

Approaches to Automated Traffic Optimal Control System

E.A. Sofronova

Abstract—The paper addresses the problems of optimal control and synthesis of traffic flow control in urban areas. It is proposed to solve these problems within the framework of an automated control system used to create an intelligent transport system (ITS) for large cities. Data on the quantitative characteristics of the flow is obtained from road infrastructure detectors. The road network is described by a directed graph, where nodes correspond to road sections and edges correspond to manoeuvres at intersections. The graph has a variable structure depending on the control. The control is determined by the duration of the traffic light phases at signalised intersections. A universal recurrent model of traffic flow, based on the theory of controlled networks, is used to describe the control of traffic flow. Optimal control problem statements are given for different phase switching modes: within a fixed cycle, without a fixed cycle, within a multicycle. The solution of the multicriteria optimal control problem is given. The use of evolutionary algorithms is proposed to solve the optimal control problem. Then, a control synthesis statement for the traffic flow is proposed. The traffic flow is controlled by selecting the duration of the working phases of the traffic lights with respect to the state of the object. The task is to find a traffic light control function that depends on the state of the traffic flow. To solve the control synthesis problem, it is proposed to use modern numerical methods of symbolic regression. Numerical solution of multicriterial optimal control problem for intersection with real field data is given.

Keywords—Optimal control, control synthesis, traffic control, traffic flow model

I. Introduction

Traffic signals are designed to order traffic flows in efficient way so that to ensure safety of all participants in the road network. There are different types of traffic signals, some comprehensive reviews are [1], [2]. Among them there is a traditional fixed-time (FT) [3] traffic signal that uses a timer to change at predetermined intervals, instead of changing according to traffic movements. An organised and predictable traffic pattern is obtained and installed by using an electro-mechanical signal controller to ensure the signal changes according to traffic engineer decision. The main advantage of fixed-time traffic signals is that the costs and maintenance efforts are relatively low. At the same time they may sometimes result in longer delays, but generally they benefit for urban areas with more or less constant and heavy traffic.

Pre-timed traffic signals are more flexible in comparison to FT since the signals may change during the day depending on the stats information, but still perform not effectively in case of some events, for example concert with many visitors

coming in limited time interval, that had not been considered while design of traffic signals.

Modern cities are equipped with road infrastructure that makes it possible to obtain estimates of the state of the network at any time. The main types of road infrastructure recording devices include radar detectors, loop detectors and video cameras. The use of various detectors makes it possible to obtain necessary characteristics of traffic flows, both at intersections and arteries.

Alternatives to the traditional approaches are actuated and adaptive traffic signal control. An actuated traffic signal changes using information on detected vehicles from sensors such as inductive loops embedded in the road or video and non-video sensors. It allows traffic move more quickly and efficiently instead of unnecessary time loss. It works best in suburban or rural areas where traffic patterns change intensely during the day. The drawback of actuated traffic signal is the costs for control systems development and maintenance of sensors as well as for service.

Examples of adaptive traffic signal control (ATSC) are SCOOT (Split, Cycle, and Offset Optimization Technique), [4], [5], and SCATS (Sydney Coordinated Adaptive Traffic System), [6], that optimize the timing of traffic signals at intersections based on real-time traffic patterns.

Then came methods based on self-organisation of traffic lights (SOTL) [7] and maximum pressure (MP) [8]. The self-organisation method involves switching traffic light phases based on data about the current state of traffic in the network and the implementation of certain rules. The maximum pressure method is based on determining the difference between the number of vehicles on exit roads and the number of vehicles on entering roads. When switching phases, the phases with the highest pressure are given priority. The MP method is characterised by high performance, but at the same time it lacks flexibility, as the duration of the traffic lights is a fixed length interval.

Later, a new adaptive approach, reinforcement learning (RL), emerged [9], [10]. The reinforcement learning methods use Q-learning tables. The state of flows at an intersection is represented discretely. Then deep reinforcement learning (DRL) emerged to handle discrete and continuous states at intersections [11], [12].

The agent (object) is an intersection. Each phase of the traffic light is associated with a specific manoeuvre. The phases are switched in a certain order. The task is to dynamically adjust the duration of the phases. The state of the agent (the number of vehicles entering the intersection) is stored in a matrix. The agent's action is the set of all phases in the cycle. The definition of the reward (quality criterion) is a key in this method. For example, the reward can be the difference between the total waiting time for

Статья получена 5 июля 2024 г.

Elena Sofronova, Federal Reserach Center "Computer Science and Control" of the Russian Academy of Sciences, (email: sofronova_ea@mail.ru).

a given number of arriving vehicles in the previous and the next cycle or time interval. The goal of the agent is to maximise the reward. At each point in time, the agent observes the state of the intersection and then chooses an action according to the strategy. The action is a set of phases in the cycle and the efficient phase is activated first. After taking the action, the agent receives a reward and the state of the intersection changes. By reacting to different scenarios, the agent gradually learns to get better rewards.

Some deep reinforcement learning methods show their effectiveness over traditional methods [13], [14]. The advantages of both SOTL and MP methods have been combined in the [15]. However, there may be situations where they become unstable. Double Q-learning network [16], dueling Q-learning network [17], and prioritised experience replay [18] have been proposed to improve the performance of such methods.

