \mathcal{M}_l, \quad l \in L, \tag{15}$$

$$n \in R_s, \quad s \in S, \tag{16}$$

$$K_1 = \frac{|\mathcal{F}(M_l, R_s)|}{|\mathcal{F}(R_s)|}, \tag{17}$$

$$K_2 = \frac{t_n - \tau}{t_{\max}}, \quad \tau \leqslant t_n, \tag{18}$$

$$K_3 = -\frac{d(v_f, v_0^n)}{d(v_0^n, v_f^N)}. \tag{19}$$

Function (14) is the utility function of assigning a locomotive l to a simple subthread n, taking into account the time τ and station v_f where the locomotive stops at the previous simple subthread.

Relation (17) allows us to estimate the "utility" of a locomotive l for transporting train s, i.e., determines what part of the train run is contained in the route of this locomotive, taking into account constraints on the legs and maintenance. Relation (18) is used to account for the time the locomotive has been idle before it is assigned to a simple subthread n, where τ is the stop at the previous simple subthread in the itinerary of the locomotive, t_n is the departure time for simple subthread n, $t_{\max}$ is the normalization factor that depends on the chosen frame of reference. Relation (19) determines the penalty for running the locomotive, where v_f is the locomotive stop station for the previous simple subthread in the route, v_0^n is the departure station for simple subthread n, and v_f^N is the destination station for the thread N that includes the simple subthread n. Function $d(v_i, v_j)$ is used to calculate the number of arcs between stations v_i and v_j respectively. Coefficients ω_1, ω_2, and ω_3 are weight coefficients that take values in the range $[0, 1]$ and satisfy condition $\omega_1 + \omega_2 + \omega_3 = 1$. Varying the weighting factors involves changing such parameters of locomotive motion as the number and duration of runs and the idle time. Note that $K_1, K_2 \in [0, 1]$, but $K_3 \in [0, |A|]$, therefore, $u(l, \tau, v_f, n) \in [-|A|, 1]$.

Thus, we can formulate an auxiliary optimization problem of choosing the optimal route for a locomotive in the sense of objective function (13):

$$U(M_l) \to \max_{M_l \in \mathcal{M}_l} \tag{20}$$

under constraints (4)–(12).

Solving the formulated problem by specifying optimal turnover routes for the sequence of locomotives that come from a depot allows us to find a suboptimal solution of the optimal assignment problem (3)–(12) for the locomotives formulated in [1]. The algorithm for obtaining this suboptimal solution is given in the next section.

4. ALGORITHM FOR LOCOMOTIVE ASSIGNMENT

In this section, we present an algorithm for finding a suboptimal solution to problem (3)–(12) based on the use of the heuristic utility function (14). We assume that the runs of all trains are defined, i.e., for each train the thread along which it moves is determined. The algorithm looks for an admissible solution to the problem of assigning locomotives to trains, which is a special case of the general mathematical formulation (3)–(12) without optimization by $R \in \mathcal{R}$. In addition, the problem is to assess the total volume of the locomotive park necessary for efficient transportation of the cargo according to the specified cargo traffic determined by the daily schedule. Therefore, there are no restrictions on the number of locomotives available in each depot. Optimization by criterion (3) will allow to estimate, with the help of the constructed suboptimal solution, the necessary number of locomotives for each depot. After this, the algorithm can be used again, taking into account the constraints on the number of locomotives found for each depot. In this case, the algorithm can begin with any initial distribution of locomotives over the considered section of the Russian railway network (a part of the rolling stock is in the depot, a part is in transit or at the stations of the network). For simplicity of exposition of algorithms, in their constructions we do not take into account restrictions on the mass of transported trains. When the masses of the trains are constrained, it is necessary to search not only for simple locomotives $l \in L$ but also for composite locomotives, i.e., combinations of several locomotives, but the algorithm requires only an insignificant modification.

