Smart Transport as an Enhancement of the Urban Infrastructure

Ekaterina Zakharova, Inna Minashina, and Fedor Pashchenko

Abstract The concept of a Smart City is to use the existing resources in an optimal way to provide the greatest convenience to its residents. This requires close integration of all components, for example, street video surveillance, public services, intelligent transport systems and others, on the scale of a megalopolis. Every year, the world's megacities are becoming more comfortable for residents due to the introduction of newest technologies. First, such technologies include intelligent control systems in the transportation field. The main goals of Smart Transportation are the efficient and coordinated movement of people, monitoring the location of objects, fast and reliable interaction of vehicles with each other, as well as guaranteeing road safety. This paper represents examples of artificial intelligence technologies and optimization methods applications to create such smart systems.

Keywords Smart city · Smart transport · Artificial intelligence · Vehicle rescheduling problem · Forecasting · Railway traffic

1 Introduction

A smart city is the integration of information and communication technologies for the management of urban infrastructure: transport, education, healthcare, networks serving residential and security. The goal of a smart city is effective governance and a high standard of living for the population using innovative technologies. The smart city infrastructure consists of the following elements:

E. Zakharova (✉) · I. Minashina
Moscow Institute of Physics and Technology, Moscow, Russia
e-mail: zakharova.em@mipt.ru

I. Minashina
e-mail: minashina.ik@mipt.ru

F. Pashchenko
Institute of Control Sciences, Moscow, Russia
e-mail: feodor@ipu.rssi.ru

Y. A. Vershinin et al. (eds.), *Technologies for Smart Cities*,
https://doi.org/10.1007/978-3-031-05516-4_7

- Water supply and energy supply management—Smart housing and communal services
- Waste Management—Smart Garbage;
- Ensuring the mobility of citizens within the city—Smart transport;
- Digitalization and provision of reliable communication;
- Citizens' participation in city management—e-government;
- Environmental protection—control of pollution and noise levels, creation of "green" neighborhoods;
- Safe City—ensuring the safety of citizens;
- Affordable e-education and healthcare—smart healthcare, telemedicine, distance learning.

Smart transport is a common name for all types of vehicles using modern communication technologies for efficient movement of people, location monitoring, the interaction between vehicles and other traffic elements, improving the environmental friendliness of transport and the safety of road use in general [1]. Because of the increasing number of vehicles on the roads, there are problems of congestion and irrational use of road resources, which entail an increase in travel time, the amount of fuel consumed, and emissions polluting the environment.

The main trend of creating smart transport systems is the creation of intelligent automated control systems, which are able not only to find optimal solutions for emerging traffic situations but also to conduct analysis to identify "bottlenecks" of technological processes on the roads.

1.1 Artificial Intelligence for Smart Transport

Trends in the introduction of modern technologies in the field of train traffic control are dictated by changes in the principles of implementing responsible technological processes aimed not just at improving all components of the railway infrastructure and rolling stock, but also at obtaining maximum effect for all participants in the transportation process [2, 3]. The most relevant areas of smart transport development are the active digitalization of archived traffic data, as well as the use of big data and artificial intelligence. More and more intelligent systems are being developed that provide automation and decision support at all stages of the organization of transportation.

In the works [4, 5], the most actual tasks of Smart Transportation are highlighted, such as automation of the main control functions; the using analysis and Big Data to find solutions; the transition to intellectual planning of maintenance; the development of expert systems capable of finding solutions in emergency situations. Most of the technological tasks solved in railway transport are quite complex, and their full formalization is almost impossible. It is almost impossible to automate such processes using only standard analytical methods and linear algorithms. Therefore, there was a need to develop and apply new methods capable of solving these problems. Such approaches, for example, are AI methods and expert systems [6, 7].

2 Practical Methods of Problem Solving in Smart Transport Systems

Improvement of the economic efficiency of any transport system is impossible without solving the problem of automatic train schedule planning and managing the maintenance of traction facilities. In order to ensure the correct operation of the transportation process, it is necessary to create appropriate conditions for the maintenance of the operational fleet, guaranteeing its supply with the required amount of traction resources in the prescribed time. In turn, when planning the maintenance of traction resources, an important condition is the behavior of traffic flows affecting the simulated system. These three tasks are interrelated and underlie the creation of modern intelligent smart transport systems.

2.1 Forecasting Problem of the Traffic Flow

Many studies have been conducted on modelling the traffic flow. Regression models, probabilistic methods and others were used for this purpose. However, the existing methods are no longer so effective due to the lack of accurate models in this area. As an alternative, methods from the field of artificial intelligence can be used to model and predict such flows. They can be considered as more flexible, reliable and suitable for solving problems in conditions of great uncertainty, in comparison with conventional deterministic statistical approaches. In addition, the combination of various artificial intelligence methods and optimization methods can lead to behavior that is more efficient and give greater flexibility when working with real and large-scale tasks.

If the transport flow is described as an incompressible fluid, then the main predicted macro parameters should be:

- Flow is the total number of people on the train per time unit.
- Density is the number of people on each unit of the train.

Based on this, the density of the passenger flow will be the output parameter for our model.

The problem under study as a whole can be formulated as follows: the density of the flow of passengers for the previous period for this train is given. It is required to give a forecast for the density of the flow of passengers in the current time period. The flow density $N(t)$ is the number of passengers on the train at the current time t per unit area. Based on these results, the amount of traction resources required in the current time interval is optimized.

Most control systems of modern technological processes are built using mathematical modelling and forecasting. Many technological processes and systems have important features: complexity, nonlinearity, poor knowledge of the connections within the system, high inertia, the presence of lag and interference, unsteadiness, etc. These uncertainty conditions lead to the fact that the standard deterministic approach to modelling may become unacceptable. One of the approaches to solving this problem is to attract high-quality information [8]. In this regard, fuzzy logic and neural networks will be used along with analytical algorithms to construct the rules of conclusions of the decision-making system. In this paper, a fuzzy difference model TSK is used to model transportation processes-a model built based on the statistical fuzzy model Takagi (Takagi) and Sugeno (Sugeno), called the TS model:

$$R^{\theta}: if\ v_1(t)\ is\ V_1^{\theta}(t),\ \ldots,\ v_{m+l+1}(t)\ is\ V_{l+m+1}^{\theta}(t),$$
$$then\ N^{\theta}(t) = w_0^{\theta} + \sum_{j=1}^{m+l+1} w_j^{\theta} v_j(t),\ \theta = \underline{1,\ n}$$

where V_k^{θ}—fuzzy sets characterized by membership functions $V_k^{\theta}\left(v_k, d_k^{\theta}\right)$.