The approaches studied in the paper for finding optimal traffic signal control programs at intersections relate to pre-timed traffic signal control, where the phase durations are “fixed” according to historical data from the network under consideration. This work is focused on creating an automated traffic flow control system. The problem of traffic flow control is considered as a mathematical control problem of a dynamic object.

The road network is divided into sections that are characterised by a maximum vehicle capacity and a numerical assessment of their current state, expressed as number of vehicles of average size. Traffic flow in an urban road network is controlled by switching the phases of traffic lights for a given duration.

The solution of the optimal control problem in the classical statement [19] requires an adequate model of the control object. When it comes to transport flows, there are many mathematical models to describe them, see [20], [21], but not all of them are suitable for numerical solution of the optimal control problem. The purpose of this paper is to formulate a number of approaches to solve the traffic flow control problem at regulated intersections using universal recurrent traffic flow model (URTFM). This model was proposed in [22] and then enhanced in [23], [24], [25]. URTFM mathematically describes the processes occurring in the road network during traffic control at controlled intersections. In the classical statement, the optimal control is sought as a time function and is called a program control. Optimal control problem is solved taking into account existing constraints on control, flows, etc. Depending on the particular task different control modes may be implemented as well as a multicriterial optimization.

Another approach is to solve the traffic flow control problem as a synthesis problem, where it is necessary to find a control as a function of the state space vector. By solving the synthesis problem, one may take into account unforeseen changes in the traffic flow. For example, in case of an accident or bad weather conditions, a sudden change in the state of the object occurs. The resulting control function will ensure that an optimal solution is obtained without additional computations, while the program control obtained by solving the optimal control problem can only be computed for limited and predetermined changes in the flow.

Combination of these approaches will automate the process of traffic light control and thereby reduce the

load on operators of traffic management centers, solve the optimization problem and improve the quality of control in accordance with selected criteria, ensure coordinated control in the considered networks, and increase the level of safety.

The rest of the paper is organized as follows. The optimal control problem including some possible phase switching modes is presented in Section II. The multicriterial optimal control problem is given in Section III. Section IV contains control synthesis problem. Certain numerical methods are proposed to solve the considered problems. Computational experiment of solution of multicriterial optimal control problem for intersection with real field data is presented in Section V.

II. Optimal Control Problem

This study proposes an approach to traffic flow control in the considered network by determining the coordinated optimal phase durations of the traffic lights based on the information obtained from video cameras and inductive loop detectors located directly at the intersections and from radar detectors located on the adjacent roads. The current coordination plan is taken as a basis and improved by evolutionary methods, taking into account the flow data obtained over a certain time interval. The coordination plans are calculated for specific days of the week and times of the day.

A. Problem Statement

The road network is described by a directed graph whose nodes correspond to road sections and whose edges correspond to manoeuvres at intersections. To describe the control of traffic flows, a universal recurrent traffic flow model is used, which is based on the controlled networks theory, see [22], [25]. In general, the model is a system of recurrent finite-difference equations

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{f}(\mathbf{x}(k), \mathbf{A}(\mathbf{u}(k))) + \delta(k). \quad (1)$$

The components of the state vector $\mathbf{x}(k) = [x_1(k) \dots x_L(k)]^T$ are quantitative estimations of traffic flow in average vehicles at all road sections L of the network.

The state vector is limited by the maximum number of vehicles that can simultaneously be on the section i

$$x_i(k) \leq x_i^+, \quad i = \overline{1, L}. \quad (2)$$

All road sections of the network are divided into entry, I_0 , exit, I_1 and internal, I . Suppose, that the value of entry traffic flow is known at each time step

$$x_{i_r}(k) = x_{i_r}(k-1) + y_{i_r}(k), \quad (3)$$

where $y_{i_r}(k)$ is the value of the entry flow on the road section i_r , $i_r \in I_0$, $r = \overline{1, L_0}$, L_0 is the number of entry road sections.

The number of vehicles on the exit road sections is not limited

$$x_{i_q}^+ = \infty, \quad i_q \in I_1, \quad (4)$$

$q = \overline{1, L_1}$, L_1 is the number of exit road sections.

Time is discretized in control steps k , $k = \overline{1, K}$, K is a given number of control steps. The change of traffic flow at

each control step is performed by function $\mathbf{f}(\mathbf{x}(k), \mathbf{A}(\mathbf{u}(k)))$, that depends on the state vector and configuration matrix $\mathbf{A}(\mathbf{u}(k))$. Configuration matrix is an adjacency matrix of the graph that presents the road sections (nodes) and maneuvers between them (edges) that are permitted by current traffic light phase. In addition, the traffic flow changes by input flow $\delta(k) = [\delta_1(k) \dots \delta_L(k)]^T$.

Traffic light phases are switched consequently

$$S_i = (0, 1, \dots, u_i^+), i = \overline{1, M}, \quad (5)$$

where u_i^+ is an index of the last phases in the sequence, M is a number of intersections in the network.

According to the classical statement of the optimal control problem, [19], for the control object (1), the initial conditions are given

$$\mathbf{x}(0) = \mathbf{x}^0 = [x_1^0 \dots x_L^0]^T, \quad (6)$$

$$\mathbf{u}(0) = \mathbf{u}^0, u_i \in U_i, i = \overline{1, M}. \quad (7)$$

Terminal state is not given.

