

**Данный текст является русскоязычной версией опубликованной на английском языке статьи и представлен в авторской редакции только на данном сайте!**

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](mailto: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 $\tau$  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 $\tau$  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

*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. Они формируются напрямую, с помощью случайных последовательностей (с равномерным законом распределения) или с помощью случайных последовательностей (с равномерным законом распределения) в пространстве вещественных чисел, а потом преобразуются 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 – неслучайный способ формирования начальных точек – ЛПт последовательность. Применялось большое количество повторных запусков алгоритмов оптимизации. Использовалась достаточно высокая точность вычислений. Исследования проводились на генетическом алгоритме глобальной оптимизации. Использовались функция Акли, функция Растригина, 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.

**Введение.** Генетический алгоритм глобальной оптимизации [1] отличается от остальных тем, что хорошо работает в экономике, финансах, банковской сфере. В генетическом алгоритме точки представлены в виде булевых строк. Эти строки можно формировать по-разному. Можно формировать напрямую, с помощью случайных последовательностей (с равномерным законом распределения). Можно разбросать точки с помощью случайных последовательностей (с равномерным законом распределения) в пространстве вещественных чисел, а потом преобразовать вещественные числа в булевы. Исследования проводились на функции Акли, функции Растригина, функции Шекеля, функции Гриванка и функции Розенброка [2]. ЛПт последовательность [3], UDC последовательность, равномерный случайный разброс – очень интересные и эффективные алгоритмы разброса начальных точек. Последние исследования в этой области проводились в работах [4]. Эти исследования применялись к конкретным практическим задачам, не ставилась цель в усреднении данных параметров, в тестировании на большом количестве практических задач сложного вида тестируемой функции [5]. ЛПт-последовательности – это алгоритм разброса точек на основе матрицы неприводимых многочленов Маршала. UDC последовательности – это алгоритм абсолютно равномерного распределения точек по всем координатам в многомерном пространстве [6] независимо от количества разбрасываемых точек [7]. Равномерный

случайный разброс [8] – это стохастический алгоритм разброса точек, использующий нормальный закон распределения.

**Описание параметров.** В работе использовались лучшие параметры генетического алгоритма глобальной оптимизации [9]:

1. Селекция – турнирная 4 участника.
2. Рекомбинация – 2-точечная, вероятность скрещивания – 0,8.
3. Вероятность мутации 0,001.
4. Бинарное кодирование.

В работе использовались дополнительные параметры:

1. Размер пространства исследуемой закономерности – 2.
2. Общее количество разбрасываемых точек (булевых строк) – 50.
3. Максимальное число шагов алгоритма (поколений) – 200.
4. Точность нахождения экстремума – 0,0001.
5. Количество повторных запусков алгоритма – 400.
6. Границы исследуемой области от –45 до +45 по каждой координате.

В работе использовались 6 способов инициализации:

1-й способ: с помощью случайных последовательностей (с равномерным законом распределения) получают вещественное число от 1 до 100. Если полученное число будет меньше 50, то соответствующий бит булевой строки [10] принимает значение 1, иначе 0. Так получают булевые строки (рис. 1).

