Hybrid approach to the dynamic systems identification based on the self-configuring genetic programming algorithm and the differential evolution method

T S Karaseva\(^1\,^2\) and O E Semenkina\(^1\)

\(^1\)Reshetnev Siberian State University of Science and Technology, Krasnoyarsky Rabochy Av. 31, Krasnoyarsk, 660037, Russian Federation
\(^2\)Siberian Federal University, 79 Svobodny avenue, Krasnoyarsk, 660041, Russian Federation

E-mail: tatyanakarasewa@yandex.ru

Abstract. This paper considers a hybrid approach to the identification of dynamical systems based on a self-configuring genetic programming algorithm and a differential evolution method. The value of this approach is in the automatic determination of the order, structure and parameters of differential equation, i.e., a model of a dynamic system. The application of the differential evolution method can significantly increase the accuracy of the resulting model confirmed by the results of the experiments presented in this paper.

1. Introduction

The process of mathematical models of dynamic systems building is an urgent field of research, since it is impossible to study properties of a real object, make forecasts and select the correct control without an adequate model [1]. Dynamic processes are met in various fields such as biology, chemistry, bacteriology, medicine, physics, and financial mathematics. A separate class of methods is related to the process of mathematical models of dynamic systems building based on measurements that describe input and output variables [2]. Methods of parametric identification are widely applied for solving the problem in this formulation [3-4]. However, with the development of scientific and technological progress, most of the newly developed solutions often do not have a mathematical model that would be derived from a physical, chemical or other nature. It is also not always possible to determine a model’s structure of a dynamic system due to the great complexity of the solutions being developed or the absence of the theory describing such processes [2]. Therefore, it is necessary to develop identification methods capable to determine both parameters and a structure of the model based on the sample data.

This paper considers the process of the model building in the form of a differential equation based on the input and output data. A structure and parameters of the model are assumed to be unknown, and the order of the model in the form of the differential equation is restricted.

An approach to the dynamical systems identification based on a self-configuring genetic programming (GP) algorithm is presented in works [5-6]. The essence of this approach is to present a dynamic system model in the form of a differential equation in symbolic form, and its main value lies in the automatic determination of the order, structure and parameters of the differential equation.
The paper proposes a modification of the approach by using a differential evolution (DE) method to optimize numerical parameters of the resulting model. Thus, the algorithm proposed in this paper is a hybrid evolutionary one, since it combines two approaches from the class of evolutionary ones, i.e., a GP algorithm and a DE method.

2. Hybrid approach to the dynamic systems identification
The structure of the differential equation is selected by the GP algorithm, i.e., the differential equation is encoded as a tree (figure 1).

![Example of the differential equation in the form of a tree.](image)

Figure 1. Example of the differential equation in the form of a tree.

However, equations constants and known initial conditions are necessary to find the correct search for solutions.

The optimization of constants vector \((c_1, c_2, ..., c_n)\) will be performed for each selected structure of the differential equation, where \(n\) is the number of constants of the currently adjusted equation. Obviously, it is impossible to obtain a model of at least acceptable accuracy without any optimization method. Therefore, such a local optimization method as a steepest descent method was applied [7]. Further, this approach to identification based on the self-configuring GP algorithm and the steepest descent method will be called as a basic approach.

The article proposes such a global optimization method to be applied as the differential evolution in the hybrid algorithm to optimize the vector constants [8]. This approach will be called a hybrid one (figure 2).

In practice, values of the initial conditions are usually unknown. Therefore, they also need to be selected. For this, add the initial conditions \((c_1, c_2, ..., c_n, t_1, ..., t_k)\) to the vector of constants adjusted by the DE method, where \(k\) is the number of initial conditions (an order of the differential equation minus 1). Thus, the parallel optimization of the equation constants and initial conditions is performed. It makes no sense to perform sequential optimization, since it is impossible to find correct constants without particular initial conditions, and vice versa.
3. Self-configuring of evolutionary algorithms
In this paper, a self-configuring type of evolutionary algorithms (GP and DE) is applied. Self-configuring means the automated selection and application of the existing variants of operators and numerical parameters [9]. Operator groups are presented in Table 1.

| Operator groups | GP | Type selected operators | DE |
|-----------------|----|-------------------------|----|
|                 |    | Selection (Tournament, rank selection, proportionate) | Mutation (Rand1, Rand2, Best1, Best2, Current to Rand, Current to Best) [10] |
|                 |    | Crossover (Standard, single-point, uniform) | Crossover probability |
|                 |    | Numerical values elected operators | Differential weight |
|                 |    | Crossover probability | Crossover probability |

For operators with a selected type, self-configuring was carried out using the Population-Level Dynamic Probabilities method [11]. The numerical parameters are set according to the Success History Adaptation algorithm [12]. The application of these procedures will reduce span time significantly and will simplify the interaction with algorithms of scientists who are not specialists in the field of evolutionary algorithms.