Note, that the control cannot change the sequence of phases (5), it changes only their duration. The control is presented in the form of a set of working phases of certain duration of traffic lights at each control step, which is called a coordination plan,

$$\tilde{\mathbf{u}}(\cdot) = (\tilde{\mathbf{u}}(0), \dots, \tilde{\mathbf{u}}(K)), \quad (8)$$

where $\tilde{\mathbf{u}}(k) = [\tilde{u}_1(k) \dots \tilde{u}_M(k)]^T$, $\tilde{u}_i(k) \in \{0, 1\}$, $i = \overline{1, M}$. Zero means that the phase remains unchanged, one stands for the change to the next phase. When the current phase number reaches its maximum value u_j^+ , it switches to the initial value. The duration of traffic light phases is limited.

Control should minimize the quality criterion

$$J = \sum_{k=1}^K f_0(\mathbf{x}(k), \mathbf{A}(\mathbf{u}(k))) \rightarrow \min. \quad (9)$$

The choice of optimisation criterion depends on the control strategy. Optimisation criteria can be used, for example, to maximise the number of vehicles on all exit road sections at the last control step or to minimise the overflow on internal road sections in total over all control steps.

Depending on the traffic light controllers available some phase switching modes may be implemented.

B. Phase switching modes

Optimal control problem 1 (without a fixed length of the cycle) - at all controlled intersections the duration of the traffic lights phases are set taking into account the constraints, so that when these working phases are cyclically repeated up to a certain time, the quality criterion (9) is met.

Optimal control problem 2 (with a fixed length of the cycle) - at all controlled intersections the duration of the traffic lights phases are set taking into account the constraints and the given total duration of the switching cycle of all phases, so that when these working phases are cyclically repeated up to a certain time, the quality criterion (9) is met.

Optimal control problem 3 (with a multicycle of fixed length). The OCP consists in determining the number of simple cycles within a complex cycle (multicycle) and in determining at all controlled intersections duration of the

traffic lights phases, taking into account the constraints. The same working phase at the same traffic light can have different duration in all simple cycles within a multicycle, so that when cyclically repeating this complex cycle until a certain time, the quality criterion (9) is met.

C. Method

Among the analytical methods for solving the optimal control problem, one of the most important is the Pontryagin maximum principle [19]. The model (1) can be represented as a system of ODEs, but even in this case the dimensionality and non-linearity of the system make it difficult to apply the maximum principle. It is proposed to use numerical methods for its solution.

The main computational problem in solving the optimal control problem in the proposed statement is the large number of unknown parameters. In general, the search space has the dimension $P = 2^{KM}$, where K is a number of control steps, M is a number of intersections in the network. For example, for a coordination plan with a duration of 1 hour and a control step of 1 second for two intersections, $P = 2^{3600 \times 2}$.

To reduce the search space a variational genetic algorithm (VarGA) is used, see [26]. Like many other evolutionary algorithms VarGA consists of the following main steps: generating a set of possible solutions, evaluating each solution according to selected criteria, selecting solutions to perform genetic operations, performing genetic operations to obtain new possible solutions, evaluating new solutions and deciding whether or not to include the resulting solutions in the set of possible solutions. Further, the algorithm is repeated for the current set of possible solutions a certain number of times, which is called generations. The condition for terminating the algorithm may be reaching a given number of generations and determining the best possible solution in the last generation or finding a solution with a given accuracy.

The feature of VarGA is that it uses the principle of small variations of the basic solution, see [27]. According to this principle, a basic solution, which represents the current coordination plan for traffic light objects at intersections, and a set of its permissible variations are specified. The search for an optimal coordination plan is performed on the set of small variations of the basic solution. After some generations, called epoch, the basic solution is changed to the best currently found one, and then variations are applied to the new basic solution, etc.

III. Multicriterial Optimal Control Problem

In the presence of several quality criteria, e.g. maximising the traffic flow on exit road sections and minimising the average manoeuvre waiting time, the solution of the optimal control problem consists in finding a set of Pareto optimal solutions. It requires definition of dominance of one solution over another to rank solutions according to their fitness.

In addition, when using VarGA the problem of changing the basic solution in the search process, it is necessary to select a solution from the Pareto set after several generations and make it the basis for further search. In this case, it is necessary to define additional conditions for the selection of the basic solution, for example, by the minimum norm among the normalised criteria on the Pareto set of zero rank.

A. Problem statement

In the road network, each traffic light system has its own set of operating phases, the order of which is cyclically repeated. Traffic flows are controlled by varying the duration of the phases. The traffic flow control is described as (1).

At all controlled intersections in the network, the traffic flow is changed depending on the traffic lights phases (8).

The state vector is limited (2). The sequence of phases is given (5). With a given initial state (6), initial values of the traffic light phases $\mathbf{v}(0)$, known entry flow (3), it is necessary to find the optimal control program to minimize several quality criteria (9). Such criteria may be maximization of the number of vehicles on all exit sections, minimization of the number of vehicles on entry sections, minimization of overflows on internal sections, etc. Criteria often do not agree with each other.