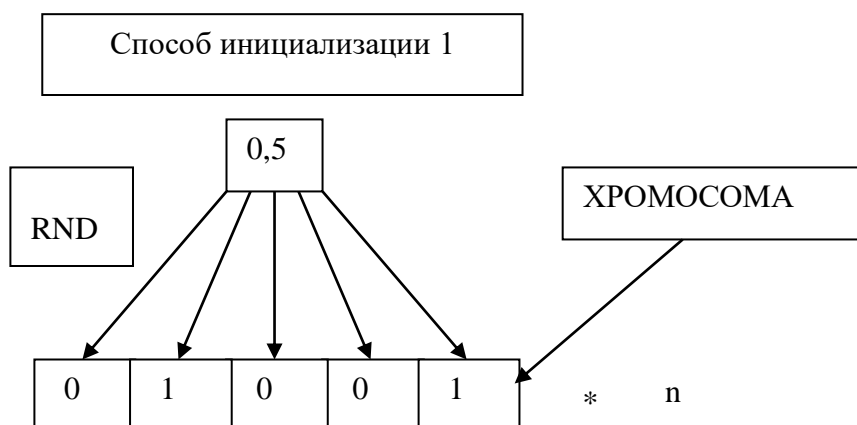


Рис. 1. Первый способ инициализации  
Fig. 1. The first way to initialize

2-й способ: с помощью случайных последовательностей [11] получают вещественное число от 1 до К, где К – максимальное вещественное число, получившееся, если бы каждый бит булевой строки был равен 1. Затем это вещественное число преобразуют в булевую строку по правилам преобразования. Так получают булевые строки (рис. 2).

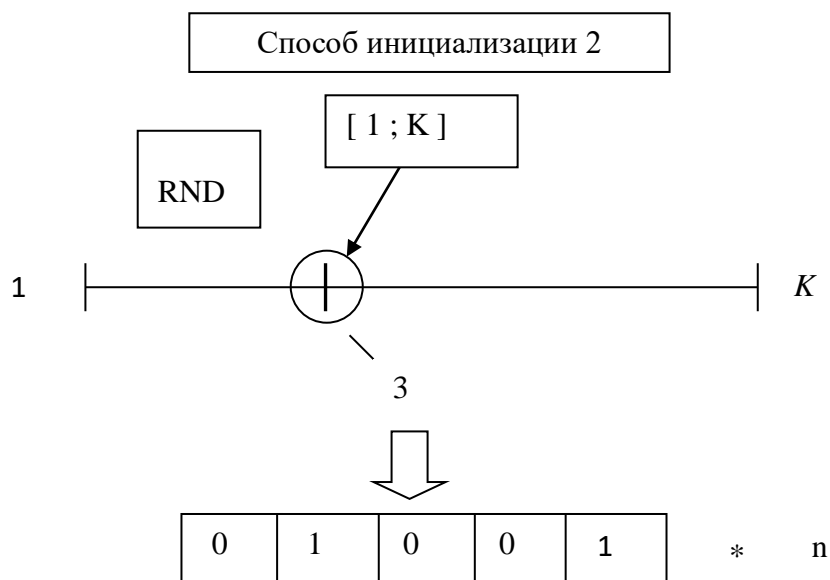


Рис. 2. Второй способ инициализации  
Fig. 2. The second way to initialize

3-й способ – это тот же 1-й способ, но с проверкой булевых строк на повторяемость (рис. 3).

