

UDC 519.6

Doi: 10.31772/2587-6066-2019-20-4-436-442

For citation: Pavlenko A. A. Comparison of methods for initializing starting points on the optimization genetic algorithm. *Siberian Journal of Science and Technology*. 2019, Vol. 20, No. 4, P. 436–442. Doi: 10.31772/2587-6066-2019-20-4-436-442

Для цитирования: Павленко А. А. Сравнение способов инициализации начальных точек на генетическом алгоритме оптимизации // Сибирский журнал науки и технологий. 2019. Т. 20, № 4. С. 436–442. Doi: 10.31772/2587-6066-2019-20-4-436-442

COMPARISON OF METHODS FOR INITIALIZING STARTING POINTS ON THE OPTIMIZATION GENETIC ALGORITHM

A. A. Pavlenko

Reshetnev Siberian State University of Science and Technology
31, Krasnoyarsky Rabochy Av., Krasnoyarsk, 660037, Russian Federation
E-mail: saaprepod@mail.ru

The way to initialize the starting points for optimization algorithms is one of the main parameters. Currently used methods of initializing starting points are based on stochastic algorithms of spreading points. In a genetic algorithm, points are Boolean sets. These lines are formed in different ways. They are formed directly, using random sequences (with uniform distribution law) or formed using random sequences (with uniform distribution law) in the space of real numbers, and then converted to boolean numbers. Six algorithms for constructing multidimensional points for global optimization algorithms of boolean sets based on both stochastic and non-random point spreading algorithms are designed. The first four methods of initialization of Boolean lines used a random distribution law, and the fifth and sixth methods of initialization used a non-random method of forming starting points-LP τ sequence. A large number of optimization algorithms were restarted. Calculations of high accuracy were used. The research was carried out on the genetic algorithm of global optimization. The work is based on Acly function, Rastrigin function, Shekel function, Griewank function and Rosenbrock function. The research was based on three algorithms of starting points spreading: LP τ sequence, UDC sequence, regular random spreading. The best parameters of the genetic algorithm of global optimization were used in the work. As a result, we obtained arrays of mathematical expectations and standard deviations of the solution quality for different functions and optimization algorithms. The purpose of the analysis of ways to initialize the starting points for the genetic optimization algorithm was to find the extremum quickly, accurately, cheaply and reliably simultaneously. Methods of initialization were compared with each other by expectation and standard deviation. The quality of the solution is understood as the average error of finding the extremum. The best way of initialization of starting points for genetic optimization algorithm on these test functions is revealed.

Keywords: genetic optimization algorithm, points initialization methods.

СРАВНЕНИЕ СПОСОБОВ ИНИЦИАЛИЗАЦИИ НАЧАЛЬНЫХ ТОЧЕК НА ГЕНЕТИЧЕСКОМ АЛГОРИТМЕ ОПТИМИЗАЦИИ

А. А. Павленко

Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева
Российская Федерация, 660037, г. Красноярск, просп. им. газ. «Красноярский рабочий», 31
E-mail: saaprepod@mail.ru

Способ инициализации начальных точек для алгоритмов оптимизации является одним из главных параметров. Сегодня используются способы инициализации начальных точек, основанные на стохастических алгоритмах разброса точек. В генетическом алгоритме точки представляют собой булевы строки. Эти строки формируются по-разному: напрямую с помощью случайных последовательностей (с равномерным законом распределения) или с помощью случайных последовательностей (с равномерным законом распределения) в пространстве вещественных чисел, а потом преобразуют вещественные числа в булевы. Спроектированы шесть алгоритмов построения многомерных точек для алгоритмов глобальной оптимизации – булевых строк, основанные как на стохастических, так и на неслучайных алгоритмах разброса точек. В первых четырех способах инициализации булевых строк использовался случайный закон распределения, а в четвертом и пятом способе инициализации использовался неслучайный способ формирования начальных точек – ЛП τ последовательность. Применялось большое количество повторных запусков алгоритмов оптимизации. Использовалась достаточно высокая точность вычислений. Исследования проводились на генетическом алгоритме глобальной