Assume, that m quality criteria are given on the set of possible solutions

$$\mathbf{f}(\mathbf{x}(\cdot), \mathbf{u}(\cdot)) = [f_1(\mathbf{x}(\cdot), \mathbf{u}(\cdot)), \dots, f_m(\mathbf{x}(\cdot), \mathbf{u}(\cdot))]^T, \quad (10)$$

$$f_i(\mathbf{x}(\cdot), \mathbf{u}(\cdot)) \in \mathbb{R}^1, \quad i = \overline{1, m}.$$

The solution is the set of Pareto optimal programs on the space of criteria

$$P = \{(\tilde{\mathbf{x}}^1(\cdot), \tilde{\mathbf{u}}^1(\cdot)), \dots, (\tilde{\mathbf{x}}^s(\cdot), \tilde{\mathbf{u}}^s(\cdot))\}, \quad (11)$$

where

$$\forall (\tilde{\mathbf{x}}^i(\cdot), \tilde{\mathbf{u}}^i(\cdot)) \in P, i \in \{1, \dots, s\},$$

$$\mathbf{f}(\tilde{\mathbf{x}}^i(\cdot), \tilde{\mathbf{u}}^i(\cdot)) \prec \mathbf{f}(\mathbf{x}(\cdot), \mathbf{u}(\cdot)),$$

$$f_k(\tilde{\mathbf{x}}^i(\cdot), \tilde{\mathbf{u}}^i(\cdot)) \leq f_k(\mathbf{x}(\cdot), \mathbf{u}(\cdot)), \quad k = \overline{1, m}, \quad \text{and}$$

$$\exists j \in \{1, \dots, m\}, \quad f_j(\tilde{\mathbf{x}}^i(\cdot), \tilde{\mathbf{u}}^i(\cdot)) < f_j(\mathbf{x}(\cdot), \mathbf{u}(\cdot)).$$

To solve the multiobjective optimization problem, a variational genetic algorithm with nondominated sorting is proposed, [28], [27].

B. Method

A distinctive feature of variational nondominated sorting genetic algorithm (VarNSGA-II) for solving the multiobjective optimal control problem of traffic flows is the representation of a set of possible solutions.

We select the Pareto set from the set of possible solutions. The Pareto set contains solutions for which the following condition is met: for each solution that is in the Pareto set, there is no solution that dominates it, and for each solution that is not in the Pareto set, there is always an element in the Pareto set that dominates it. Solutions are ranked.

Thus, each possible solution is evaluated using the Pareto rank and an additional indicator that evaluates the proximity of the solution to neighboring solutions that have the same Pareto rank.

The Pareto front with rank equal to one, created after the last generation, is considered a solution to the problem.

IV. Control Synthesis Problem

The traffic flow control synthesis problem is considered with respect to the duration of traffic light phases depending on the state of traffic flow. The entry flow vector δ in (1) in general is the uncertainty of the mathematical model. In the optimal control problem, the δ is a constant most probable value. If a random vector variable has a normal Gaussian distribution, then the most probable value is the average value observed over the interval for which the optimal control problem was solved. In real conditions, the entry flow may change sharply and then the program control $\mathbf{u}(k)$ will no longer be optimal and will require adjustment. Such adjustment may be calculated in advance, at the stage of solving the optimal control problem. In this case, it is necessary to solve the control synthesis problem and find the control as a function of state space coordinates. If the control depends on the state, then firstly it is necessary to find a control function and then apply it in case of a sharp change in the state of the network.

A. Problem statement

Let us consider the control synthesis of traffic flow in the urban network. Let us apply a universal recurrent model for the control of traffic flows (1). Traffic light phases are switched consequently (5). Every road section is characterized by the maximal number of vehicles that can be there at one time (2), (4).

The domain of initial states is given

$$X \in \mathbb{R}^L. \quad (12)$$

The control is a function of state

$$\mathbf{u} = \mathbf{h}(\mathbf{x}) \in U, \quad (13)$$

that provides the minimum value of a given quality criterion

$$J = \sum_{k=1}^K f_0(\mathbf{x}(k), \mathbf{A}(\mathbf{h}(\mathbf{x}(k)))) \rightarrow \min. \quad (14)$$

B. Method

The problem of synthesising a control system as a function of state for an initial state was formulated by [29]. To solve the synthesis problem, a dynamic programming method has been developed which produces a set of control vector values depending on the values of the state vector. The dynamic programming method is most effective for discrete values of the state vector. With many initial states and a large number of state vector values, the dynamic programming method results in a large amount of data, which has been called the ‘‘curse of dimensionality’’. The resulting solution is sensitive to the initial state. For other initial states, the resulting control will not be optimal and the synthesis problem will have to be solved again.

To solve the synthesis problem, the Bellman equation has been proposed, from which it is possible to construct a state control function. In known examples, the Bellman equation has a standard quadratic form and in general there is no approach to finding it.

In addition, synthesis seeks a function of many variables rather than a function of one time variable as in the optimal control problem, which is computationally more complex.

The problem of synthesising a control system for the entire state space as a general synthesis problem was formulated in the work of [30]. The maximum principle of L.S. Pontryagin was used as a solution method. It does not allow finding multidimensional control functions, but can be used for simple models of low-dimensional control objects. The result of solving the general synthesis problem is a control function which, when substituted into the right-hand sides of the object model, one can obtain a particular solution which, from any initial state, provides the optimal value of the quality criterion for a given time.