4.1. The Assignment Algorithm

Consider a non-empty set of trains $S = \{s_i \mid i = \overline{1, |S|}\}$ with nonempty runs. Let v_f^l, τ^l be the final station of the turnover route M_l (or the initial station in case of an empty route) and the time of arrival at this station for locomotive $l \in L$. Note also that $||L|| = 0$, and the value $|S|$ corresponds to the number of trains in the transportation plan.

Algorithm 1.

0. Set $i := 1, j := 1, k := 1$.

1. Fix a train $s_i \in S$ and a simple subthread from the run $n_j := (v_0(n_j), t(n_j), v_f(n_j), \tau(n_j))$, $n_j \in \mathcal{F}(R_{s_i})$. Fix a locomotive $l_k \in L$.

2. If $j > |\mathcal{F}(R_s)|$, then go to the next train $i := i + 1$, $j := 1$. If $i > |S|$ then go to the next locomotive $k := k + 1$, $i := 1$. If $k > ||L||$, then take a new locomotive $L := L \bigcup \{l_k\}$ from the nearest depot $v_0(n_j)$ under the condition $(v_0(n_j), v_f(n_j)) \in \mathcal{P}_{l_k}$, denote for the locomotive l_k the set of selectable simple subthreads as $N(l_k)$, and let $N(l_k) := \varnothing$.

3. If $\tau^{l_k} \leqslant t(n_j)$, $(v_0(n_j), v_f(n_j)) \in \mathcal{P}_{l_k}$, go to step 4. Otherwise, go to step 5.

4. If $v_f^{l_k} \neq v_0(n_j)$, search for the thread N^* to move the locomotive l_k to the beginning of the simple subthread n_j according to Section 4.2. If $v_f^{l_k} = v_0(n_j)$, we assume $N^* := \varnothing$. If a thread N^* is found, then check the maintenance constraints according to Section 4.3. If maintenance constraints are satisfied, fix the simple subthread n_j in the set of subthreads $N(l_k) := N(l_k) \cup \{n_j\}$ admissible for this locomotive.

5. If $i = |S|$ and $j = |\mathcal{F}(R_s)|$, then go to step 6, otherwise set $j := j + 1$ and go to step 2.

6. Find a simple subthread n from the set of admissible subthreads $N(l_k)$ that has the maximum value of the utility function (14) compared to the others; if values of the utility function for different simple subthreads coincide, preference is given to the subthread with earlier time of movement. Enter the found simple subthread n and, if necessary, the corresponding thread for movement N^* in the locomotive's route $M_{l_k} := M_{l_k} \cup N^* \cup \{n\}$. Remove simple subthread n from the run $R_s := R_s \setminus \{n\}$, where s is the train that is moving along this thread. If $\mathcal{F}(R_s) = \varnothing$ then remove the train from the set of trains $S := S \setminus \{s_i\}$. If $S = \varnothing$, go to step 7, otherwise let $N(l_k) := \varnothing$, $i := 1$, $j := 1$ and go to step 2.

7. End of the algorithm; a suboptimal solution for problem (3) has been obtained.

4.2. Finding a Composite Thread

To move a locomotive, we have to find a thread N^* that connects station v_f^l, where the locomotive l is located, and station $v_0(n)$, from which the simple subthread n departs. Let $t(n)$ be the start time of the simple subthread n, $\tau(n)$—end time of simple subthread n, and τ^l—the stopping time for the locomotive at station v_f^l. We will denote by $\mathcal{N}_a$ the set of simple subthreads corresponding to the arc $a \in A$. We associate with each arc of the graph G a weight characteristic equal to the average time of moving along it:

$$w_a = \frac{1}{|\mathcal{N}_a|} \sum_{n \in \mathcal{N}_a} (\tau(n) - t(n)). \tag{21}$$

Set N^* equal to the thread that runs along the shortest path in column G weighted according to (21), which connects stations v_f^l and $v_0(n)$ with start time not earlier than τ^l and end time no later than $t(n)$. To find the shortest paths between vertices of a weighted directed graph one can use, for example, the Floyd–Warshell algorithm [9].