оптимизации. Использовались функция Акли, функция Растргина, функция Шекеля, функция Гриванка и функция Розенброка. Исследования проводились с использованием трех алгоритмов разброса начальных точек: ЛПт последовательность, UDC последовательность, равномерный случайный разброс. В работе использовались лучшие параметры генетического алгоритма глобальной оптимизации. На выходе получены массивы математических ожиданий и среднеквадратических отклонений качества решения для разных функции и оптимизационных алгоритмов. Цель анализа способов инициализации начальных точек для генетического оптимизационного алгоритма заключалась в нахождении экстремума одновременно быстро, точно, дешево и надёжно. Способы инициализации сравнивались между собой по математическому ожиданию и среднеквадратическому отклонению. Под качеством решения понимается среднестатистическая ошибка нахождения экстремума. Выявлен лучший способ инициализации начальных точек для генетического алгоритма оптимизации на данных тестовых функциях.

Ключевые слова: генетический алгоритм оптимизации, способы инициализации точек.

Introduction. The genetic algorithm of global optimization [1] differs from the others in Economics, Finance, and banking. In the genetic algorithm, points are represented as boolean sets. These lines can be formed in different ways: directly, using random sequences (with regular distribution law) or spreading points using random sequences (with a regular distribution law) in the space of real numbers, and then convert them into Boolean ones. The research is based on Acly function, Rastrigin function, Shekel function, Griewank function and Rosenbrock function. [2]. LPт sequence [3], UDC sequence, uniform random spread are very interesting and efficient algorithms for spreading starting points. Recent studies in this area were carried out in the works [4]. These studies were applied to specific practical problems, there was no goal in averaging these parameters, in testing on a large number of practical problems of a complex type of the tested function [5]. LPт sequences is a point spreading algorithm based on the matrix of irreducible Marshall Polynomials. UDC sequence is an algorithm for regular distribution of points over all coordinates in multidimensional space [6] regardless of the number of spread points [7]. Regular random spread [8] is a stochastic point spreading algorithm using the normal distribution law.

The best parameters of the genetic algorithm of global optimization were used [9].

Parameter description. The paper used the best parameters of the genetic algorithm of global optimization:

1. Selection – tournament 4 participants.

2. Recombination – 2-point, the probability of crossing – 0.8.

3. Probability of mutation 0.001

4. Binary coding.

Additional parameters used:

1. The size of the space of the studied regularity – 2.
2. Total number of spreading points (boolean sets) – 50.
3. The maximum number of steps of the algorithm (generations) – 200.
4. The accuracy of finding the extremum – 0.0001.
5. The number of repeated runs of the algorithm – 400.
6. Boundaries of the study area отfrom –45 to +45 on each coordinate.

We used six methods of initialization:

Method 1: using random sequences (with regular distribution law), a real number from 1 to 100 is obtained. If the resulting number is less than 50, the corresponding bit of the boolean set [10] takes the value 1, otherwise 0. This produces boolean sets (fig. 1).

Method 2: using random sequences [11], we obtain a real number from 1 to K, where K is the maximum real number obtained if each bit of the boolean set is equal to 1. Then, this real number is converted to a boolean set by the conversion rules. This produces boolean sets (fig. 2).

Method 3. The third method is equal to the first but with the checking of boolean sets for repeatability (fig. 3).

