A MULTI-OBJECTIVE APPROACH WITH A RESTART META-HEURISTIC FOR THE LINEAR DYNAMICAL SYSTEMS INVERSE MATHEMATICAL PROBLEM

[1]Ivan Ryzhikov, [2]Christina Brester, [3]Eugene Semenkin

Institute of Computer Science and Telecommunications,
Reshetnev Siberian State University of Science and Technologies,
e-mails: [1]ivan.s.ryzhikov@gmail.com, [2]christina.brester@gmail.com,
[3]eugenesemenkin@yandex.ru
Russia

Abstract: This study is focused on the automatic modelling of linear dynamical systems based on system output observations. Identification of a differential equation degree and its parameters requires the initial point because the output of the dynamical system for the given input depends on the initial point value. Generally, it is impossible to estimate the system state and its derivatives using the observation values, so the inverse mathematical modelling problem for dynamical systems can be reduced to a simultaneous approximation of the differential equation and the initial point. Solving the reduced two-criterion extremum problem requires a searching algorithm with specific problem-oriented modifications.

Key words: dynamical system modelling, linear time-invariant systems, initial point estimation, multi-objective optimization, restart meta-heuristic

1. INTRODUCTION

The mathematical modelling of dynamical systems, as with mathematical modelling in general, is a complex and important problem, since the model is the basis for the problems of system control, analysis and prognoses. At the same time, the tools for solving these kinds of problems requires a model in a specific form, such as differential equations or difference equations, so it is preferable for the modelling approach to receive an identity of this class of models. The basic way of solving the system identification problem is its reduction to an extremum problem with some criterion, and that is why the goal of this work is to propose the reduction scheme and improve the extremum-seeking approach applied to the reduced identification problem.

This study comes from an intersection of the system identification and multi-criterion optimization fields. This intersection is well-known in some particular cases, such as static system parameter identification or statistical modelling, and inverse mathematical modelling for static systems, but here we consider dynamical systems in a more generalised way than is common in other approaches. Moreover, we need to point out that dynamical system modelling is related to multi-criterion optimization problem solving because in the case of the dynamics, we need to identify not just some set of parameters, which determines the transient process, but also an initial point, which is used for the numerical estimation of the dynamical system output. Actually, many different approaches are based on the hypothesis that this initial value is known or can be estimated, but from a practical point of view, this assumption can lead to a significant error rate arising, since the numerical schemes used in the approximation of the model's parameters are sensitive to the initial value.

Statistical methods are not a general tool to approximate the initial value, because there is a need to estimate the function with, for example, some of its derivatives, and in practice, the sample size can be small or the sample data can be flat, so the estimation of the initial value is a distinct problem. At the same time, we could decide if the initial point is good or bad on the basis of how the model fits the sample data for some particular initial point. Here we can see that initial value estimation is not an isolated problem.

As can be seen, the inverse mathematical modelling problem is related to the initial value estimation and the initial value estimation is related to inverse mathematical modelling, so each of these problems should be solved simultaneously. For solving the proposed two-objective optimization problem, a specific extremum-seeking tool is required. Firstly, previous studies have shown that the efficiency of this tool is provided by its problem-orientation - [7] and [6], since the dynamical system identification problem is a hard optimization problem. Secondly, implementing meta-heuristics, such as a restart meta-heuristic, provides an improvement in the efficiency of the multi-objective extremum-seeking algorithm [4].

In this study, we used Preference Inspired Co-Evolutionary Algorithms-goals (PICEA-g) with a restart meta-heuristic. The proposed approach was applied for a set of identification problems for the systems of various differential equation degrees, their input functions and initial values. The proposed approach performance was estimated and analysed, and the experimental results are presented with figures and tables.

This paper consists of an introduction, three chapters: Inverse Linear Time Invariant System Modelling, Multi-Criteria Optimization and Restart Meta-Heuristic, and Experimental Results, and conclusions. In the first section, the main principles of the proposed meta-heuristic are stated, in the second, the problem definition is given and in the third, the numerical results are presented.