4.3. Checking Maintenance Constraints

To ensure timely maintenance of locomotives according to constraints specified in Section 2, we have to carry out the corresponding checks. All constraints are checked when the locomotive is assigned to a simple subthread.

To simplify the description of the algorithm, we describe the version that checks constraints associated with one type of maintenance. In this case, $T^{\mathrm{MR}}, \tau^{\mathrm{MR}}$ are scalar values corresponding to the type of maintenance being considered.

Consider a locomotive l with a specified turnover route M_l, the time τ_l^{MR} is when the last maintenance was conducted; the time T^{MR} determines the frequency of maintenance; V^{MR} is the set of stations where maintenance can be executed; and $n(v_0, t, v_f, \tau)$ is the simple subthread to which this locomotive should be assigned. Let v_f^l, τ^l be the final station of the turnover route M_l (or the initial station in case of an empty route) and the time of arrival at this station of the locomotive l. We describe the algorithm for checking maintenance constraints.

Algorithm 2.

1. If $\tau \leqslant \tau_l^{\mathrm{MR}} + T^{\mathrm{MR}}$ or $\tau^l + T^{\mathrm{MR}} \leqslant t$, go to step 2, otherwise go to step 4.

2. If $v_f \notin V^{\mathrm{MR}}$, search for thread N^* to move the locomotive to the nearest maintenance station according to Section 4.2. If such a thread is found, then check constraints related to the maintenance for the locomotive l assigned to each simple subthread of the thread N^*, similar to step 1 of the algorithm. If there is at least one simple subthread of the thread N^* for which the constraints are not satisfied, go to step 4. Go to step 3.

3. Simple subthread n satisfies maintenance constraints and can be included in the locomotive's turnover route; the algorithm halts.

4. Restrictions are not satisfied; the algorithm halts.

5. RESULTS OF NUMERICAL EXPERIMENTS

We have carried out a numerical experiment using sample data from a segment of the Moscow Railway (MZD) over a certain period of time. Characteristics of the input data are given in Table 1.

Calculations were carried out taking into account constraints maintenance. We assumed that the locomotives must pass maintenance of the same type, with a duration of at least 8 hours, at intervals not exceeding 48 hours. For all locomotives, admissible stations for passing the maintenance are

Table 1. Input data characteristics

Number of stations	40
Number of depot stations	16
Number of sorting stations	16
Number of trains in daily assignment	598
Number of threads per day	1254
Planning period (days)	10

Table 2. Varying the weights of the utility function

ω	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Number of locomotives	442	439	445	439	439	441	443	450	444	443	449
Average idle time (%)	50	50	52	53	53	54	53	52	53	53	55
Average time for maintenance (%)	17	17	17	17	17	17	17	18	17	17	18
Average time spent on useful mileage (%)	21	21	20	21	21	20	20	20	20	20	19
Average time spent on idle run (%)	12	11	11	9	9	9	10	10	10	10	8

Table 3. Using the locomotives

Number of trains	Day	Number of locomotives used in day i and entered operation in day j											Locomotives in day i
		1	2	3	4	5	6	7	8	9	10	11	
299	1	321	0	0	0	0	0	0	0	0	0	0	312
598	2	319	82	0	0	0	0	0	0	0	0	0	401
598	3	320	74	14	0	0	0	0	0	0	0	0	408
598	4	320	75	6	9	0	0	0	0	0	0	0	407
598	5	321	77	8	7	14	0	0	0	0	0	0	420
598	6	320	77	11	9	10	1	0	0	0	0	0	416
598	7	320	78	12	9	14	0	3	0	0	0	0	420
598	8	321	71	10	9	11	0	1	0	0	0	0	408
598	9	320	75	7	7	14	0	1	0	2	0	0	417
598	10	321	70	8	6	13	0	0	0	0	0	0	401
299	11	294	52	4	4	9	0	0	0	0	0	0	350