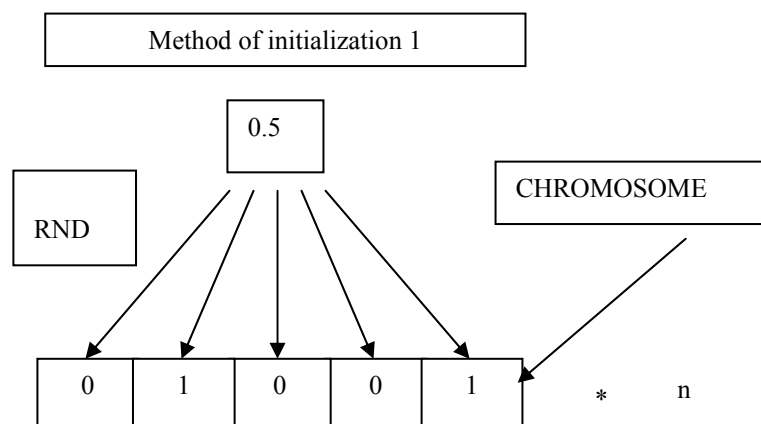


Fig. 1. The first method of initialization

Рис. 1. Первый способ инициализации

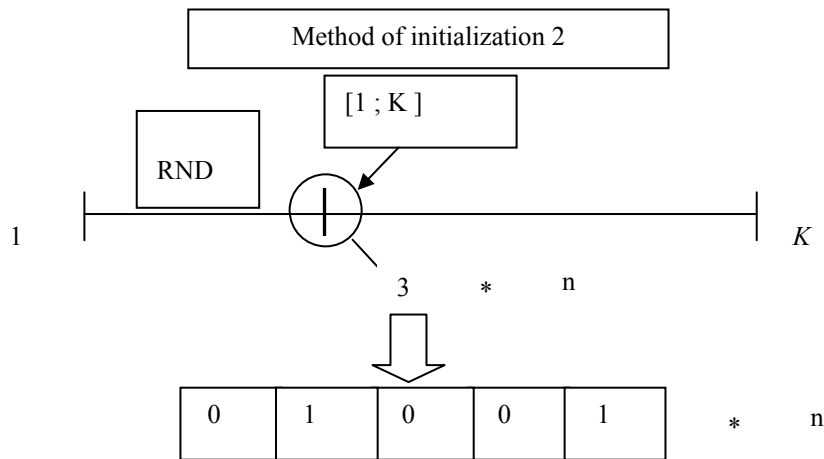


Fig. 2. The second method of initialization

Рис. 2. Второй способ инициализации

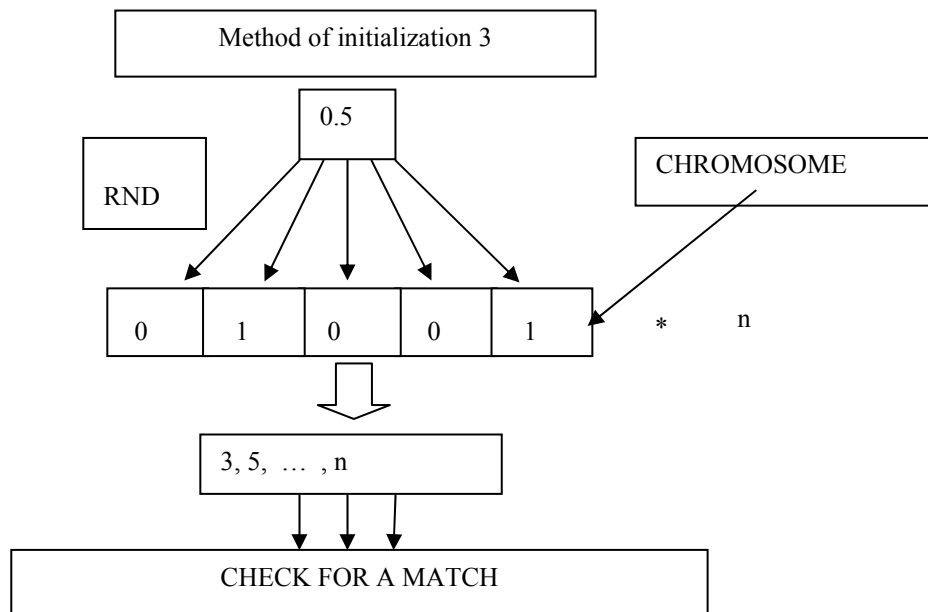


Fig. 3. The third method of initialization

Рис. 3. Третий способ инициализации

Method 4. The fourth method is equal to the second one but with the checking of real numbers for repeatability (fig. 4).

Method 5: with the help of the $LP\tau$ sequence, a real number from 1 to 100 is obtained. If the resulting number is less than 50, the corresponding bit of the boolean set will be 1, otherwise 0. This produces boolean sets (fig. 5).

Method 6: using the $LP\tau$ sequence, we obtain a real number from 1 to K, where K is the maximum real number obtained if each bit of the boolean set was equal to 1. Then, this real number is converted into a boolean set (fig. 6).

As a result, we get arrays of mathematical expectations and standard deviations of the solution quality for different functions and optimization algorithms [12]. The quality of the solution is the average error of finding the extremum [13].

Recent experiments in this area were carried out in the works of the scientist [7]. In his works, binary lines were formed without taking into account the check for repeatability.

Experimental part. The initialization methods are compared with each other on a genetic algorithm, according to math expectation, on the first function. The best way to initialize on a six-point scale is determined. Then they are compared on the second function, etc.