The advantages of the TS model are as follows: parametric identification can be used for a fuzzy dynamic model; it describes nonlinear dynamic processes with high accuracy; the averaging properties of the output mechanism and the specific type of membership functions provide high noise immunity of the model [9].

In the fuzzification block, the values of variables $v_1(t), v_2(t), \ldots, v_{M+L+1}(t)$ are converted into a matrix

$$V = \begin{bmatrix} V_1^1(v_1(t)) & \cdots & V_{m+l+1}^1(v_{m+l+1}(t)) \\ \vdots & \ddots & \vdots \\ V_1^n(v_1(t)) & \cdots & V_{m+l+1}^n(v_{m+l+1}(t)) \end{bmatrix}$$

In the fuzzy inference block, the value of the truth of the θ-th rule is calculated by the formula

$$\beta^{\theta}(t) = V_1^{\theta}(v_1(t)) \oplus V_2^{\theta}(v_2(t)) \oplus .. \oplus V_{l+m+1}^{\theta}(v_{l+m+1}(t))$$

from where the fuzzy function is determined

$$u^{\theta} = \frac{\beta^{\theta}(t)}{(\beta^1(t) + \beta^2(t) \ldots + \beta^n(t))} +, \theta = 1,\ \ldots n$$

In the defuzzification block, the specific value of the N_t output is determined by the formula (1).

2.1.1 Predicting Traffic Flow Approach

Hybrid approach of model identification is proposed to use structural and parametric identification [10, 11]. Structural identification is the determination of the number of rules and the order of a fuzzy model. Parametric identification—determination of coefficients of linear difference equations, as well as parameters of membership functions. The stochastic approximation method is used as a parametric identification algorithm [12].

Initial data for the algorithm: the number of rules n, the order r, s of the fuzzy model, the value of the coefficient vectors c and the parameters of the membership functions. The criterion for stopping the algorithm is the average modular error:

$$J(c) = \frac{1}{T}\sum_{t=1}^{T}\left(\left|N(t) - \hat{N}(t)\right|/N(t)\right)$$

The operation of the entire hybrid algorithm ends when J < 0.03.

2.1.2 Passenger Traffic Simulation

Training data is needed to start the hybrid algorithm. Input data for the hybrid algorithm: time for prediction t, the flow density at the moment $t - 1$ and $t - 2$, maximum number of standing and sitting places. Membership functions and a set of fuzzy rules are calculated for each of these parameters.

Forecasting the density of passenger traffic over the entire time period of the selected train determines the optimal planning of transportation, as well as the necessary amount of trains and wagons to ensure the transportation process. The train and wagon numbers may vary depending on the magnitude of the flow density.

This model can be modified in the following ways. It is possible to train the system to predict the density of passenger traffic not only by the time periods of the day but also by the days of the week, as well as by various seasons. This is applicable when the flow changes depending on the day of the week and the period of the year. For example, during the summer season, the number of people travelling from the city to the region for the weekend increases dramatically. Therefore, it is worth running longer trains on Friday evening from the city and on Sunday evening to the city.

Thus, the proposed model makes it possible to describe the behavior of traffic flows with sufficient accuracy using intelligent systems. This advantage can be used to solve the traction resources control problem on the transport network. Especially in the case when the transport situation is changing dramatically and it is necessary to find the optimal solution according to some parameters.

2.1.3 Using a Passenger Flow Simulation Algorithm to Regulate the Interval of Movement and the Number of Subways Trains

Planning the schedule for the transportation process is the basis for correct work of all transport systems, including the subway. At the same time, this process always requires compliance with a large number of technological factors, such as ensuring passenger transportation demand, traffic safety, efficient use of the capacity of sections and stations, rational use of traction resources. The main global trends in the development of rail transport systems suggest an increase in the automation of the transportation process. One of the profitable improvements in this area can be calculating the required number of trains and the time of their departure from the depot for efficient transportation of all passenger traffic.

The discussed above method of passenger traffic forecasting allows one to solve the problem of smart predicting the number of trains and their departure time from the depot for optimal covering the entire passenger traffic, taking into account its intensity and unevenness. The implemented in such a way smart transport management complex can automatically coordinate, by means of a schedule, the processes of movement and departure at stations, the delivery, and dispatch of rolling stock from the depot, taking into account the specified pair and passenger traffic on the line.

Thus, the proposed model provides an opportunity to solve the specified problem of regulating the interval of movement and the number of trains, from which we can conclude about the potential of using the built systems in this area.

2.1.4 Providing Railway Traffic with Transport Resources Problem

Planning the Locomotives Fleet Quantity for a Specified Period

To solve this problem, it is necessary to automate the calculation of the forecast locomotive number and build a plan of operations to ensure this at a specified time interval. All calculations are carried out separately for each directorate (except the case of moving locomotives) [13]. Locomotives differ in the type of traction and the type of movement. The following information is received at the input of this module:

- The number of the required locomotive for the period by directorate and type;
- Locomotives with their attributes and current location;
- The railway graph with its parameters.

The main steps to bring the number of the locomotive fleet to the required are:

Stage 1. The number of the required locomotive is calculated as following (given the percentage of defective locomotives for the period)

$$N_{plan} = N(k + 1),$$

where N is the number of the required locomotives, k is the factor accounting for defective locomotives (for the past year).

Stage 2. The forecasted locomotives number is calculated and their state (for example, for which of them will be decommissioned):

$$N_{cur} = N - N_{cons} - N_{dec} + N_{bt}$$

where N is the total number of all locomotives of this directorate at the moment; N_{cons} is the number in unexploited fleet; N_{dec} is the number for decommissioning or transferring into the unexploited fleet; N_{bt} is the number for purchasing.