Attempts at solving the considered problem were based on solving an additive form of the considered two-criterion extremum problem, but this approach allows only one point on the non-dominated field to be estimated, depending on the coefficients of criteria linear combination [6] and [7]. There are many related works on inverse mathematical modelling for dynamical systems, but with a different problem definition. Linear time-invariant system identification problems, where an evolution-based or a nature-inspired optimization tool is applied, are stated in [4] references. A deeper look at dynamical system identification shows that these areas have a lot of application fields and include many different problem definitions, which differ also in model class, but each problem needs its own high performance processing and optimization algorithm.

The PICEA-g with the restart meta-heuristic was successfully implemented for solving the hexadecane disintegration reaction modelling problem [5], and proved its high effectiveness. The studies on restart meta-heuristic development are stated in [1], [2] and

[3]. In these studies, the authors demonstrate the benefits of applying a restart meta-heuristic, but it is designed and implemented for single-objective optimization problems.

2. INVERSE LINEAR TIME INVARIANT SYSTEM MODELLING

In this chapter, we give a problem definition for linear dynamical system modelling. Let the dynamical system observation be represented with sample data: dynamical system output observations $Y = y_i, i = \overline{1,s}$ and time points $T = t_i, i = \overline{1,s}$, where s is the sample size. It is also known that the system inputs are determined with a vector-function $U(t) = (u_1(t) \quad \ldots \quad u_m(t))$, where m is the number of inputs. Due to the assumption of system linearity, its model can be determined as follows:

$$\sum_{i=0}^{d} \tilde{a}_i \cdot x^{(i)}(t) = \sum_{k=1}^{m} \tilde{b}_k \cdot U(t)_k , \quad (1)$$

where d is the highest degree of the equation. After dividing (1) by $\tilde{a}_d \neq 0$, we receive the equation in a more suitable form:

$$x^{(d)}(t) + \sum_{i=0}^{d-1} a_i \cdot x^{(i)}(t) = \sum_{k=1}^{m} b_k \cdot U(t)_k , \quad (2)$$

The output of the real system is the solution of the Cauchy problem for equation (2) with the initial value

$$x^{(i)}(t_0) = p_i, i = \overline{1, d-1}, \quad (3)$$

where $p \in R^{d-1}$ is the vector of initial values.

Now we assume that the observations are random values

$$y_i = x(t_i) + \xi_i, \xi \sim N(0, \sigma_\xi), \quad (4)$$

where $N(0, \sigma_\xi)$ is a normal distribution with a zero expected value and standard deviation $\sigma_\xi < \infty$.

According to (2) and (3), we need to identify the differential equation maximum degree $\hat{d} \in N, \hat{d} \leq D$, which is limited by a positive value D, vectors $\hat{a} \in R^{\hat{d}}$ and $\hat{b} \in R^m$, and the initial value $\hat{p} \in R^{\hat{d}-1}$. All of these parameters can be identified with a minimization of the following criteria

$$C_1(\hat{d}, \hat{a}, \hat{b}, \hat{p}) = \sum_{i=1}^{s} \left\| y_i - \hat{x}(t_i) \right\| \to \min_{\hat{d} < D,\, \hat{a}, \hat{b} \in R^{\hat{d}},\, \hat{p} \in R^{\hat{d}-1}},$$
$$C_2(\hat{p}) = \left\| y_{\arg\min_i (t_i)} - \hat{p}_0 \right\| \to \min_{\hat{p} \in R^{\hat{d}-1}}, \quad (5)$$

where $\hat{x}(t)$ is the solution of the Cauchy problem

$$\hat{x}^{(d)}(t) + \sum_{i=0}^{\hat{d}-1} \hat{a}_i \cdot \hat{x}^{(i)}(t) = \sum_{k=1}^{m} \hat{b}_k \cdot U(t)_k , \quad (6)$$

$$\hat{x}^{(i)}\left(\min_i (t_i)\right) = \hat{p}_i,\ i = \overline{1, \hat{d}-1}.$$

As can be seen, the linear dynamical system can be fully determined with a set of the considered parameters, which brings the extremum to a two-objective optimization problem (5). The first objective function of this problem is the inadequacy degree estimation, which indicates how well the model output fits the observation data. The second objective function is the distance between the first observations and the estimated initial value first coordinate, which is the approximation of the system output. Any other criteria could be used here instead of the proposed one. However, this one was chosen because of its applicability in a general case when the sample size is small.