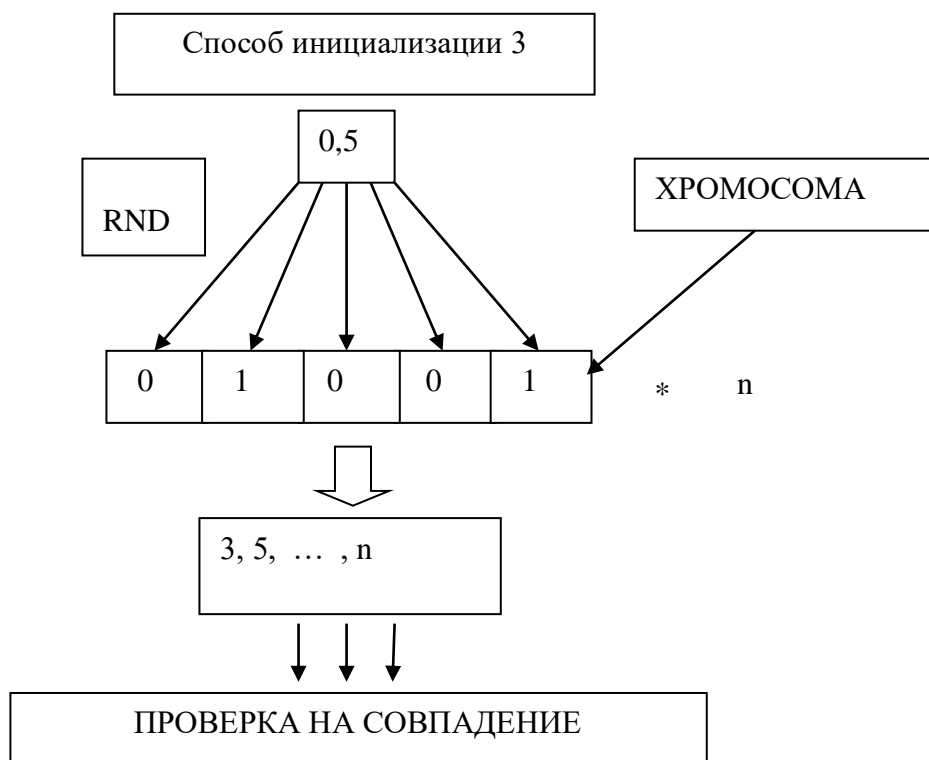


Рис. 3. Третий способ инициализации  
Fig. 3. The third way to initialize

4-й способ – это тот же 2-й способ, но с проверкой вещественных чисел на повторяемость (рис. 4).

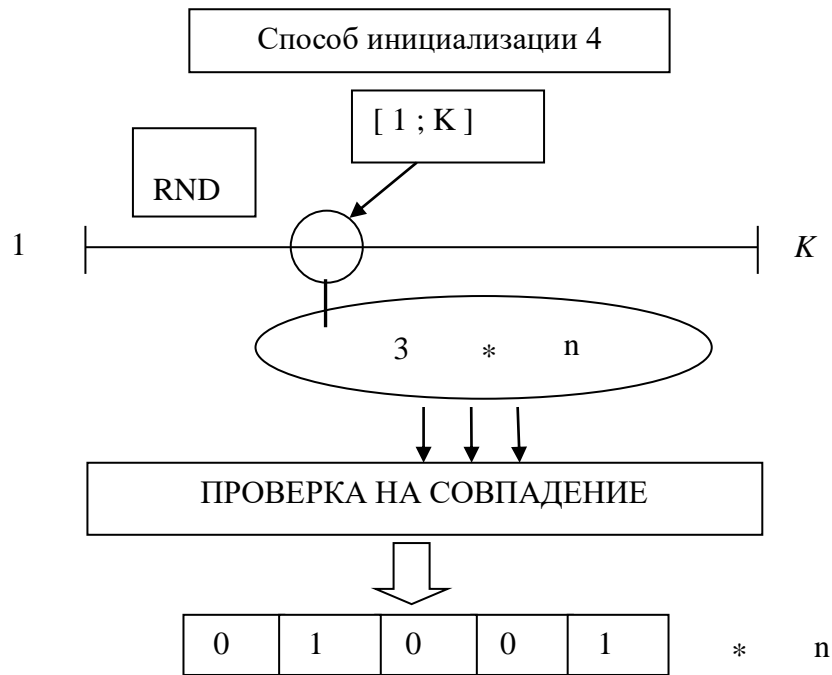


Рис. 4. Четвёртый способ инициализации  
Fig. 4. Fourth Initialization Method

5-й способ: с помощью ЛПτ-последовательности получают вещественное число от 1 до 100. Если полученное число меньше 50, то соответствующий бит булевой строки будет 1, в противном случае 0. Так получают булевые строки (рис. 5).