Therefore, the directorate state is calculated at the period start time, taking into account factors affecting the locomotive number. Further, if there is a deficit in any directorate, a certain list of operations is being taken.

Multi-agent Approach for the Locomotive Relocation Task

To solve this problem, the following algorithm was compiled based on technological experts' experience. The algorithm describes all available activities, sorting and search parameters. The architecture of this system based on multi-agent technologies, in which the following types of agents were defined:

- The planner is a central agent that synchronizes the work of all other agents, manages the sequence of events, and processes input and output information;
- The directorate is regional directorates' agents, which contains information about the locomotive fleet and their technical state, and the behavior of the agent when receiving various commands from the main one.

Since the system has multiagency, each agent in it strives to achieve its specific goal, using the strategies laid down in it to achieve him or her, but at the same time taking into account the main limitations of the entire system. Almost all operations for the shortage or getting rid of the excess are carried out within the directorate. However, there is a locomotive relocation from. This task was reduced to a transport problem and its solution using an auction algorithm.

The Solution to the Transportation Problem for the Locomotives Relocation

The task is formulated as follows: it is necessary to make an optimal plan for the relocation of locomotives from one directorate to another.

To solve it, the algorithm of asynchronous parallel auctions was used under the condition of uniformity of objects, described in [14]. This method determines flows that maximize the overall utility of moving similar objects at a given cost of transporting one unit from supplier i to consumer j, varying flows. Each supplier and consumer is assigned an intelligent agent responsible for storing internal information for him and exchanging information with other agents and a managing agent to coordinate the work of other agents.

Directorates with a deficit are consumers in terms of the auction method, and directorates with excess are suppliers. Let's denote the variables s—flow, u—value of utility function. Consumers contain information about a set of already distributed pairs $\langle s, u\rangle$, upon receipt of which suppliers select locomotive available for relocation. However, there are often cases of "competition" of directorates with a deficit for the same locomotives. In this case, it is necessary to determine where it is more profitable to send this transport unit. This decision is made based on the calculation of the new utility value according to the following formula:

$$u_i^{new} = u_i^{old} + b_{ij} + \varepsilon,$$

where u_i^{new} and u_i^{old} are the new and old values of the utility function, respectively, ε—is an arbitrarily small value, and bij is the difference between the best choice and the choice that is second after the best. This value is constantly growing, and then eventually this process stops. Huang's method is used to estimate the completion phase of the algorithm.

The algorithm main steps are the following:

Step 1. Initialization. Each participant contains information about the pairs < s, u > . At the initial moment of time, there is only one such pair, the utility of which is 0. With changes in the course of the algorithm, the utility and power change, and at the same time this participant sends the rest an updated set of pairs < s, u > indicating its own unique identifier.

Step 2. Formation of applications. If there are unallocated repairs at some point in time, then the process of creating and sending applications to all repair companies takes place. After that, the process of calculating the utility function is carried out, as a result of which flows are formed, sent to repair companies with the selected values of the utility function and repair ids.

Step 3. Processing of applications. Each repair company, upon receipt of such a flow, selects repairs for the distribution of its repair capacity. After that, messages are sent about updating information about the flows and values of the utility function of the assignment for them until either all repairs are placed or all capacities are filled.

Algorithm Testing Results

The result of solving this transport problem is the optimal distribution of locomotives from the directorates with an excess to the directorates with a disadvantage, indicating specific locomotives that allow you to bring the current number of locomotives to the required in the forecast period. Table 1 shows an example of the result of the algorithm for a year.

Table 1 Results of the system of construction and distribution of repairs

Input data			
Region	Required quantity	Forecasted quantity	Deficit/excess
Reg_1	1061	1214	153
Reg_2	1290	1118	-172
Reg_3	1601	1674	73
Reg_4	1664	940	-724
Output data			
Reg_1	1061	1214	0
Reg_2	1290	1118	0
Reg_3	1601	1674	0
Reg_4	1664	940	0

Table 2 Results of the locomotive fleet planning for a specified period

Region	From an unused fleet	Relocation		ESLR	Decommission
		Receiving	Sending		
Reg_1	6	0	293	130	0
Reg_2	2	164	0	6	0
Reg_3	6	0	86	7	0
Reg_4	8	215	0	0	0

The required quantity and the forecasted fleet quantity are estimated in locomotives. The deficit and excess of the fleet is estimated by taking into account the purchase of new locomotives, the decommissioning of old ones and providing an extended service life repair (ESLR). In addition to this, the following information is displayed for each operation performed (Table 2).

In addition, for each event and each directorate, there is an opportunity to view a list of locomotives that need to hold an event, indicating the numbers of locomotives and their parameters. At the same time, the user has the opportunity to change the decision and start recalculation.

2.1.5 Providing the Required Resource Quantity for Transport Traffic Nodes

Due to the large volume of freight transportation by rail, transport companies are tasked with providing such a number of locomotives that are able to realize the transportation plan. The volume change is carried out by transferring units between the operated and non-operated fleet. In addition, an analysis of the condition of each locomotive.

The main goal of solving this task is increasing the economic efficiency of fleet use and transportation planning by automating locomotive fleet planning and management by improving the quality, reliability and systematization of information for management decision-making. The main technological requirements are the following: the selection of locomotives should be carried out according to certain criteria and must take into account their technical condition and current location. In addition, it is necessary to take into account the possibility of manually adjusting the list of locomotives in connection with the necessary repairs. Therefore, the task is divided into three stages.

Stage 1. An initial solution is created in which there is no decommissioning (this locomotive has a special status for manual correction by the user. In addition, an ESLR schedule is planned.

Stage 2. The user adjusts the schedule and manually sets locomotives for decommission or an ESLR.

Stage 3. The schedule is submitted taking into account the immutability of user edits. After that, the graph is checked at the time of the technical feasibility of carrying out an ESLR of the specified locomotives and a decision is issued on the possibility (or impossibility) of carrying it out, after which the final schedule of repairs is displayed.