In studies [6] and [7] the degree is the variable of the extremum problem (5), but in the current study, we propose enumerative searching algorithm, which enumerates the degree. There are two reasons for this. The first is an investigation of the approach performance in approximating the whole set of parameters, and the second is based on fact that different degrees combined with different parameters could give us the same outputs, which decreases our confidence in it. The searching is organized in the following way:

$$\hat{d} = \overline{1, D},\ C_1\left(\hat{a}, \hat{b}, \hat{p}\right) \to \min_{\hat{a}, \hat{b} \in R^{\hat{d}},\ \hat{p} \in R^{\hat{d}-1}},\ C_2\left(\hat{p}\right) \to \min_{\hat{p} \in R^{\hat{d}-1}}. \quad (7)$$

This problem is complex, it is multimodal and these criteria are evaluated via numerical integration schemes solving (7) and require a multi-objective optimization algorithm of high performance. In the next chapter, we describe the algorithm and proposed meta-heuristic, which increased the effectiveness of the algorithm and made it more problem-oriented.

3. MULTI-CRITERIA OPTIMIZATION AND RESTART META-HEURISTIC

Let us start with the problem definition. Let X be the space of alternatives and there is a need to find the solution for the following two-criterion extremum problem

$$C_1(x) \to \underset{x \in X_1}{extrem},\ C_2(x) \to \underset{x \in X_2}{extrem}, \quad (8)$$

where $C_i(\cdot): X \to R,\ i = \overline{1,2}$ are the objective functions and $X_1, X_2 \subseteq X$ are subspaces of the whole search space X.

The functions (8) are calculable, but, generally, perform a Black-Box Optimization Problem (BBOP), which is multi-modal, complex and there is no additional information we have about this problem or the alternative space. What we actually have is (5) or (7).

Evolution-based and nature-inspired optimization algorithms tend to be the most powerful tools to solve problems of this kind, but as with any other 0-th order stochastic algorithms, there is a risk of it becoming stagnated in some basin of a local optimum. Of course, this risk is not as high as for deterministic algorithms, but it is still crucial in particular cases. The reducing of this risk leads to a contradiction between the search in breadth and the search in depth and the restart operator is an easy way out.

In this study, we consider a particular restart meta-heuristic which is used to improve the initial optimization algorithm performance. This meta-heuristic works on a simple idea: the seeking algorithm restarts if some condition is met, and can be applied to any iterative and stochastic extremum-seeking algorithm. It has various implementations for one-criterion extremum problems, but for multi-criterion problem solving, we suggested a new one in [4].

Let the set of non-dominated solutions, which are the Pareto set estimation, be denoted as S_p and the Pareto front estimation be denoted as F_p. The front and the set can change after each iteration of algorithm, and the next front estimation dominates the previous one. In a similar way to the restart operator for a single-objective optimization problem, we need the characteristics of the searching process, and that is why we designed a specific metric between the front estimations

$$\rho(u,v) = R^c \times R^c \to R,$$
$$\rho(u,v) = \frac{1}{card(u)} \cdot \sum_{i=1}^{card(u)} \min_{j \le card(v)} \left(\left\| u_i - v_j \right\| \right), \tag{9}$$

where $card(\cdot)$ is the function that returns the cardinality of an argument, and u, v are the next and the previous front approximations, respectively.