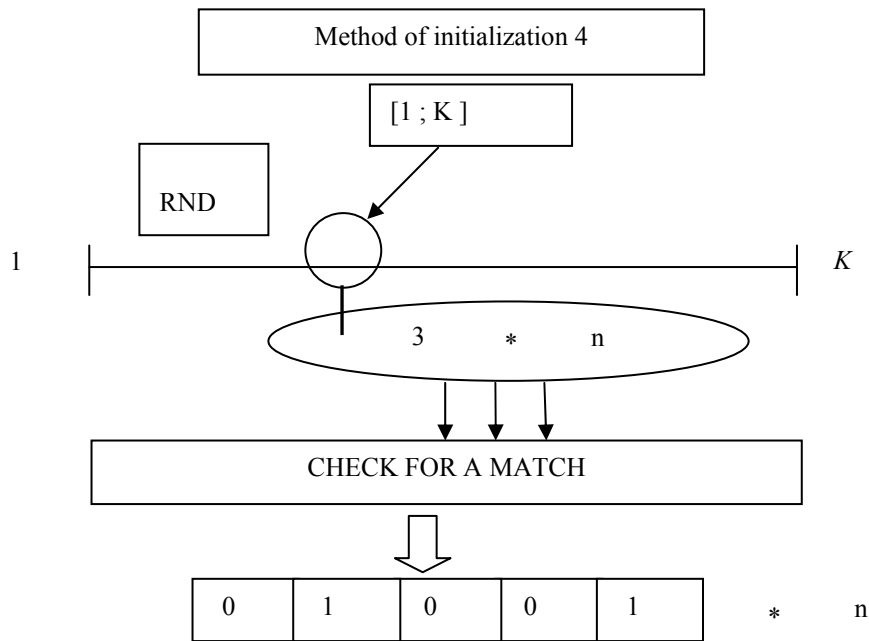


Fig. 4. The fourth method of initialization

Рис. 4. Четвёртый способ инициализации

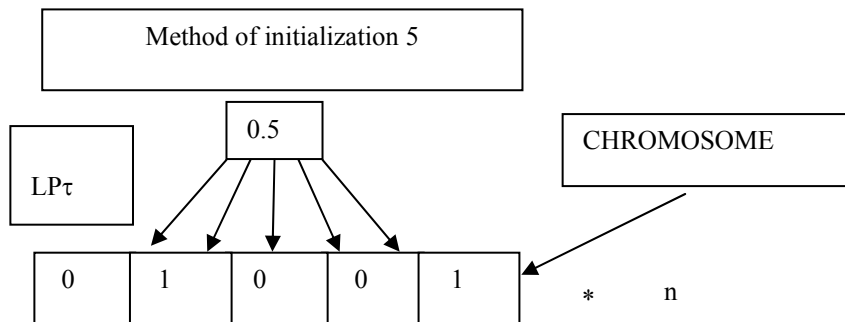


Fig. 5. The fifth method of initialization

Рис. 5. Пятый способ инициализации

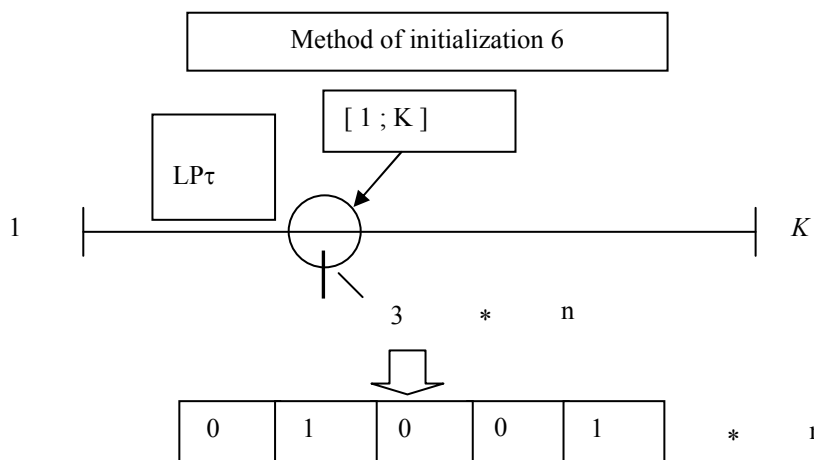


Fig. 6. The sixth method of initialization

Рис. 6. Шестой способ инициализации

Table 1

Percentage of finding the extremum for the algorithm by absolute value

		Method of initialization (I)						Best I
		1	2	3	4	5	6	
Function	1	4.00	3.50	2.75	2.25	3.50	3.25	1
	2	7.25	7.50	7.25	7.75	8.50	7.25	5
	3	0	0	0	0	0	0	—
	4	0	0.25	0	0.75	1.00	0.25	5
Total								5