Locomotive Repairs Schedule Making

When solving the problem of ensuring the necessary volume of traffic, it is not enough just to bring the maintenance of the locomotive fleet to the required number, it is also necessary to maintain each element of this fleet in working condition. To do this, each section of the locomotive must periodically undergo maintenance in accordance with technological standards, and therefore a schedule of repairs must be calculated. This task formulates as follows: it is necessary for each locomotive in the operated fleet to make a schedule for each type of repair, taking into account the adjustments of repairs according to the priority of repairs. All repairs are divided by seniority, each of which has its own frequency; each type of locomotive has its own mileage rate for them.

The capabilities of repair facilities are limited and to ensure the optimal distribution of repairs to the relevant facilities in order to maximize the workload. This task is reduced to a transport task, which in turn can be formulated as follows: it is necessary to distribute the repair schedule to suitable repair facilities, taking into account the type of repair, the locomotive parameters, as well as the maximize distribution efficiency [15, 16]. Each enterprise has the ability to repair a limited set of series for specified repair types and a capacity. In addition, for the specified period, the relocation norm also serves as an additional restriction.

This approach provides automation of the following technological processes:

- Creation senior factory repairs schedule and additional maintenance, which should be carried out independently;

- Creation of an ESLR schedule for locomotives;
- Creation repairs scheduled for various time periods with the indication of detailed information about every repair.

Solving the Task of Planning Locomotives Repairs

To solve this problem an object-oriented software architectural model was implemented with the following objects classes:

- Classes for the sequence of execution of the general algorithm for creating a repairs schedule and drawings up individual repair plans for each locomotive;
- Classes for storing information about locomotives and its current parameters;
- Classes for storing information for all types of repairs for each locomotive;
- Classes for recording and storing all calculated repairs by type.

The Task of Repairs Distribution by Repair Facilities Solving

The task of distributing repairs to repair companies can be reduced to solving a multi-threaded transport problem, which can be formulated as follows. There is a capacitive network consisting of nodes and branches, to which objects are transferred along with branches. The main task is to determine the maximum capacity for each arc. The transport distribution problem with minimal costs is positioned as a linear programming model. A modified auction method is used to solve this problem [17]. The auction method is a parallel relaxation method for solving the classical assignment problem. After calculating all the permissible utility functions, the repair company with the maximum utility function is selected. To reassign a repair its utility function must be higher than the previous one. This leads to the convergence of the algorithm to obtain the final solution.

The algorithm can be represented as follows.

There are N repair facilities and N repairs to distribute the task. For each factory, there is a non-empty subset of $A(i)$ objects that can be assigned to it. Solution S is a set of <repair company-repair> (i, j) pairs such that:

- $j \in A(i)$ for all $(i, j) \in S$;
- each repair company i has at most one pair $(i, j) \in S$;
- each object j has at most one pair $(i, j) \in S$.

The full purpose is a set of N pairs <repair company, repair>. In the context of this assignment S, a repair facility i is assigned if there is such a repair j that $(i, j) \in S$. There is some given integer value a_{ij} in which repair company i is associated with repair j ∈ A(i). It is necessary to find a complete assignment that: maximizes the following function for all completed assignments S:

$$\sum_{(i,j)\in S} a_{ij}$$

Table 3 Results of repair schedule making for the year planning period

Input data	
Sections	5017
Repair facilities	570
Repairs type systems	Old and new
Repairs entered by the user	0
Output data	
Required repairs quantity	18,614
Distributed repairs quantity	17,205
Undistributed repairs quantity	287

Table 4 Results of repair schedule making for the 10 days

Input data	
Sections	5056
Repair facilities	570
Repairs type systems	Old and new
Repairs entered by the user	10
Output data	
Required repairs quantity	486
Distributed repairs quantity	392
Undistributed repairs quantity	72

The auction algorithm is used here, which solves the dual-use problem: minimize the function under the condition. A detailed description of the algorithm of auctions and dual price is given in the articles [16].

Results of Planning Repairs Schedule

The results of the work of this module are the schedule of repairs, indicating the start and end dates of repairs, the repair facilities, the type of repair, the ID of the locomotive. Consider as an example the solution obtained based on the following data for annual planning (Table 3).

Table 4 shows the solution for the ten days period.

At the same time, in addition to obtaining total indicators for repair needs, as well as for occupied and unoccupied capacities, the user can view the schedule of repairs for each locomotive on the form.

2.2 *Freight Scheduling Problem*

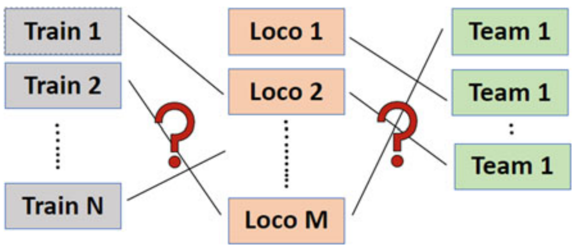
While managing the transportation process the task of finding an effective timetable and providing trains with traction resources is decisive. Figure 1 illustrates the main idea of the freight scheduling problem.

Having received the generated train routes, locos and locomotive crews' parameters, the system should build a list of trains that need traction and a list of available locomotives and crews. For trains from the first list, it is necessary to find the optimal attachment of locomotives and crews from the other two lists. In other words, an assignment that provides the best implementation of the railway transportation management problem and complies with specified restrictions and technical regulations in relation to traction resources.

Freight scheduling problem includes three sub problems: train scheduling problem, locos assignment problem and locomotive crew assignment problem. Here, for brevity, we will dwell in more detail on only one of them—the locos assignment problem.

Figure 2 illustrates the problem of assigning locomotives to trains.