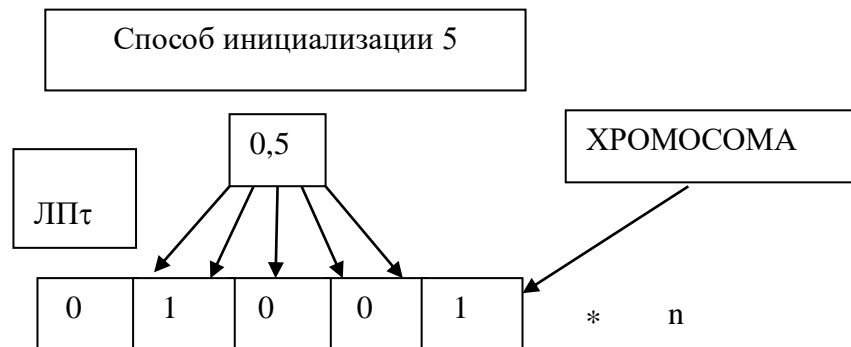


Рис. 5. Пятый способ инициализации  
Fig. 5. Fifth Initialization Method

6-й способ: с помощью ЛПτ-последовательности получают вещественное число от 1 до K, где K – максимальное вещественное число, получившееся, если бы каждый бит булевой строки был равен 1. Затем это вещественное число преобразуют в булевую строку (рис. 6).

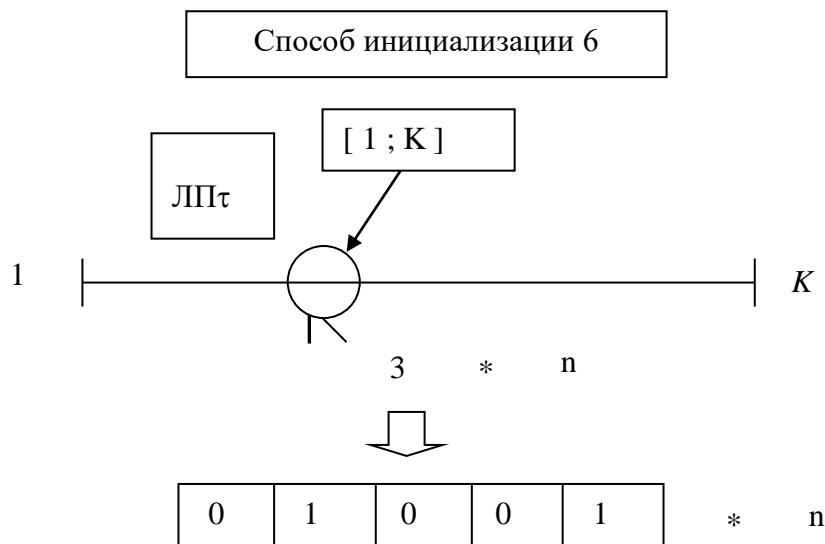


Рис. 6. Шестой способ инициализации  
Fig. 6. Sixth way to initialize

На выходе мы получаем массивы математических ожиданий и среднеквадратических отклонений качества решения для разных функций и оптимизационных алгоритмов [12]. Под качеством решения понимается среднестатистическая ошибка нахождения экстремума [13].

Последние эксперименты в данной области проводились в работах [7]. В данных трудах бинарные строки формировались без учета проверки на повторяемость.

**Экспериментальная часть.** Способы инициализации сравниваются между собой на генетическом алгоритме, по математическому ожиданию, на первой функции. Определяется лучший способ инициализации по шестибальной шкале. Затем они сравниваются между собой на второй функции и т. д.

После этого суммируются заработанные баллы по всем функциям для каждого способа инициализации [14]. Чем больше получится сумма, тем лучшее место займёт тот или иной способ инициализации. Заработанные таким образом места фиксируются в табл. 1 и 2.

Таблица 1

**Процент нахождения экстремума для алгоритма по абсолютному значению**

		Способы инициализации (I)						Лучшая I
		1	2	3	4	5	6	
Функции	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
Итого								5

Таблица 2

**Качество решения для Генетического алгоритма по абсолютному значению**

Функция		Способы инициализации (I)						Лучший I по М	Лучший I по σ
		1	2	3	4	5	6		
1	М	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	М	0,0613	0,0585	0,0629	0,0586	0,0533	0,0598	5	5
	$\sigma$	0,0512	0,0495	0,0725	0,0553	0,0310	0,0456		
3	М	0,5001	0,5478	0,5091	0,4904	0,4433	0,4455	5	5
	$\sigma$	0,3947	0,5188	0,3718	0,3459	0,0055	0,0422		
4	М	1,7635	1,7237	1,7866	1,7827	0,9066	1,9681	5	5
	$\sigma$	1,5675	1,7093	1,6367	1,6068	0,4152	1,8819		
Места по М		5	4	6	2	1	3		
Места по $\sigma$		3	3	5	4	1	2		
Итого мест		5	4	6	3	1	2		