In this paper it is proposed to solve the problem of synthesising traffic flow control numerically [31] in the form of a multidimensional control function. The dimension of the control function is determined by the dimension of the control vector. The number of arguments is determined by the dimension of the state vector. As a solution method, it is proposed to use machine learning control using the method of symbolic regression, which allows to search for mathematical expressions in coded form using a special genetic algorithm, specially designed in such a way that, after the basic operations of evolution, obtain the correct function codes of mathematical expressions.

C. Alternative approach

Alternative approach is the solution of the synthesis problem via adaptive approach. First, the optimal coordinated signal timings for a number of possible situations, different initial conditions and different distribution parameters are calculated. The task of control synthesis will be to determine that the current coordinated signal timing is no longer optimal and select the optimal coordinated signal timing from the database. The scheme is presented in Figure 1.

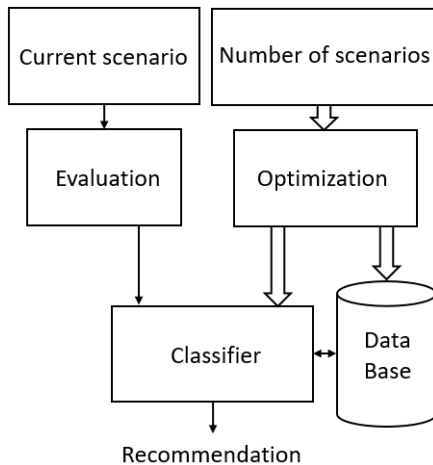


Fig. 1. Scheme of adaptive approach

The control synthesis problem is a complex computational problem that cannot be solved in real time. At the same time, solving the general synthesis problem makes it possible to obtain optimal control values for any admissible state of the object. However, this advantage of this task is negated if the model parameters change, for example, a change in the maneuver capacity parameter due to weather conditions.

Thus, the solution to the synthesis problem depends on the model parameters. The approach based on the selection of optimal coordination plans seems to be more appropriate

and feasible. With this approach, optimal coordination plans can be calculated in advance for various values of model parameters and selected in real time using a decision function that is also calculated in advance.

V. Computational Experiment

The proposed approaches were implemented in CTraf software, [32]. Here the solution of the multicriteria optimal control problem for X-type intersection is provided. Intersection with road sections and maneuvers and its graph are presented in Figure. 2.

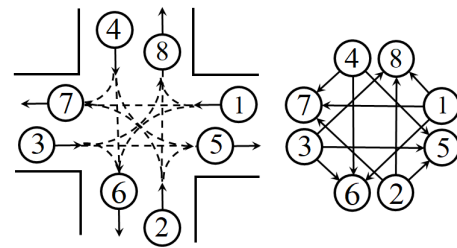


Fig. 2. X-type intersection and its graph

Two quality criteria were used:

J_1 – maximization of overall throughput of intersection, i.e. maximization of number of vehicles at exit sections (I_1) at the final control step K

$$J_1 = - \sum_{i \in I_1} x_i(K) \rightarrow \min;$$

J_2 – equal throughput of directions to avoid queues on certain entry sections (I_0) during the performance

$$J_2 = \frac{1}{C} \sum_{i \neq j, k \bmod T_c = 0} (x_i^{\max}(k) - x_j^{\min}(k)) \rightarrow \min, i, j \in I_0,$$

where C is a number of control cycles, T_c is the duration of one control cycle.

The solution of the multicriterial optimal control problem for the control program 1 (CP1 (6.00-7.30, Mo-Su)) after 32 generations of VarNSGA-II is a Pareto front presented on Fig. 3.

For solution №27 the basic control program and resulting optimal control program are given in Table. I. The overall throughput of intersection J_1 was improved by 3%, equality of throughput of directions J_2 by 7.7% respectively. The results for all control programs are given in Table. II. Parameters of VarNSGA-II are given in Table. III. As it can be seen from the results of experiment all five control programs were improved by both criteria.

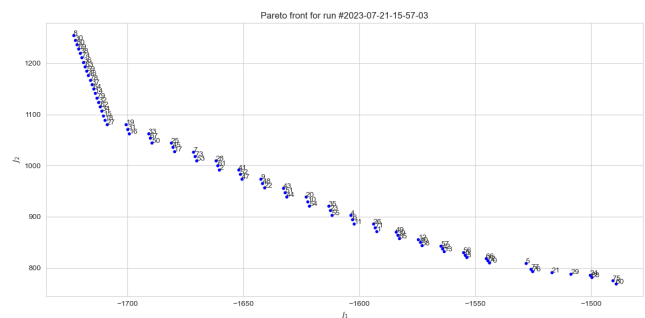


Fig. 3. Pareto front for CP1

Таблица I
Control Program 1 (6.00-7.30, Mo-Su)

Phase	T_{bas}, s	T_{opt}, s
1	43	37
2	19	19
3	19	25
4	29	29
T_c	110	110

Таблица II
Performance of Optimal CP 1-5

Control Program	$J_1, \%$	$J_2, \%$
1 (06.00-07.30, Mo-Su)	+3	+7.7
2 (07.30-10.00, Mo-Su)	+1	+12
3 (10.00-16.30, Mo-Su)	+3	+8.25
3 (20.00-21.30, Mo-Su)	+3	+8.25
4 (16.30-20.00, Mo-Su)	+3	+2
5 (21.30-06.00, Mo-Su)	+2	+13.65

Таблица III
Parameters of optimization algorithm

Parameter	Value
Size of population, H	128
Number of generations, G	32
Number of crossovers, R	128
Depth of variation, d	16
Type of variation	point
Probability of mutation, p_μ	0,75
Type of mutation	even
Type of optimization	within one control cycle

Acknowledgement

The research was carried out using the infrastructure of the Shared Research Facilities “High Performance Computing and Big Data” (СКР “Informatics”) of FRC CSC RAS (Moscow).