4. Testing of approach
To study the effectiveness of the proposed approach, we compared the basic approach to identification and the hybrid one. Differential equations of various orders and number of input variables were used to...
identify differential equations (table 2). According to the conditions from table 2, samples \( \{x_i, y_i\}, i = 1, s \) were formed, where \( s \) is a sample size, for testing the proposed approach.

### Table 2. Test identification problems.

| №  | Differential equation | Starting points | Interval of \( x_j \) |
|----|-----------------------|-----------------|------------------------|
| 1  | \( y' = \frac{y \ln y}{x} \) | \( y(1) = e \) | \( x \in [1; 7] \) |
| 2  | \( y' = \frac{y^2}{x^3} + x \) | \( y(1,1) = -11.49 \) | \( x \in [1; 1.5] \) |
| 3  | \( y' = \frac{y + x^2 \cos (x)}{x} \) | \( y(0,1) = 0.06 \) | \( x \in [0; 1; 3.5] \) |
| 4  | \( y' = \frac{3x^2 e^{-x} - (x + y)y}{x} \) | \( y(1) = 0 \) | \( x \in [1; 6] \) |
| 5  | \( y' = -xe^y \) | \( y(0) = 0 \) | \( x \in [0; 5] \) |
| 6  | \( y'' = 2(y')^2 \) | \( y(0) = 2 \) | \( x \in [0; 6] \) |
| 7  | \( y'' = 6y' - 9y \) | \( y(0) = 1 \) | \( x \in [0; 6] \) |
| 8  | \( y'' = 6y' - 9y + 6xe^{3x} \) | \( y(0) = 0 \) | \( x \in [0; 6] \) |
| 9  | \( y'' = \frac{y' + x^2yy'}{x} \) | \( y(1) = 0 \) | \( x \in [1; 7] \) |
| 10 | \( y'' = \cos(3x) - 9y \) | \( y(0) = 0 \) | \( x \in [0; 6] \) |
| 11 | \( y''' = e^{2x} \) | \( y(0) = 1, 125 \) | \( x \in [0; 4] \) |
| 12 | \( y'''' = \frac{6}{x^3} \) | \( y(0) = 0 \) | \( x \in [1; 6] \) |
| 13 | \( y'''' = 4y' + 24e^{2x} - 4 \cos(2x) + 8 \sin (2x) \) | \( y(0) = 1 \) | \( x \in [0; 4] \) |
| 14 | \( y'''' = y + 2 \cos (x) \) | \( y(0) = 0 \) | \( x \in [0; 5] \) |
| 15 | \( y'''' = y'' + 2x + 3 \) | \( y(0) = 2 \) | \( x \in [0; 6] \) |
| 16 | \( y' = \frac{y + x_1^2 \cos(x_2)}{x_3} \) | \( y(0) = 0.06 \) | \( x_1 \in [0; 1; 6] \) |
| 17 | \( y' = \frac{3x_2^2 e^{-x_2} - (x_1 + y)y}{x_2} \) | \( y(0) = 0 \) | \( x_1 \in [-1; 1] \) |
| 18 | \( y' = 2x_2 \sqrt{y} + \frac{2y}{x_1} \) | \( y(0) = 1 \) | \( x_1 \in [1; 6] \) |
| 19 | \( y'' = -2y' - 2y + 2x_1^2 + 8x_2 + 6 \) | \( y(0) = 1 \) | \( x_1 \in [0; 5] \) |
| 20 | \( y'' = \frac{y' + x_2yy'}{x_2} \) | \( y(0) = 0 \) | \( x_1 \in [1; 6] \) |
The following criteria were selected to compare the efficiency of the approaches. They are an error and averaged generation number where the solution was found. The error was taken as a standard deviation of the output of the differential equation from its estimation. To identify each differential equation, 100 individuals were selected, and the number of generations was 200. The number of runs on each function was 30.

It should be noted that the resistance of the method to the noise presence in the data is an advantage over other identification methods [13-14]. Therefore, table 3 presents the results of comparing the basic algorithm and the hybrid one, both on data without noise, and with it.