depot stations. In order to adequately simulate the process of locomotive maintenances at the time when a locomotive enters operation, the time elapsed since the last maintenance order is randomly generated according to the uniform distribution over $[0, 48]$ h. The daily schedule of trains is assumed to be known and remains the same for the entire planning period. These conditions for carrying out a numerical experiment differ from the ones in [1] since now we have constraints on the stations where maintenance can be done. The decision time depends on the input data; for example, increasing the planning period greatly increases the running time of the algorithm, and a change in the number of trains in the daily task practically does not affect the running time of the algorithm. For the considered numerical experiment, we found a solution in under 10 min.

We show Table 2, "Varying the weights of the utility function," which presents average values of the basic characteristics of locomotive motion and the value of the main criterion for the values of

Table 4. Comparison of results

Characteristics	Solution from [1]	Solution with UF	Solution with UF without add. constraints
Number of locomotives	369	439	370
Average idle time	35%	53%	34%
Average maintenance time	16%	17%	17%
Average time spent on useful mileage	29%	21%	32%
Average time spent on idle runs	20%	9%	17%

weights in the utility function (14). For the sake of simplicity of calculations, we make the following substitutions: $\omega_1 = (1-\omega)^2$, $\omega_2 = (1-\omega)\omega$, $\omega_3 = \omega$. Thus, we will only change ω.

Let us analyze the results presented in Table 2. It is not difficult to see that the largest relative spread of values occurs in average idle time and average time spent on the idle run. Note that relations (18) and (19) used in the utility function (14) assume that we control precisely these characteristics; therefore, the hypothesis that we can control basic characteristics of the movement of locomotives by changing the corresponding weights is confirmed with our numerical experiment, which, in turn, also justifies the use of our proposed utility function. We will consider the solution obtained with $\omega = 0.3$ as the best one presented and use it for further analysis.

Next we show Table 3, "Using locomotives," for the resulting solution. It has eleven lines since the first and eleventh days are incomplete (we consider a period of 12 hours).

It is easy to see that the trace of the locomotive usage matrix is the total number of locomotives used. Note that at the end of the control horizon (in this case, the 11th day), new locomotives are not put into operation, which indicates that the control process has "stabilized"; this gives us a reason to believe that the criterion value obtained is close to the total size of the locomotive park used.

The resulting solution can be compared with real data on the use of locomotives on the MZD section in question. The total locomotive fleet on the Moscow Railways is about 900 locomotives, and about 700 locomotives are used daily. Thus, we can conclude that our resulting solution is approximately 1.5 times better with respect to the minimum number of locomotives. However, it should be noted that the model does not take into account all limitations that occur in real work of a railway.

Table 4 compares three results of solving problem (3): two results obtained using utility function (14) (UF), one of which does not take into account additional constraints imposed on the model as compared to [1], and the result obtained in [1]. Table 4 presents average values of locomotive motion characteristics and the mean value of the main criterion obtained from 100 realizations of each algorithm. The solutions obtained under the same conditions (columns 1 and 3) practically do not differ in the value of the main criterion, the number of locomotives, which confirms that the new problem statement, with the utility function, is adequate. Note that for a certain set of weights in the utility function (14) it is easy to get a solution that completely matches the solution obtained in [1]. Next, we compare columns 1 and 2 of Table 4. Note that the total number of locomotives used has increased by approximately 20%, and as a result, the average idle time of locomotives has also increased, but the percentage of auxiliary (idle) runs has significantly decreased. Considering the ratio of useful runs to auxiliary runs, it is easy to see that in comparison with the solution obtained in [1] it takes a value better by a factor of more than 1.5. Thus, we can conclude that the locomotives work more efficiently, since with a significant reduction in the number and duration of runs we obtain only a slight decrease in useful mileage for each locomotive. When planning for a much longer period than 10 days, the advantage of the new solution obtained using utility function (14) becomes even more obvious.