Список литературы

- [1] Eom M. Kim B.I. The traffic signal control problem for intersections: a review // *Eur. Transp. Res. Rev.* — 2020. — Vol. 12, no. 50.
- [2] Wei H. Zheng G. Gayah V. Li Z. A survey on traffic signal control methods // *arXiv:1904.08117 [cs.LG]*. — 2020.
- [3] Global practices on road traffic signal control: Fixed-time control at isolated intersections / Keshuang Tang, Manfred Boltze, Hideki Nakamura, Zong Tian. — Elsevier, 2019.
- [4] Hunt P.B. Robertson D.I. Bretherton R.D. Royale M.C. The SCOOT on-line traffic signal optimization technique // *Traffic Engineering Control*. — 1982. — Vol. 23. — P. 190–192.
- [5] Robertson D.I., Bretherton R.D. Optimizing networks of traffic signals in real-time SCOOT method // *IEEE Trans. Veh. Technol.* — 1991. — Vol. 40. — P. 11–15.
- [6] Sims A. The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits // *IEEE Trans. Veh. Technol.* — 1980. — Vol. 29. — P. 130–137.
- [7] Cools S.B., Gershenson C., D’Hooghe B. Self-organizing traffic lights: A realistic simulation // *Advances in Applied Self-Organizing Systems*. — 2013. — P. 45–55.
- [8] Varaiya P. Max pressure control of a network of signalized intersections // *Transp. Res. Part C Emerg. Technol.* — 2013. — Vol. 36. — P. 177–195.
- [9] Reinforcement learning in urban network traffic signal control: A systematic literature review / M. Noaen, A. Naik, L. Goodman et al. // *Expert Syst. Appl.* — 2022. — Vol. 199. — P. 116830.
- [10] Intellilight: A reinforcement learning approach for intelligent traffic light control / H. Wei, G. Zheng, H. Yao, Z. Li // *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, London, UK, 19–23 August 2018.* — 2018. — P. 2496–2505.
- [11] Haydari A., Yilmaz Y. Deep reinforcement learning for intelligent transportation systems: A survey // *IEEE Trans. Intell. Transp. Syst.* — 2022. — Vol. 23. — P. 11–32.
- [12] Shabestary S.M.A., Abdulhai B. Deep learning vs. discrete reinforcement learning for adaptive traffic signal control // *Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018.* — 2018. — P. 286–293.
- [13] Zeng J., Hu J., Zhang Y. Adaptive traffic signal control with deep recurrent q-learning // *Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 26–30 June 2018.* — 2018. — P. 1215–1220.
- [14] Chen P., Zhu Z., Lu G. An adaptive control method for arterial signal coordination based on deep reinforcement learning // *Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 27–30 October 2019.* — 2019. — P. 3553–3558.
- [15] Expression might be enough: Representing pressure and demand for reinforcement learning based traffic signal control / L. Zhang, Q. Wu, S. Jun et al. // *Proceedings of the 39th International Conference on Machine Learning, Baltimore, MD, USA, 17–23 July 2022.* — 2022. — P. 26645–26654.
- [16] Van Hasselt H., Guez A., Silver D. Deep reinforcement learning with double q-learning // *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI’15), Austin, TX, USA, 25–30 January 2015.* — 2015. — P. 2094–2100.
- [17] Dueling network architectures for deep reinforcement learning / Z. Wang, T. Schaul, M. Hessel et al. // *Proceedings of the 33rd International Conference on Machine Learning (ICML’16), New York, NY, USA, 19–24 June 2016.* — 2016. — P. 1995–2003.
- [18] Prioritized experience replay / T. Schaul, J. Quan, I. Antonoglou, D. Silver // *Proceedings of the 4th International Conference on Learning Representations (ICLR’16), San Juan, PR, USA, 2–4 May 2016.* — 2016.
- [19] *The Mathematical Theory of Optimal Processes* / L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, E.F. Mishchenko. — VIII + 360 S. : New York/London, 1962.
- [20] Genealogy of traffic flow models / F. Van Wageningen-Kessels, H. van Lint, K. Vuik, S. Hoogendorn // *EURO Journal on Transportation and Logistics*. — 2015. — Vol. 4. — P. 445–473.
- [21] Introduction to mathematical modeling of traffic flows: textbook / A.V. Gasnikov, S.L. Klenov, E.A. Nurminskij et al. — 362 p. [in Russian] : MIPT, Moscow, 2010.
- [22] Diveev A.I. Controlled networks and their applications // *Computational Mathematics and Mathematical Physics*. — 2008. — Vol. 48, no. 8. — P. 1428–1442.
- [23] Sofronova E.A. Hybrid recurrent traffic flow model (URTFM-RNN) // *Intelligent Systems and Applications, Proceedings of the 2021 Intelligent Systems Conference (IntelliSys)*. — 2021. — Vol. 2.
- [24] Sofronova E.A., Diveev A.I. Traffic flows optimal control problem with full information // *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), Rhodes, Greece.* — 2020. — P. 1–6.
- [25] Sofronova E., Diveev A. Controlled networks to solve traffic flows problem // *2022 International Conference on Modern Network Technologies (MoNeTec)*. — 2022.
- [26] Sofronova E.A., Belyakov A.A., Khamadiyarov D.B. Optimal control for traffic flows in the urban road networks and its solution by variational genetic algorithm // *Procedia Computer Science*. — 2019. — 01. — Vol. 150. — P. 302–308.
- [27] Sofronova E.A., Diveev A.I. Universal approach to solution of optimization problems by symbolic regression // *Appl. Sci.* — 2021. — Vol. 11. — P. 5081.
- [28] A fast and elitist multi-objective genetic algorithm: NSGA-II / K. Deb, A. Pratap, S. Agarwal, T. Meyarivan // *IEEE Transactions on Evolutionary Computation*. — 2002. — Vol. 6, no. 2. — P. 182–197.
- [29] Bellman R. *Dynamic Programming*. — 340 p. : Princeton University Press, Princeton, New Jersey, Sixth Printing, 1972.
- [30] Boltyanskiy V.G. *Mathematical methods of optimal control*, 2nd Edition. — 408 p. [in Russian] : Nauka, Moscow, 1969.
- [31] Diveev A.I. *Numerical methods for control synthesis problem: monograph*. — 192 p. [in Russian] : RUDN University, Moscow, 2019.
- [32] Sofronova E.A., A.I. Diveev. Package for simulation and search for the optimal control program for groups of traffic lights using variational genetic algorithm. certificate of state registration of the computer program no. 2020619911 dated August 25, 2020. — 2020.