Fig. 1 Main idea of freight scheduling problem



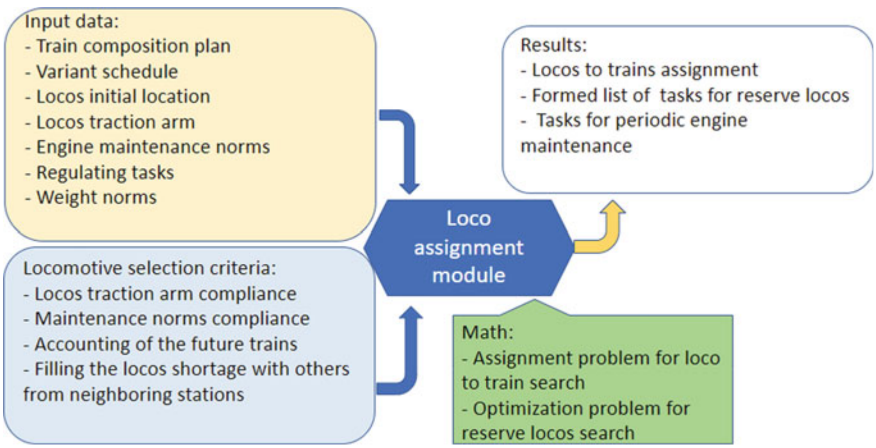


Fig. 2 Problem of assigning locomotives to trains

The specified data is supplied to the module input. The task is to assign locomotives to trains in the best possible way according to the specified criteria and constraints.

2.2.1 Formalization of the Locomotives Assignment Task

The technological process of transportation entails many restrictions that are imposed on the train and traction resources schedule generation rules. To take into account these limitations and features of the railway area, a task formalization method was developed, which formed the basis for creation of a mathematical model of the described control system.

Formalization of the task includes building a model of the freight management system, forming a utility function and formulation of an optimization problem for an effective train and traction resources scheduling.

Multi-agent Transportation Management System

Due to a large number of technological limitations in cargo transportation planning, it was necessary to develop a system that would take into account and comply with all requirements and constraints of the described task. For this purpose, a multi-agent control system (or "planner") was proposed. It consists of the following components:

- The execution environment, which interacts with external systems, receives all the input data from them and sends the received data to the main agent.
- Agents:
  - Station agents, which contains their parameters, data about neighboring stations, the number of available locomotives and locomotives teams for each hour of the planning period. Restrictions for them:

    The trains number at the station should not exceed the number of station tracks at any given time
    The locomotive's number and locomotives teams at the station should not be negative at any hour

  - Line sections agents. Two agents are created for each line (in the forward and reverse direction). The line agent knows the transfer capacity for each hour on the planning horizon. Restrictions:

    The number of trains sent to the line every hour should not exceed the capacity value for this hour.

  - Train agents. Agent's goals:

    Get departure and arrival times on each included in the route line.
    Minimize the dwell time of the train on intermediate stations of the route.

  - Loco agents. Restrictions and goals:

The railway traffic norms are should not exceed
Minimize the loco's journey time without a train.
Minimize locomotive downtime at stations.
Maximize the effective work of the locomotive

– Team agents. Its similar restrictions to locomotives.

To solve the entire planning period of the railway modeling problem, the system was decomposed into 4 subsystems:

- The Volumetric Train Planner operates with train, loco, team and stations. This part of the system splits the initial number of trains into hourly intervals, taking into account the restrictions of all agents.
- The Volumetric Loco Planner includes locomotive and station agents that eliminate conflicts related to the lack of locomotives.
- The Volumetric Team Planner is similar to the previous one.
- Object-by-Object Planner contains station, train, locomotive and team agents. It binds trains to the train schedule paths, assigns locomotives, and crews to their abstract transposition plans (which are the results of the work of the described above planners). Optimization of the locomotives (teams) assignment is the main goal of this work. It is a subtask of the general freight scheduling problem that the optimization algorithms are aimed at.

The Utility Function Definition

The key question for taking into account technological constraints and formulating train scheduling tasks as an optimization problem is the utility function formation. The implementation of the railway control system is formed on its basis. The main criteria of the utility functions are given in Table 5. A normalized numerical value *uk* is assigned to each criterion. After that the value for the utility function of the pair <locomotive, locomotive slot (an abstract plan for moving a locomotive between stations)> is calculated as $U = \sum_k c_k u_k$, where c_k is the weight of the corresponding k-th criterion.

Thus, the evaluation rules for the described above subtask solution efficiency were formed. After the Object-by-Object Planner has compiled a list of available slots and locomotives, one can proceed to the heart of planning—solving the assignment of traction resources to the appropriate slots task. This problem should be formulated as a general mathematical optimization problem with constraints.

Reducing the Task of Attaching Traction Resources to Trains to the Optimization Problem

Above, we have considered the type of utility functions for pairs <locomotive slot, assigned to it locomotive> . Considering the optimization of the trains to traction resources attachment in the entire planning period, cumulative utility function was

Table 5 Information on the calculation of utility function criteria for the locomotives to a trains linking task

№	Criteria name	Description	Sign
1	Working time	The difference between the arrival time at the terminal station of the slot and the departure time from the initial station	+
2	Locomotive waiting time	The difference between the departure time of the slot and the arrival time of the locomotive at the departure station. If this difference is negative, then take the value of the criterion = 0	–
3	Slot waiting time	The difference between the arrival time of the locomotive at the departure station and the departure time of the slot to the first section. If this difference is negative, then take the value of the criterion = 0	–
4	Time till service station	The time of the engine relocation from the end station of the train route to the nearest station providing the required service type	–
5	Time until the next periodic engine maintenance	The operating time remained until the next periodic engine maintenance at the moment of arrival at the terminal station of the slot. If the value is <0, then the value of the criterion is set to $\boxed{\text{OBJ}}$	–
6	Weight restrictions	Difference between maximum permissible weight of the train that this locomotive can carry on this route and the weight of the train corresponding to the locomotive slot. If the value is <0, then the criterion value is set to -infinitive. Further calculations on this branch are not required	+
7	Time to transfer the locomotive by reserve	The travel time between the station where the locomotive is currently located and the station from which the locomotive slot begins	–
8	Future plans	Consideration of possible "future" slots from the terminal station	+

used, which is equal to the sum of all objective functions of constituents locomotive slots in the whole schedule. Thus, the total objective function is a multidimensional function of the vector argument, where vector—any solution of the locomotives to slots assignment problem, presented in the form, where—the number of the locomotive (crew) assigned to the i-th slot,—number of slots in the current planning period.

Then, in order to solve the optimal railway scheduling problem, it is necessary to solve the general optimization problem of finding the extremum of the above-mentioned function with given constraints:

$$F(X) = \sum_i U_i(x_i) \to extr(max);\ X^* = \arg extr F;\ r_j\left(X^*\right) \le b_j,\ j = 1, \ldots, l$$

where X^*—the desired solution; $r_j(X^*)$—given set of constraints; l—number of constraints.