Table 2

The quality of the solution for the Genetic algorithm by absolute value

Function		Method of initialization (I)						Best I by M	Best I by σ
		1	2	3	4	5	6		
1	M	0.2431	0.2453	0.2511	0.2401	0.2231	0.2370	5	2
	σ	0.2127	0.2035	0.2221	0.2177	0.2076	0.2077		
2	M	0.0613	0.0585	0.0629	0.0586	0.0533	0.0598	5	5
	σ	0.0512	0.0495	0.0725	0.0553	0.0310	0.0456		
3	M	0.5001	0.5478	0.5091	0.4904	0.4433	0.4455	5	5
	σ	0.3947	0.5188	0.3718	0.3459	0.0055	0.0422		
4	M	1.7635	1.7237	1.7866	1.7827	0.9066	1.9681	5	5
	σ	1.5675	1.7093	1.6367	1.6068	0.4152	1.8819		
Places by M		5	4	6	2	1	3		
Places by σ		3	3	5	4	1	2		
Total		5	4	6	3	1	2		

After that, the points obtained for all functions for each initialization method are summed up [14]. The larger the amount, the better place takes one or another method of initialization. The places got in this way are recorded in the tab. 1 and 2.

Then the whole cycle is repeated for the standard deviation (σ). After that, the obtained places by math expectation (M) and by σ are summed up and compared again [15]. Eventually we will know:

1. The best initialization method for each of the functions when using a genetic algorithm, both by math expectation and by σ .

2. The best and the worst methods of initialization on average for all four optimization functions, both by math expectation and by σ , and in general, when using a genetic algorithm.

Results. As we know, the goal of any optimization algorithm is to find the extremum as quickly, more accurately, cheaper and more reliable as possible.

The percentage of finding the extremum, on average, for all optimization functions is better for the initialization method 5 (tab. 2), it follows that the initialization method 5 has the greatest reliability (probability) of finding the extremum.

We prove that at 50 points, the algorithm calculates 200 steps (generations), on average, for 1/400 seconds.

Accuracy is the inverse of an error. The higher the accuracy, the smaller the error. The higher the accuracy obtained, the smaller the number takes the characteristic on a six-point scale, which means that the characteristic is better.

For the genetic algorithm (tab. 1) the smallest error of finding the extremum (quality of solution) according to the expected value for the first function characterizes the method of initialization 5, this means that the method 5 has the highest accuracy of finding the extremum. For the genetic algorithm, the initialization method 5 also has the highest accuracy of finding the extremum by mathematical expectation for the function 2–4. On average, for the genetic algorithm, the initialization method 5 has the highest accuracy of finding the extremum.

For the genetic algorithm, the smallest error of finding the extremum by σ for the function 1 (error) is in the initialization method 2, and for the function 2–4 – in the initialization method 5. On average, for a genetic algorithm, the smallest error is that the resulting expectation value is the average expectation value of the initialization method 5.

Since, on average, for the genetic algorithm, the highest accuracy of finding the extremum and the smallest error of the result obtained is in the initialization method 5, then, on average, the initialization method 5 is the best initialization method for the genetic algorithm and takes the first place.

To compare the method of initialization 5, for example, with the applied now numerical method of initialization 1, the comparison of these methods in percentage being conducted in relation to the method of initialization 1 for each function, and then taken the arithmetic average of these percentages for all functions.

As a result, it became obvious that in the genetic algorithm, the initialization method 5 is better than the initialization method 1 by an average of 20 % and the application of check in boolean initialization does not lead to positive result, and in real – leads.