Подходы к автоматизированному оптимальному управлению транспортными потоками

Софронова Е.А.

Abstract—В статье рассматриваются задачи оптимального управления и синтеза управления потоками транспорта в сети городских дорог. Данные задачи предложено решать в рамках автоматизированной системы управления, используемой при создании интеллектуальной транспортной системы (ИТС) больших городов. Объектом управления является транспортный поток. Состояние объекта представляется в виде числовой оценки величины потока на каждом участке дороги в каждый момент времени. Информация о количественных характеристиках потока поступает с детекторов дорожной инфраструктуры. Сеть дорог описывается ориентированным графом, вершинам которого соответствуют участки дорог, а дугам – маневры на перекрестках. Граф имеет переменную структуру, зависящую от управления. Управление определяется длительностями фаз светофоров на регулируемых перекрестках. Для описания управления транспортными потоками используется универсальная рекуррентная модель управления транспортными потоками (УРМ УТП), построенная на основе теории управляемых сетей. Приведены постановки задачи оптимального управления для различных режимов переключения фаз: внутри фиксированного цикла, без фиксированного цикла, внутри мультицикла. Приведена постановка задачи многокритериального оптимального управления. Для решения задач оптимального управления предложено использовать современные эволюционные алгоритмы. Далее приведена постановка задачи синтеза управления транспортными потоками. В задаче синтеза управление транспортными потоками осуществляется за счет выбора длительностей рабочих фаз светофоров в зависимости от состояния объекта. Для решения задачи синтеза управления предложено использовать численные методы символьной регрессии. Приведено численное решение задачи многокритериального оптимального управления потоками транспорта на перекрестке по данным с детекторов.

Keywords—оптимальное управление, синтез управления, управление транспортными потоками, модель транспортных потоков

Работа выполнялась с использованием инфраструктуры Центра коллективного пользования «Высокопроизводительные вычисления и большие данные» (ЦКП «Информатика») ФИЦ ИУ РАН (г. Москва).