To estimate if there is stagnation, which can be seen by the slow changing of the metric (8), we need to hold a set containing these metrics for the several previous algorithm iterations to the *i*-th one

$$M_i = \left\{ \rho\left(\left(F_p\right)_i, \left(F_p\right)_{i-1} \right) : i - l_o < j \le i \right\}, \tag{10}$$

where $l_o > 0$, the controlling parameter, is the number of metric observations. Using this set, we can predict if stagnation will take place by checking the following condition

$$\max\left(M_i\right) - \min\left(M_i\right) < \delta, \tag{11}$$

where δ is another controlling parameter.

After each restart, the algorithm's non-dominated set is stored in its memory and used for performing the initial population on the next step. The next two parameters α and p_s control the probability of generating each individual in the initial population randomly against making it a modified copy of a randomly chosen solution from the memory. Each gene of the solution is mutated with probability p_s.

The so-called fitness function in this work is the mapping, based on criteria (7) and is calculated for each criterion as follows

$$f(x) = \frac{1}{1 + C(x)}, \tag{12}$$

where $C(x)$ is the criterion in terms of (8) and x is a point on the subspace X_1 or X_2, depending on which criterion is being evaluated.

The restart meta-heuristic is implemented into PICEA-g as proposed by Wang in 2013 [8]. This algorithm relates to a class of preference-inspired co-evolutionary algorithms (PICEAs) which are based on the concept of co-evolving the population with decision-maker preferences. It was used because there are a variety of successfully solved problems with this algorithm and its speed.

To estimate the algorithm efficiency, it was applied to a set of identification problems for systems of different orders and with different numbers of control inputs.

4. EXPERIEMENTS AND RESULTS

To provide the investigation of the algorithm with efficiency and the benefits from solving the considered identification problem as a multi-objective one, we performed two sets: a set of dynamical systems with its parameters and initial values and a set of control inputs with control functions and control parameters. The first one is given in Table 1.

Table 1. Linear dynamical system.

№	Differential equation	Initial value, $x(0)$
1	$x^{(4)}(t)+x'''(t)+4\cdot x''(t)+2\cdot x'(t)+x(t)=u(t)$	$(2\quad 0\quad 0\quad 0)^T$
2	$x'''(t)+4\cdot x''(t)+3\cdot x'(t)+x(t)=u(t)$	$(0\quad 0\quad 0)^T$
3	$x''(t)+2\cdot x'(t)+x(t)=u(t)$	$(0\quad 0)^T$

Table 2, similar to Table 1, is given below and here the right-side functions are presented.

Table 2. Control inputs.

№	Number of inputs	Control function
1	2	$u(t)=2+\sin(t)$
2	3	$u(t)=1+\sin(t)\cdot\sin(t)+\sin(2\cdot t)\cdot\sin(t)$
3	4	$u(t)=1+0.5\cdot t+0.25\cdot\sqrt{t}+\cos(t)$

In this investigation, we compare the values of fitness functions (12). According to the given determination of the fitness function, the closer its value to 1, the better solution is.

We investigate the algorithms with different settings for the problems below, where for each particular identification problem we take 200 points randomly from the time interval $[0, 20]$, on which the systems were integrated numerically. The standard deviation of error was set to 0. For all restarts, the border value for changing the metrics was set at $\delta=10^{-5}$, according to the fitness function.

At first, we would like to show the benefits provided by applying the multi-objective approach instead of the single-criterion approach. In Figure 1, the Pareto front estimation is given, which is the union of all estimations from 20 algorithm runs with restart and settings: $l_o=100$, $\alpha=1$ and $p_s=1$. We receive a set of solutions which differ in their fitting criteria and in their parameter values.