Conclusion. The genetic algorithm of global optimization based on Acly function, Rastrigin function, Shekel function, Griewank function and Rosenbrock function is analyzed. The research was carried out on six methods of initialization of boolean sets, using three algorithms of spreading initial points: LP τ sequence, UDC sequence, uniform random spread. Studies show that, on average, the initialization method 5 is the best initialization method for the genetic algorithm, and it follows that the initialization method 5 is very promising. It must be taken into account that the initialization method 5 is based on a non – random algorithm of spreading points LP τ sequence. This provides an incentive for further study of the LP τ sequence, and once again confirms its effectiveness.

References

1. Zaloga A. N., Yakimov I. S., Dubinin P. S. Multi-population Genetic Algorithm for Determining Crystal Structures Using Powder Diffraction Data. *Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques*. 2018, Vol. 12, No. 1, P. 128–134.
2. Stanovov V., Akhmedova S., Seminkin E. Automatic Design of Fuzzy Controller for Rotary Inverted Pendulum with Success-History Adaptive Genetic Algorithm. *2019 International Conference on Information Technologies (InfoTech)*. IEEE, 2019. P. 1–4.
3. Zaloga A. et al. Genetic Algorithm for Automated X-Ray Diffraction Full-Profile Analysis of Electrolyte Composition on Aluminium Smelters. *Informatics in Control, Automation and Robotics 12th International Conference, ICINCO 2015 Colmar, France, July 21–23, 2015 Revised Selected Papers*. Springer, Cham, 2016, P. 79–93.
4. Du X. et al. Genetic algorithm optimized non-destructive prediction on property of mechanically injured peaches during postharvest storage by portable visible/shortwave near-infrared spectroscopy. *Scientia Horticulturae*. 2019, Vol. 249, P. 240–249.
5. Akhmedova S., Stanovov V., Seminkin E. Soft Island Model for Population-Based Optimization Algorithms. *International Conference on Swarm Intelligence*. Springer, Cham, 2018, P. 68–77.
6. Yakimov I. et al. Application of Evolutionary Rietveld Method Based XRD Phase Analysis and a Self-Configuring Genetic Algorithm to the Inspection of Electrolyte Composition in Aluminum Electrolysis Baths. *Crystals*. 2018, Vol. 8, No. 11, P. 402.
7. Chen W. et al. Applying population-based evolutionary algorithms and a neuro-fuzzy system for modeling landslide susceptibility. *Catena*. 2019, Vol. 172, P. 212–231.
8. Karabadzha G. et al. Semi-Empirical Method for Evaluation of a Xenon Operating Hall Thruster Erosion Rate Through Analysis of its Emission Spectra. *Spacecraft Propulsion*. 2000, Vol. 465, P. 909.
9. Akhmedova S., Stanovov V., Seminkin E. Success-History Based Position Adaptation in Co-operation of Biology Related Algorithms. *International Conference on Swarm Intelligence*. Springer, Cham, 2019, P. 39–49.
10. Yakimov I., Zaloga A., Dubinin P., Bezrukova O., Samoilov A., Burakov S., Seminkin E., Semenkina M., Andruschenko E. Application of Evolutionary Rietveld Method Based XRD Phase Analysis and a Self-Configuring Genetic Algorithm to the Inspection of Electrolyte Composition in Aluminum Electrolysis Baths. *Crystals*. 2018, Vol. 8, No. 11, P. 402.
11. Brester C., Rönkkö M., Kolehmainen M., Seminkin E., Kauhanen J., Tuomainen T. P., Voutilainen S., Ronkainen K. Evolutionary methods for variable selection in the epidemiological modeling of cardiovascular diseases. *BioData Mining*. 2018, Vol. 11, No. 18.
12. Mamontov D. et al. Evolutionary Algorithms for the Design of Neural Network Classifiers for the Classification of Pain Intensity. *IAPR Workshop on Multimodal Pattern Recognition of Social Signals in Human-Computer Interaction*. Springer, Cham, 2018, P. 84–100.
13. Brester C. et al. On a restart metaheuristic for real-valued multi-objective evolutionary algorithms. *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. ACM, 2019, P. 197–198.
14. Sharifpour E. et al. Zinc oxide nanorod loaded activated carbon for ultrasound assisted adsorption of safranin O: Central composite design and genetic algorithm optimization. *Applied Organometallic Chemistry*. 2018, Vol. 32, No. 2, P. e4099.
15. Penenko A. V. 2019 Newton–Kantorovich method for solving inverse problems of source identification in product–destruction models with time series data. *Siberian J. of computational mathematics*. 2019, No. 1, P. 57–79.