The search is performed on a set of constraints related to technological features in the railway complex. Then it is necessary to find such a kind of iterative operator for finding the extremum of the function F(X) to make the iterative process converge to the best solution or, in other words, to the extremum of the total utility function.

2.2.2 Solving the Problem of Building an Effective Schedule

The assignment problem is a special case of a transport problem. It, in turn, is a special case of a linear programming problem. Such problems are solved using classical methods, but using a specialized method gives faster results.

Hungarian Algorithm

The most common way to solve the assignment problem is the "Hungarian algorithm", which is a combinatorial optimization algorithm.

It was proposed in 1955 and its computational complexity is, where n—is the number of tasks and employees (the number of jobs must be equal to the number of employees). The algorithm can be modified up to the level of complexity. The Hungarian method was implemented to solve the problem of assigning traction resources to trains, but its speed turned out to be unacceptable. The results of this experiment are discussed in more detail below. Therefore, another algorithm was found, that is, solving the assignment problem using auctions.

Auctions Algorithm for Assignment Problem Solution

D. Bertsekas proposed the solution of the assignment problem using auctions. The main ideas of the auction algorithm are described in [14]. In relation to our problem, the steps of the algorithm can be described as follows [15, 16]:

Step 1. The best slot j is selected for the detached locomotive i:

$j_i = \arg \max\limits_{j \in A(i)} \left(U_{ij} - p_j\right)$,where p_j—-th slot price, $A(i)$—set of slots allowed for the i-th locomotive.

Step 2. The second-highest value of the "benefit" from the assignment of the locomotive to the j_i-th slot is calculated:

$$\omega_i = \max_{j \in A(i), j \neq j_i} U_{ij} - p_j$$

If except j_i, there are no other possible locomotive assignments, then $\omega_i = -\infty$.

Step 3. Next, the new price for the slot j is calculated:

$$p_j = \max\{\lambda; U_{ij} - \omega_i + \varepsilon\}$$

where λ—the threshold value (constant) below which it is forbidden to set the price, ε—infinitesimal number of the order 1/N.

Step 4. If $\lambda \le u_{ij_1} - \omega_i + \varepsilon$ then locomotive i is assigned to the slot j, and the previous assignment to slot j is reset.

Steps 1–4 are repeated until all locomotives are assigned to locomotive slots.

A rigorous proof of the convergence of this algorithm, as well as a justification for the choice of values and ε is given in the article [13].

This method allows you to quickly get a solution, but it is not able to meet a number of planning criteria, as well as to solve the optimization problem in compliance with the maintenance percentages of crews and reserve locomotives. The mentioned criteria include:

- Obtainment of several solutions with specified accuracy for selection of the best according to additional technological criteria;
- Solving a multidimensional assignment problem for the simultaneous solution of all three subtasks of the freight scheduling problem;
- Adjustment of the resulting solution accuracy.

Simulated Annealing Method

Due to the inability of the auction algorithm to satisfy the listed above criteria, there is a need to develop another method of multidimensional optimization that solves the described situations. As a trade-off, the simulated annealing method was chosen. In general, the algorithm for assigning traction resources to trains with the method of simulated annealing can be represented as follows:

- The initial temperature and minimum temperature are initialized together with the temperature change function.
- The first locomotives assignment is chosen randomly and its overall utility is calculated.
- As long as the temperature is greater than the minimum value, the new state and its overall utility are calculated. If the utility of the new state is greater than the utility of the current one, then it switches to the new one. If the new state has a lower utility, the transition is carried out with a certain probability, depending on the temperature. For example: $P = exp(-\Delta t), where \Delta = F_1 - F_0$—is the difference between the total utility of the new and current states, and t is the current temperature.
- Transition to a new temperature according to the temperature change function.

To obtain a new state from the current one, the following algorithm was proposed: the coordinate of the utility matrix is randomly selected and any two assignments at this coordinate are swapped.

The peculiarity of the considered optimization task is the diversity of the utility function, which appears from launch to launch at various planning steps. The described method of simulated annealing, based on a random approach, copes well with cases of non-unimodal utility function, and allows you to meet the additional planning criteria described above. On other hand, its main disadvantage is a long operating time, which can be reduced by adjusting the appropriate parameters. In this regard, it is convenient to use a hybrid of the two methods described in this paragraph, using each of them for the purposes corresponding to their capabilities.

Genetic Optimization Algorithm

The optimization problem is reformulated as the problem of finding the maximum of some function $f(x_1, x_2, ..., x_n)$, called the fitness function. It is necessary that on a bounded domain of definition $f(x_1, x_2, ..., x_n) \geq 0$, while continuity and differentiability are not required. A string of bits encodes each parameter of the fitness function. An individual will be called a string that is a concatenation of strings of an ordered set of parameters $|x_1|x_2|...|x_n|$. The universality of GA lies in the fact that only the fitness function and coding of solutions depend on a specific task. The remaining steps for all tasks are performed the same way.

With the help of the fitness function, among all individuals of the population, the most (crossing capable) and the least (removed from the generation) adapted individuals are distinguished. Thus, the fitness of the new generation is on average higher than the previous one. The final solution to the problem is the fittest individual of the last generation with the greatest fitness function.

The algorithm step consists of the following stages:

Stage 1. Setting initial states and parameters. Before starting the GA, the program receives a list of stations, locomotives, and a set of volumetric plans for the purpose of locomotives. In addition, to start the GA, the following parameters are set: the maximum population size P, the number of populations M, the number of the worst individuals D.

Stage 2. Creation of the initial population. At the start of the generation of the initial parent population of each individual, all agents of stations and locomotives are created based on the starting information indicating the number of the current population $Pn_{Id} = 0$.

To create one individual of the parent population, the following mechanism is used. The system receives an input volume plan for the transfer of locomotives, train routes formed with the threads of the variant schedule. The route lists all stations in the order of route. When linking trains to threads, there are the following features:

- The train can change the threads during the movement.
- Different trains can travel on the same line in different sections.