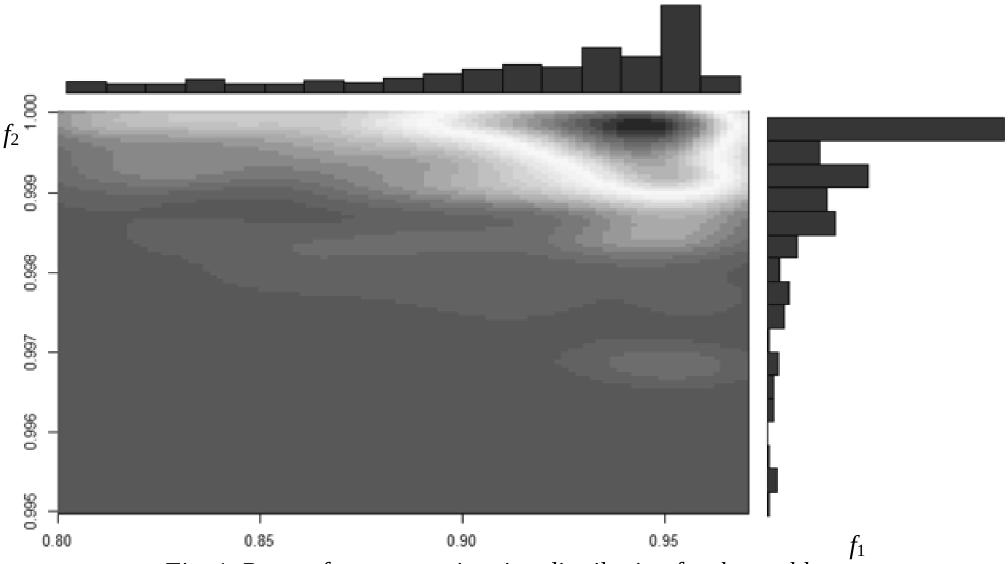

Fig. 1. Pareto front approximation distribution for the problem: dynamical system 1 (Table 1) and inputs 1 (Table 2).

Solving the identification problem as a two-objective optimization problem gives us the possibility to estimate the set of solutions, which could be closer to fitting the observation data, but distant from the initial value measurement and vice-versa. The result of a single algorithm with a restart run also proves this assumption and it is shown in Figure 2.

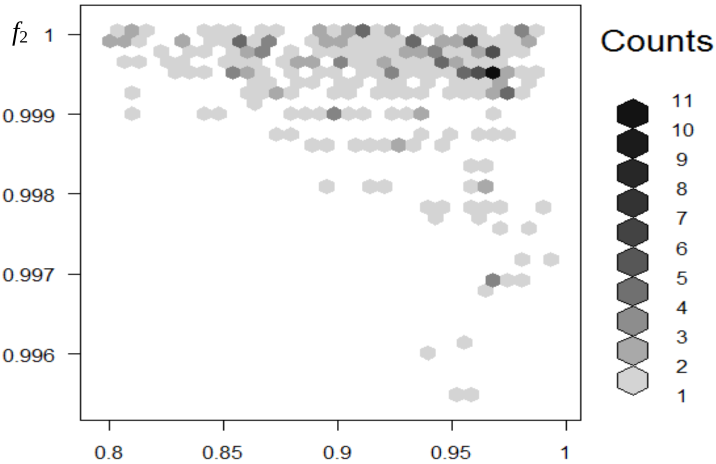

Fig. 2. Pareto front approximation. Single algorithm run.

However, Figure 1 shows that all the solutions give very good results for the second criterion and this is proven by all the experiments in which the values of the second criterion vary slightly. Because of this, and since the first criterion is actually more important, the performances of the algorithms will be approximated and compared by their first criterion values.

Next, we considered the combinations of systems given in Table 1 and control inputs given in Table 2 to estimate algorithm performance. Now we can compare the algorithm with a restart meta-heuristic and one without it by varying different parameters. A summary demonstrating the characteristics of all the algorithms is given in Figure 3, where the fitness function medium values, 10 and 90 percent quintiles and their minimum and maximum values are given. The results were achieved on the basis of 20 runs for each algorithm for every particular identification problem; the number of objective function evaluations is $2^5 \cdot 10^5$. To fit the picture, each algorithm is named in the following way: population size, l_o, and optionally α and p_s if these do not equal 1. With dashed and dotted lines, we marked the best and medium values of the best considered algorithm without a restart.