6. CONCLUSION

We have proposed a mathematical model of assigning locomotives for the transportation of freight trains that takes into account a heuristic utility function. We have developed an algorithm for solving this problem. Results of numerical experiments have shown that variation of weight coefficients in the utility function allows to control the characteristics of locomotive motion and leads only to an insignificant increase in their total number. Comparing with the results obtained in [1], we have confirmed the efficiency of the resulting solutions. The algorithm has exhibited good stability under randomly distributed initial time of passing maintenance, which let us suggest that this approach will also be possible to use with a number of other random factors. However, one should take into account that in practice obtaining such an effect is difficult, since the considered examples have not taken into account a number of restrictions; in particular, we have discarded restrictions on the mass of transported trains and considered only maintenance carried out at certain intervals, disregarding all other types of maintenance. But our results indicate that even under the influence of all these factors the resulting solution will ensure efficient use of the locomotive park.

ACKNOWLEDGMENTS

Results of this work were obtained within the framework of the governmental order of the Ministry of Science and Education, no. 2.2461.2017/PCh, and with the support of the Russian Foundation of Basic Research and JSC “RZhD” as part of research project no. 17-20-03050ofi_m_rzhd.

REFERENCES

1. Azanov, V.M., Buyanov, M.V., Gaynanov, D.N., and Ivanov, S.V., Algorithm and Software Development to Allocate Locomotives for Transportation of Freight Trains, *Bull. South Ural State Univ., Ser. Math. Modell. Programm. Comput. Software*, 2016, vol. 9, no. 4, pp. 73–85.

2. Kibzun, A.I., Naumov, A.V., and Ivanov, S.V., A Two-Stage Optimization Problem for the Activity of a Railroad Transportation Node, *Upravlen. Bol'shimi Sist.*, 2012, no. 38, pp. 140–160.

3. Gainanov, D.M., Ivanov, S.V., Kibzun, A.I., and Osokin, A.V., An Optimal Assignment Model for Locomotives for the Construction of Freight Trains, *Proc. 4th Intl. Sci.-Tech. Conf. “Intelligent Control Systems on Railroad Transportation,”* Moscow: SJC NIIAS, 2015, pp. 45–47.

4. Azanov, V.M., Buyanov, M.V., Ivanov, S.V., et al., Optimizing a Locomotive Park Intended for Carrying Freight Trains, *Proc. 5th Intl. Sci.-Tech. Conf. “Intelligent Control Systems on Railroad Transportation,”* Moscow: SJC NIIAS, 2016, pp. 94–96.

5. Belyi, O.V. and Kokurin, I.M., Organization of Freight Railroad Transportation: Methods of Optimization, *Transport RF*, 2011, no. 4 (35), pp. 28–30.

6. Cacchiani, V., Galli, L., and Toth, P., A Tutorial on Non-Periodic Train Timetabling and Platforming Problems, *EURO Transport. Logist.*, 2015, vol. 4, no. 3, pp. 285–320.

7. Lazarev, A.A. and Musatova, E.G., Integer-Valued Problem Settings for Constructing Railroad Trains and Their Schedules, *Upravlen. Bol'shimi Sist.*, 2012, no. 38, pp. 161–169.

8. Lazarev, A.A., Musatova, E.G., Gafarov, E.R., and Kvaratskheliya, A.G., *Teoriya raspisanii. Zadachi zheleznodorozhnogo planirovaniya* (Scheduling Theory. Railroad Planning Problems), Moscow: Inst. Probl. Upravlen., 2012.

9. Floyd, R.W., Algorithm 97—Shortest Path, *Commun. ACM*, 1962, vol. 5, no. 6.

This paper was recommended for publication by A.A. Lazarev, a member of the Editorial Board