For each volume plan, the algorithm replaces the planned departure and arrival times with real ones taken from the detailed route of the train. Then there is an object-by-object assignment of real locomotives. For each time at least one locomotive is planned to depart, a departure stations list is compiled and the evaluation function of all existing locomotive-plan combinations at this station is calculated.

The evaluation function has the following form:

$$u(Plan, Loco) = C_1\frac{T}{N_1} + C_2\frac{D}{N_2} + C_3\frac{S}{N_3}$$

where $Plan$ is the identifier of the binding plan, $Loco$ is the locomotive identifier assigned to this $Plan$, C_1, C_2, C_3 are configurable coefficients, N_1, N_2, N_3 are normalization coefficients, T is the time remaining until the next repair, given the time of the relocation to the repair facility, D is the distance remaining until the next repair, given the distance of the relocation to the repair facility, S is the coefficient of correspondence between the train route and the traction arm of the locomotive. The selected locomotives are assigned to the plans, their location and parameters are updated according to the forwarding plan for further work.

If there is no one locomotive suitable for the train according to the restrictions, a return to the previous time step occurs and the assigned task is solved again.

For the found solution the following assignments matrix is created:

$$M_{Pn_{id}} = (m_1(Plan_1, Loco_1, PnId, A_1)\ldots, m_i(Plan_i, Loco_i, PnId, A_i))$$

where $Plan$ is the locomotive forwarding plan, $Loco$ is the locomotive assigned to the $Plan$ (if no suitable locomotive is found Loco $= -100$), $PnId$ is the population ordinal number, A is the identifier of assign (if the locomotive is assigned it is equal to 1, otherwise—0).

Thus, one of the initial individuals of the population is created. For each individual, the following total utility function is calculated:

$$F_i = \sum_{i,j} u_{i,j}$$

Stage 3. Crossover. Two parents are selected from the starting population—P_1 and P_2. The child C_1 is generated in the following way:

$$m_i{}^{C_1} = \{m_i{}^{P_1}, if\ An_{P_1} = 1\ \ m_i{}^{P_2}, if\ A_{P_2} = 1 \text{ и } Loco_{Plan}^{P_1} \neq -100\ m_i{}^{P_1}, if\ A_{P_2} = 0\ and\ A_{P_1} = A_{P_2} = 0$$

The usefulness of each individuals is calculated as follows:

Table 6 Comparison of the Hungarian method and the auction method

Quality indicators	Hungarian algorithm	Auction method
Run time (simulation of the assignment task on AgentSpeak)	3–4 min	20 s
Complexity	O(n3)	Nlog N

$$U_{C_1}(Loco, Plan) = \{U_{P_1}(Loco, Plan), if\ m_i{}^{C_1} = m_i{}^{P_1}\ и\ A_{P_1} = 1\ U_{P_2}(Loco, Plan), if\ m_i{}^{C_1} = m_i{}^{P_2}\ and\ A_2 = 1\ 0, if\ A_{P_1} = A_2 = 0$$

that is, if $m_i^{child\_1}$ was copied from the genotype of any parent, the usefulness of a pair $(Loco_i, Plan_i)$, selected based on the genotype of a parent was used for this.

Stage 4. Mutation. With some probability, a mutation may occur in the population. The parameter TPlanId is time, which is randomly determined for each individual. The TPlanId redesigns assignments for plans.

Stage 5. Selection and removal of the worst individuals. After all the individuals for a given population have been generated, the planner selects a given number of the worst individuals in terms of a utility function from all the individuals and then removes them from the offspring.

Stage 6. Stop criterion. The generation of new populations stops if one of the following stopping criteria is reached: the final population size is reached or the improvement of the utility function with a new individual does not exceed the specified.

Stage 7. Final selection of the population. The best individual with the maximum utility function is selected.

Results

Comparison of the Hungarian Algorithm and the Auction Method for Solving the Assignment Problem

Comparing the proposed approach to solving the assignment problem by the auction method with the application of the Hungarian algorithm, it can be noted that the auction algorithm shows a significant increase in the convergence rate compared to the Hungarian algorithm: 20 s instead of 3–4 min (these time indicators should not be evaluated in absolute terms, since this study was conducted in the AgentSpeak language, which itself is a rather "slow" language) [17].

This time estimate shows that the auction algorithm converges in $Nlog(N)$ time, which is a significant improvement over the $O\left(n^4\right)$ Hungarian method (Table 6).

Despite the noticeable advantage in speed, the auction method in some cases is powerless to fulfill the mentioned above additional planning requirements. Let's

Table 7 Comparative analysis of the auction method and the simulated annealing method

Criteria	Auction method	Simulated annealing
Capability to return several solutions	No	Yes
Working with "elongated" matrices	Good	Time costs increasing
Working with sparse matrices	Manual verification is required for some special cases	Time costs increasing
Run time	36 s	3 min
Scalability of the dimension of the problem	No	Yes, but also time costs increasing
Accuracy adjustment	No	Yes

consider the results of another optimization algorithm that would take into account these requirements.

Comparison of the Auction Method and the Simulated Annealing Method

To evaluate the characteristics of these algorithms, a series of experiments was carried out, in which the following criteria of their work were considered: run time, capability to return several solutions, working with sparse matrices, adjusting accuracy, the possibility of increasing the dimension of the problem, working with "elongated" matrices [18]. A comparative analysis of these characteristics is shown in Table 7.

The test data used for the analysis are close to the real traffic and distribution of transportation resources on the Eastern Range of the Russian Railways network. The traction resources to trains assignment was carried out on the planning horizon in 24 h.

On this data set, the freight transportation control module was executed twice: in the first run, the auction algorithm performed the assignment of traction resources, in the second by the annealing simulation algorithm. The main advantage of the auction algorithm is its operating time. As can be seen from the table, according to the carried out tests, planning ends almost four times faster with auctions than with the simulated annealing method. The auction also shows better performance on datasets, where the number of locomotives significantly exceeds the number of trains. On the other hand, the simulated annealing method allows high flexibility in adjusting the accuracy of the final solution. If the set of acceptable solutions includes a number of solutions with very similar total utility, the simulated annealing method can be run with adjusted parameters in order to reduce the number of iterations, which leads to a significant reduction of the execution time without loss of optimality of the solution. This is possible, for example, when there are several assigned teams that come to the depot of the registry at about the same time.