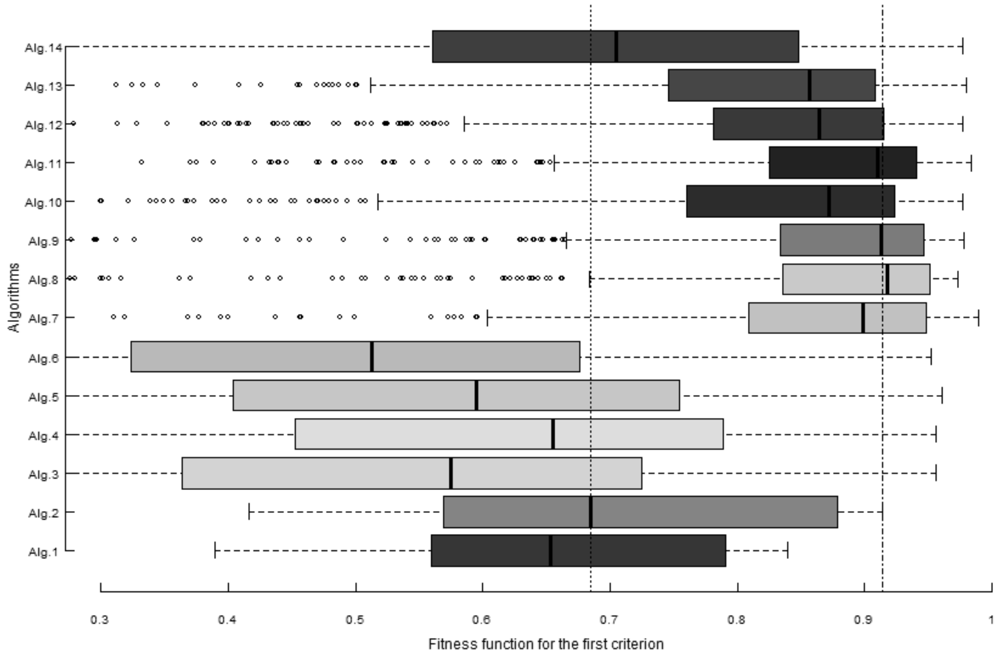

Fig. 3. Comparing the algorithm performance.

In this figure, the following notation is used: algorithms 1-2 are PECIA-g without a restart and have 4000 generations against 800 individuals and 8000 generations against 400 individuals, respectively; the other algorithms are with restart meta-heuristics. The other algorithms with the restart implementation have the settings given in Table 3, where the rows are: 1 – generations number (p.s.), 2 – population size (i.n.), 3 – restart set cardinality (r.s.c.), 4 – probability of individual random generation (p.r.g.) and 5 – super-mutation parameter (s-m.p.).

The numerical results let us make a conclusion that l_0 is the parameter, the adjusting of which significantly improves the algorithm performance and makes this algorithm more efficient than the algorithm without a restart. Increasing the number of generations and decreasing the population size gives better results. It can also be seen that controlling the initial population generation parameters could improve the performance too, which suggests that adjusting the restart parameters, making them adaptive or problem-oriented, could provide a very promising research direction.

Table 3. Algorithms.

Alg. №	3	4	5	6	7	8	9	10	11	12	13	14
p.s. $\times 10^3$	4	8	8	8	8	8	8	8	8	8	8	16
i.n. $\times 10^2$	8	4	4	4	4	4	4	4	4	4	4	2
r.s.c. $\times 10$	10	10	5	2	10	10	10	10	10	10	10	5
p.r.g	1	1	1	1	0.1	0.5	0	0.1	0.5	0.1	0.5	0.1
s-m.p.	-	-	-	-	0.2	0.2	0.2	0.5	0.5	1	1	0.7

It is important to point out that implementing the restart operator makes it possible to receive a greater set of solutions, including the approximation of the Pareto front, not only in its cardinality, but also in its width.

As can be seen, the algorithm with the restart gives more variations of the parameters, which achieve relatively the same results.

5. CONCLUSION

In this study, we proposed an approach based on the multi-criterion optimization problem reduction of the inverse mathematical modelling problem for a linear dynamical system, and solving the reduced problem with a population-based optimization algorithm with restart meta-heuristic implementation. The considered problem is complex and it requires some problem-oriented improvements of a searching algorithm.