**Table 3. Test results of the approaches to the identification.**

| №  | Without noise in sample data | With noise in sample data |
|----|------------------------------|---------------------------|
|    | Basic algorithm (GP) | Hybrid algorithm (GP+DE) | Basic algorithm (GP) | Hybrid algorithm (GP+DE) |
| 1  | 0.0013 | 0.0000 | 0.0784 | 0.0000 |
| 2  | 0.0072 | 0.0000 | 0.0955 | 0.0012 |
| 3  | 0.0037 | 0.0000 | 0.0496 | 0.0000 |
| 4  | 0.1047 | 0.0001 | 0.1854 | 0.0032 |
| 5  | 0.0077 | 0.0000 | 0.2071 | 0.0009 |
| 6  | 0.0983 | 0.0001 | 0.3792 | 0.0008 |
| 7  | 0.0065 | 0.0000 | 0.0140 | 0.0009 |
| 8  | 0.1200 | 0.0000 | 0.2552 | 0.0024 |
| 9  | 0.1742 | 0.0002 | 0.3250 | 0.0027 |
| 10 | 0.0084 | 0.0000 | 0.0737 | 0.0008 |
| 11 | 0.0035 | 0.0000 | 0.0091 | 0.0013 |
| 12 | 0.3202 | 0.0002 | 0.3945 | 0.0000 |
| 13 | 0.1462 | 0.0000 | 0.1974 | 0.0005 |
| 14 | 0.0893 | 0.0000 | 0.4438 | 0.0001 |
| 15 | 0.2540 | 0.0000 | 0.2984 | 0.0008 |
| 16 | 0.2985 | 0.0046 | 0.3877 | 0.0055 |
| 17 | 0.2562 | 0.0095 | 0.3648 | 0.0098 |
| 18 | 0.3724 | 0.0103 | 0.4500 | 0.0097 |
| 19 | 0.0566 | 0.0011 | 0.2003 | 0.0033 |
| 20 | 0.1499 | 0.0017 | 0.2223 | 0.0047 |

It can be seen from the data in table 3 how the hybridization of the original algorithm reduced the error. The Student’s t-test was used to test hypotheses about the significance of differences in error. Testing the hypothesis of equality of mathematical expectations by the Student’s t-test proves the presence of a statistically significant difference between errors resulted by basic and hybrid algorithms.
Figure 3. Generation number where the approaches found individuals.

Figure 3 demonstrates a generation number of the evolutionary approaches where the approaches found individuals providing an error of less than 0.1 for each task. If an error exceeds this value, a number of the generation is indicated where the individual providing the smallest error was found. Figure 3 shows that a hybrid approach requires significantly fewer iterations of the evolutionary algorithm.

5. Conclusion
The paper presents a hybrid algorithm for the dynamical systems identification, combining a self-configuring genetic programming algorithm for finding the structure of the model in the form of the differential equation and differential evolution method for optimizing numerical parameters of the differential equation and starting point. The application of the global optimization method made it possible to reduce an error of the resulting models significantly.

The further research will be related to modifying the approach for models building in the form of the differential equations system.

Acknowledgments
The research was carried out with the financial support of the project № FEFE-2020-0013.

References
[1] Pashchenko F F 2006 Introduction to consistent methods of systems modeling Mathematical foundations of systems modeling (Moscow) p 328
[2] Ryzhikov I and Semenkin E 2013 Evolutionary strategies algorithm based approaches for the linear dynamic system identification Lecture Notes in Computer Science 7824 477-84
[3] Kantor O G, Spivak S I and Podvalnyi S L 2020 An analysis of the models informativeness in parametric identification problems Journal of Physics: Conf Series Applied Mathematics Computational Science and Mechanics: Current Problems 012090
[4] Sokolov S V and Kucherenco P A 2007 Nonlinear parametric identification based on the criterion of minimum probability of estimation error Automatic Control and Computer Sciences 41(6) 299-304
[5] Karaseva T S 2020 Automatic differential equations identification by self-configuring genetic programming algorithm IOP Conf Series: Materials Science and Engineering 12093
[6] Karaseva T S 2019 Identification of first-order differential equations by the self-configuring genetic programming algorithm Materials of the International Scientific and Research Conference Topical Issues in Aeronautics and astronautics 4-6
[7] Deift P and Zhou X 1993 A steepest descent method for oscillatory Riemann-Hilbert problems.
Asymptotics for the MKdV equation *The Annals of Mathematics* 137(2) 295-368

[8] Storn R and Price K 1997 Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces *Journal of Global Optimization* 11 341-59

[9] Semenko E S and Semenkina M E 2012 Self-configuring Genetic Algorithm with Modified Uniform Crossover Operator *Lecture Notes in Computer Science* 7331 414-21

[10] Mitrofanov S A 2018 Optimization of a spacecraft’s flight trajectory with differential evolution technique *Materials of the International Scientific and Research Conference Topical Issues in Aeronautics andastronautics* 4(14) 81-3

[11] Niehaus J and Banzhaf W 2001 Adaption of Operator Probabilities in Genetic Programming *Lecture Notes in Computer Science* 2038 325-36

[12] Tanabe R and Fukunaga A 2013 Success-History Based Parameter Adaptation for Differential Evolution *Graduate School of Arts and Sciences* (Tokyo) 71-8

[13] Medvedev A V 1975 Identification and control for linear dynamic system of unknown order *Optimization Techniques IFIP Technical Conf* (Berlin–Heidelberg–New York: Springer-Verlag) p 48-56

[14] Karaseva T S 2018 About efficiency of identification of first-order differential equations by the self-configuring genetic programming algorithm *Materials of the International Scientific and Research Conference Reshetnev readings* p 168-9