Список литературы

- [1] Eom M. Kim B.I. The traffic signal control problem for intersections: a review // Eur. Transp. Res. Rev. — 2020. — Vol. 12, no. 50.
- [2] Wei H. Zheng G. Gayah V. Li Z. A survey on traffic signal control methods // arXiv:1904.08117 [cs.LG]. — 2020.
- [3] Global practices on road traffic signal control: Fixed-time control at isolated intersections / Keshuang Tang, Manfred Boltze, Hideki Nakamura, Zong Tian. — Elsevier, 2019.
- [4] Hunt P.B. Robertson D.I. Bretherton R.D. Royale M.C. The SCOOT on-line traffic signal optimization technique // Traffic Engineering Control. — 1982. — Vol. 23. — P. 190–192.
- [5] Robertson D.I., Bretherton R.D. Optimizing networks of traffic signals in real-time SCOOT method // IEEE Trans. Veh. Technol. — 1991. — Vol. 40. — P. 11–15.
- [6] Sims A. The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits // IEEE Trans. Veh. Technol. — 1980. — Vol. 29. — P. 130–137.
- [7] Cools S.B., Gershenson C., D’Hooghe B. Self-organizing traffic lights: A realistic simulation // Advances in Applied Self-Organizing Systems. — 2013. — P. 45–55.
- [8] Varaiya P. Max pressure control of a network of signalized intersections // Transp. Res. Part C Emerg. Technol. — 2013. — Vol. 36. — P. 177–195.
- [9] Reinforcement learning in urban network traffic signal control: A systematic literature review / M. Noeen, A. Naik, L. Goodman et al. // Expert Syst. Appl. — 2022. — Vol. 199. — P. 116830.
- [10] Intellilight: A reinforcement learning approach for intelligent traffic light control / H. Wei, G. Zheng, H. Yao, Z. Li // Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, London, UK, 19–23 August 2018. — 2018. — P. 2496–2505.
- [11] Haydari A., Yilmaz Y. Deep reinforcement learning for intelligent transportation systems: A survey // IEEE Trans. Intell. Transp. Syst. — 2022. — Vol. 23. — P. 11–32.
- [12] Shabestary S.M.A., Abdulhai B. Deep learning vs. discrete reinforcement learning for adaptive traffic signal control // Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018. — 2018. — P. 286–293.
- [13] Zeng J., Hu J., Zhang Y. Adaptive traffic signal control with deep recurrent q-learning // Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 26–30 June 2018. — 2018. — P. 1215–1220.
- [14] Chen P., Zhu Z., Lu G. An adaptive control method for arterial signal coordination based on deep reinforcement learning // Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 27–30 October 2019. — 2019. — P. 3553–3558.
- [15] Zhang L., Wu Q., Jun S., Lü L., Du B., Wu J. Expression might be enough: Representing pressure and demand for reinforcement learning based traffic signal control // Proceedings of the 39th International Conference on Machine Learning, Baltimore, MD, USA, 17–23 July 2022. — 2022. — P. 26645–26654.
- [16] Van Hasselt H., Guez A., Silver D. Deep reinforcement learning with double Q-learning // Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI’15), Austin, TX, USA, 25–30 January 2015. — 2015. — P. 2094–2100.
- [17] Dueling network architectures for deep reinforcement learning / Z. Wang, T. Schaul, M. Hessel et al. // Proceedings of the 33rd International Conference on Machine Learning (ICML’16), New York, NY, USA, 19–24 June 2016. — 2016. — P. 1995–2003.
- [18] Prioritized experience replay / T. Schaul, J. Quan, I. Antonoglou, D. Silver // Proceedings of the 4th International Conference on Learning Representations (ICLR’16), San Juan, PR, USA, 2–4 May 2016. — 2016.
- [19] Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. М.: Наука, 1983.
- [20] Genealogy of traffic flow models / F. Van Wageningen-Kessels, H. van Lint, K. Vuijck, S. Hoogendorn // EURO Journal on Transportation and Logistics. — 2015. — Vol. 4. — P. 445–473.
- [21] Гасников А.В., Кленов С.Л., Нурминский Е.А., Холодов Я.А., Шамрай Н.Б. Введение в математическое моделирование транспортных потоков. М.: МФТИ, 2010.
- [22] Дивеев А.И. Управляемые сети и их приложения // Журнал вычислительной математики и математической физики. 2008. 48(8). С. 1510–1525.
- [23] Sofronova E.A. Hybrid recurrent traffic flow model (URTFM-RNN) // Intelligent Systems and Applications, Proceedings of the 2021 Intelligent Systems Conference (IntelliSys). — 2021. — Vol. 2.

- [24] Sofronova E.A., Diveev A.I. Traffic flows optimal control problem with full information // 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), Rhodes, Greece. — 2020. — P. 1–6.
- [25] Sofronova E., Diveev A. Controlled networks to solve traffic flows problem // 2022 International Conference on Modern Network Technologies (MoNeTec). — 2022.
- [26] Sofronova E.A., Belyakov A.A., Khamadiyarov D.B. Optimal control for traffic flows in the urban road networks and its solution by variational genetic algorithm // Procedia Computer Science. — 2019. — 01. — Vol. 150. — P. 302–308.
- [27] Sofronova E.A., Diveev A.I. Universal approach to solution of optimization problems by symbolic regression // Appl. Sci. — 2021. — Vol. 11. — P. 5081.
- [28] A fast and elitist multi-objective genetic algorithm: NSGA-II / K. Deb, A. Pratap, S. Agarwal, T. Meyarivan // IEEE Transactions on Evolutionary Computation. — 2002. — Vol. 6, no. 2. — P. 182–197.
- [29] Bellman R. Dynamic Programming. — 340 p. : Princeton University Press, Princeton, New Jersey, Sixth Printing, 1972.
- [30] Болтянский В.Г. Математические методы оптимального управления, Рипол Классик, 2013, 414 с.
- [31] Дивеев А.И. Численные методы решения задачи синтеза управления. — М.: РУДН, 2019. 189 с.
- [32] Софронова Е.А., Дивеев А.И. Программный комплекс для моделирования и поиска оптимальной программы управления группами светофоров методом вариационного генетического алгоритма. Свидетельство о государственной регистрации программы для ЭВМ №20200619911 от 25 августа 2020 г.

Автор

Софронова Елена Анатольевна, к.т.н., доцент, старший научный сотрудник ФИЦ ИУ РАН. Область научных интересов: исследование и разработка математических моделей и алгоритмов для оптимального управления потоками транспорта в сети городских дорог; методы символьной регрессии и эволюционные алгоритмы для идентификации, оптимального управления и синтеза управления.