Experimental results show that with adjusted settings, the algorithm with restart meta-heuristics works better in terms of median and best value, and it gives the more explored Pareto set and front. However, at the same time, meta-heuristics parameter adaptation, along with optimization algorithm parameter adaptation, becomes an important problem, which will be investigated in further studies.

In addition, the further work will be related not only to algorithm and meta-heuristics improvement, but also to a generalization of the proposed approach and its application to multi-output systems.

ACKNOWLEDGEMENTS

This research is performed with the financial support of the Ministry of Education and Science of the Russian Federation within state project assignment № 2.1680.2017/ПЧ.

REFERENCES

[1] Beligiannis G. N., Tsirogiannis G. A., Pintelas P. E. Restartings: a technique to improve classic genetic algorithms' performance. *J. Glob. Optim.*, 1, 2014, pp. 112-115.

[2] Fukunaga A. S. Restart Scheduling for Genetic Algorithm. *Lecture Notes of Computer Science*, Vol. 1498, 1998, pp. 357-369.

[3] Loshchilov I., Schoenauer M., Sebag M. Alternative Restart Strategies for CMA-ES. *Lecture Notes in Computer Science*, vol. 7491, 2012, pp. 296-305.

[4] Ryzhikov I., Brester C. and Semenkin E. Multi-objective Order Reduction Problem Solving with Restart Meta-heuristic Implementation. In *Proceedings of the 14th International Conference on Informatics in Control, Automation and Robotics*, Volume 1: ICINCO, 2017, p. 270-278.

[5] Ryzhikov I., Brester C. and Semenkin E. Multi-objective Dynamical System Parameters and Initial Value Identification Approach in Chemical Disintegration Reaction Modelling. In *Proceedings of the 14th International Conference on Informatics in Control, Automation and Robotics*, Volume 1: ICINCO, 2017.

[6] Ryzhikov I., Semenkin E. Restart operator meta-heuristics for a problem-oriented evolutionary strategies algorithm in inverse mathematical MISO modelling problem solving. *IOP Conference Series: Materials Science and Engineering*, 2017, vol. 173.

[7] Ryzhikov I., Semenkin E., Panfilov I. Evolutionary optimization algorithms for differential equation parameters, initial value and order identification. *Proceedings of the 13th International Conference on Informatics in Control, Automation and Robotics* (ICINCO 2016), Volume 1, 2016, pp. 168-176.

[8] Wang R. Preference-Inspired Co-evolutionary Algorithms. A thesis submitted in partial fulfillment for the degree of the Doctor of Philosophy, University of Sheffield, 2013, 231 p..

Information about the authors:

Ivan Ryzhikov is a research fellow at the Reshetnev Siberian State University of Science and Technology (Krasnoyarsk, Russia). He received his PhD in Computer Science in 2016 from the Reshetnev Siberian State Aerospace University. His areas of research include global black-box optimization, meta-heuristics for natural computing extremum-seeking algorithms, inverse mathematical modelling and dynamical system identification.

Christina Brester is an associate professor at the Department of Higher Mathematics, Reshetnev Siberian State University of Science and Technology (Krasnoyarsk, Russia). She received her PhD in Computer Science in 2016 from the Siberian Federal University and the Institute of Computational Modelling of the Siberian Branch of the Russian Academy of Sciences. Her research interests include evolutionary computation, neuro-evolutionary algorithms, machine learning and speech analysis.

Eugene Semenkin is a professor of system analysis at the Reshetnev Siberian State University of Science and Technology (Krasnoyarsk, Russia). He received his PhD in Computer Science from Leningrad State University (Leningrad, USSR) in 1989 and his DSc in Engineering and Habilitation from the Reshetnev Siberian State Aerospace University (Krasnoyarsk, Russia) in 1997. His areas of research include the modelling and optimization of complex systems, computational intelligence and data mining.

Manuscript received on 11 January 2018