Библиографические ссылки

1. Zaloga A. N., Yakimov I. S., Dubinin P. S. Multi-population Genetic Algorithm for Determining Crystal Structures Using Powder Diffraction Data // *Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques*. 2018. Vol. 12, No. 1. P. 128–134.
2. Stanovov V., Akhmedova S., Seminkin E. Automatic Design of Fuzzy Controller for Rotary Inverted Pendulum with Success-History Adaptive Genetic Algorithm // *2019 International Conference on Information Technologies (InfoTech)*. IEEE, 2019. P. 1–4.
3. Genetic Algorithm for Automated X-Ray Diffraction Full-Profile Analysis of Electrolyte Composition on Aluminium Smelters / A. Zaloga et al. // *Informatics in*

Control, Automation and Robotics 12th International Conference, ICINCO 2015 Colmar, France, July 21–23, 2015 Revised Selected Papers. Springer, Cham, 2016. P. 79–93.

4. Genetic algorithm optimized non-destructive prediction on property of mechanically injured peaches during postharvest storage by portable visible/shortwave near-infrared spectroscopy / X. Du et al. // *Scientia Horticulturae*. 2019. Vol. 249. P. 240–249.

5. Akhmedova S., Stanovov V., Semenkin E. Soft Island Model for Population-Based Optimization Algorithms // *International Conference on Swarm Intelligence*. Springer, Cham, 2018. P. 68–77.

6. Application of Evolutionary Rietveld Method Based XRD Phase Analysis and a Self-Configuring Genetic Algorithm to the Inspection of Electrolyte Composition in Aluminum Electrolysis Baths / I. Yakimov et al. // *Crystals*. 2018. Vol. 8, No. 11. P. 402.

7. Applying population-based evolutionary algorithms and a neuro-fuzzy system for modeling landslide susceptibility / W. Chen et al. // *Catena*. 2019. Vol. 172. P. 212–231.

8. Semi-Empirical Method for Evaluation of a Xenon Operating Hall Thruster Erosion Rate Through Analysis of its Emission Spectra / G. Karabadzha et al. // *Spacecraft Propulsion*. 2000. Vol. 465. P. 909.

9. Akhmedova S., Stanovov V., Semenkin E. Success-History Based Position Adaptation in Co-operation of Biology Related Algorithms // *International Conference on Swarm Intelligence*. Springer, Cham. 2019. P. 39–49.

10. Application of Evolutionary Rietveld Method Based XRD Phase Analysis and a Self-Configuring Genetic Algorithm to the Inspection of Electrolyte Composition in Aluminum Electrolysis Baths / I. Yakimov, A. Zaloga, P. Dubinin et al. // *Crystals*. 2018. Vol. 8, No. 11. P. 402.

11. Evolutionary methods for variable selection in the epidemiological modeling of cardiovascular diseases / C. Brester, M. Rönkkö, M. Kolehmainen et al. // *BioData Mining*. 2018. Vol. 11, No. 18.

12. Evolutionary Algorithms for the Design of Neural Network Classifiers for the Classification of Pain Intensity / D. Mamontov et al. // *IAPR Workshop on Multimodal Pattern Recognition of Social Signals in Human-Computer Interaction*. Springer, Cham, 2018. P. 84–100.

13. On a restart metaheuristic for real-valued multi-objective evolutionary algorithms / C. Brester et al. // *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. ACM, 2019. P. 197–198.

14. Zinc oxide nanorod loaded activated carbon for ultrasound assisted adsorption of safranin O: Central composite design and genetic algorithm optimization / E. Sharifpour et al. // *Applied Organometallic Chemistry*. 2018. Vol. 32, No. 2. P. e4099.

15. Penenko A. V. Newton–Kantorovich method for solving inverse problems of source identification in product–destruction models with time series data // *Siberian J. of computational mathematics*. 2019. No. 1. P. 57–79.

© Pavlenko A. A., 2019

Pavlenko Alexander Alexandrovich – Senior Lecturer; Reshetnev Siberian State University of Science and Technology. E-mail: saaprepod@mail.ru.

Павленко Александр Александрович – старший преподаватель кафедры информационных экономических систем; Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева, инженерно-экономический институт. E-mail: saaprepod@mail.ru.