The main advantage of the simulated annealing method is that it can be easily extended for a multidimensional task. For example, all three subtasks of the freight

Table 8 Simulation data. Simulation data and work results (P is the maximum size of one population, M is the maximum number of populations, D is the number of deaths per generation)

	Without GA	Genetic algorithm			
Parameter		P = 15, M = 10, D = 4 with F	P = 15, M = 20, D = 4 with F	P = 20, M = 10, D = 4 with N_p	P = 15, M = 20, D = 4 with N_p
Summary F	139,440	335,760	414,960	−109,920	−109,920
N_p	27	14	13	47	47
Working time	1 m, 13 s	15 m, 24 s	10 m, 31 s	19 m, 13 s	9 m, 29 s

scheduling problem can be performed simultaneously using this method (four-dimensional assignment). The disadvantage of the simulated annealing method is that the scheduler's operating time increases significantly in the case of an dimension increase (rough estimate:O(nm), where n is the maximum dimension for one of the coordinates, and m is the number of coordinates). For this reason, one has to introduce additional heuristics to achieve an acceptable operating time.

Taking into account the described characteristics of the proposed algorithms, it is convenient to use a hybrid solution of the assignment problem using auction and simulated annealing methods, applying each of them in accordance with the current set of requirements.

Results of Attaching Traction Resources Problem Solving Using GA

As a simulation, a railway section was selected, including 566 stations, 772 volumetric locomotive plans, 469 train plans, 4723 train route lines, 783 real locomotives with their parameters, 44 real trains.

The standard planning mechanism has the following features: all utility functions for pairs $(Plan, Loco)$ are calculated several steps ahead, then the total utility function is calculated relative to this depth and the option with the highest value of this function is selected. Table 8 shows the results of the algorithm with the optimization criterion by F—by the utility function or by Np—the total number of assignments.

From the simulation results, it seems that the locomotive binding plan changes depending on the choice of the parameter for which optimization was carried out.

3 Conclusion

A smart city is a combination of smart management systems and technologies for managing various areas of activity of an urban agglomeration. The main task of creating such a smart city is to improve the quality of life of the population using the latest technologies and methods. Smart transport is one of the main directions of a smart city responsible for the movement of people, optimization of the interaction between vehicles and efficient use and change of transport infrastructure. The tasks that arise in this area are characterized by the complexity of their formalization and solution, often associated with increased security requirements. To solve them, the main trend today is the development of intelligent automated control systems. capable of not only finding optimal solutions for emerging emergency situations but also providing analysis for the premature identification of "bottlenecks" in the technological processes of management and on the roads.

References

1. Valdez AM, Cook M, Potter S (2018) Roadmaps to utopia: tales of the smart city. Urban Studies 55(15):3385–3403
2. Nikitas A, Michalakopoulou K, Njoya ET, Karampatzakis D (2020) Artificial intelligence, transport and the smart city: definitions and dimensions of a new mobility era. Sustainability 12(7):2789
3. Rjab AB, Mellouli S (2018) Smart cities in the era of artificial intelligence and internet of things: literature review from 1990 to 2017. In: Proceedings of the 19th annual international conference on digital government research: governance in the data age, pp 1–10
4. Thaduri A, Galar D, Kumar U (2015) Railway assets: A potential domain for big data analytics. Procedia Comput Sci 53:457–467
5. Reinhold S, Laesser C, Bazzi D (2015) The intellectual structure of transportation management research: A review of the literature
6. Pashchenko FF, An BT, Hieu TD, Pashchenko AF, Van Trong N (2020) Intelligent technologies in decision-making support systems. In 2020 international conference engineering and telecommunication (En&T), pp 1–4
7. Pashchenko AF, Pashchenko FF (2013) Smart technologies for solving the transport problems of the city. In: Proceedings of the seventh international conference management of the development of large-scale systems (MLSD'2013), vol 2, pp 141–144
8. Van Gerven M, Bohte S (2017) Artificial neural networks as models of neural information processing. Front Comput Neurosci 11:114
9. Pashchenko FF (2007) Introduction to consistent methods of system modeling. Finance and Statistics, Moscow
10. Pashchenko FF, Minashina IK, Zakharova EM (2013) Neuro-fuzzy modeling of passenger flows, management of large-scale systems development (MLSD'2013). V.A. Trapeznikov Institute of Management Problems, 2:144, ISBN 978-5-91450-138-6
11. Pashchenko FF, Kuznetsov NA, Minashina IK, Zakharova EM (2013) Using relaxation algorithms for estimating parameters of neuro-fuzzy models. In: 4th international conference ICDQM-2013 proceedings. Belgrade, Serbia, 94–100
12. Wasan MT (2004) Stochastic approximation. Cambridge University Press, 58
13. Zakharova EM (2018) Development of planning and management algorithms in timetable tasks in railway transport. MIPT, Dolgoprudny
14. Bertsekas DP, Castanon DA (1989) The auction algorithm for the transportation problem. Annal Oper Res 67–96
15. Pashchenko FF, Pashchenko AF, Kuznetsov NA, Minashina IK, Zakharova EM (2017) Analysis of the adaptive algorithms behaviour applied to the railway optimization problems. In: The 8th international conference on ambient systems, networks and technologies, vol 10. ANT 2017, Madeira, pp 560–567
16. Pashchenko FF, Takmazian AK, Kuznetsov NA, Minashina IK, Zakharova EM (2017) Intelligent control systems for the rolling equipment maintenance of rail transport. In: AICT2017

11th international conference on application of information and communication technologies, Moscow

17. Kuznetsov NA, Pashchenko FF, Ryabykh NG, Minashina IK, Zakharova EM, Tsvetkova OA (2015) Implementation of train scheduling system in rail transport using assignment problem solution. Procedia Comput Sci 63:154–158
18. Kuznetsov NA, Minashina IK, Ryabykh NG, Zakharova EM, Pashchenko FF (2016) Design and comparison of freight scheduling algorithms for intelligent control systems. Procedia Comput Sci 98:56–63