Потом весь цикл повторяется для среднеквадратического отклонения ( $\sigma$ ). После этого заработанные места по математическому ожиданию (М) и по  $\sigma$  суммируются и снова сравниваются между собой [15]. В итоге мы будем знать:

1. Какой способ инициализации лучший для каждой из функций при использовании генетического алгоритма, как по математическому ожиданию, так и по  $\sigma$ .

2. Какие способы инициализации лучшие, какие худшие в среднем для всех четырёх оптимизационных функций, как по математическому ожиданию и  $\sigma$ , так и в общем при использовании генетического алгоритма.

**Результаты.** Как мы знаем, цель любого оптимизационного алгоритма заключается в том, чтобы найти экстремум как можно быстрее, точнее, дешевле и надёжнее.

Процент нахождения экстремума, в среднем, по всем оптимизационным функциям лучше у способа инициализации 5 (табл. 2). Отсюда следует, что у способа инициализации 5 самая большая надёжность (вероятность) отыскания экстремума.

С точки зрения скорости можно сказать, что при 50-ти точках алгоритм просчитывает 200 шагов (поколений), в среднем, за 1/400 с.

Точность – это обратная характеристика ошибки. Точность тем выше, чем меньше ошибка. Чем выше получается точность, тем меньше число принимает характеристика по шестизначной шкале, а это означает, что характеристика лучше.

Для генетического алгоритма (табл. 1) самая маленькая ошибка нахождения экстремума (качество решения) по математическому ожиданию для 1-й функций у способа инициализации 5. Это означает, что у способа инициализации 5 самая высокая точность нахождения экстремума. Для генетического алгоритма самая большая точность нахождения экстремума по математическому ожиданию для функций 2, 3, 4 тоже у способа инициализации 5. В среднем, для генетического алгоритма самая высокая точность нахождения экстремума у способа инициализации 5.

Для генетического алгоритма самая маленькая ошибка нахождения экстремума по  $\sigma$  для функции 1 (погрешность) у способа инициализации 2, а для функции 2, 3, 4 – у способа инициализации 5. В среднем, для генетического алгоритма самая маленькая погрешность того, что полученное значение математического ожидания есть среднестатистическое значение математического ожидания у способа инициализации 5.

Так как, в среднем, для генетического алгоритма самая высокая точность нахождения экстремума и самая маленькая погрешность полученного результата у способа инициализации 5, то, в среднем, способ инициализации 5 является лучшим способом инициализации для генетического алгоритма и занимает первое место.

Чтобы попытаться сравнить способ инициализации 5, например, с используемым сейчас способом инициализации 1 в числовом виде, были сделаны сравнения этих способов инициализации в процентах относительно способа инициализации 1 по каждой функции, а потом взято среднее арифметическое этих процентов по всем функциям.

В результате оказалось, что в генетическом алгоритме способ инициализации 5 лучше способа инициализации 1 в среднем на 20 %.

По результатам можно сделать вывод, что применение проверки в булевой инициализации не приводит к положительному результату, а в вещественной – приводит.

**Заключение.** Проанализирован генетический алгоритм глобальной оптимизации. Исследования проводились на функции Акли, функции Растригина, функции Шекеля, функции Гриванка на шести способах инициализации булевых строк с использованием трех алгоритмов разброса начальных точек: ЛПт последовательность, UDC последовательность, равномерный случайный разброс. Проведённые исследования показали, что в среднем способ инициализации 5 является лучшим способом инициализации для генетического алгоритма, из чего следует, что способ инициализации 5 является очень перспективным. Необходимо учесть, что способ инициализации 5 построен на неслучайном алгоритме разброса точек – ЛПт-последовательности. Это даёт стимул для дальнейшего исследования ЛПт-последовательности и лишний раз подтверждает эффективность её применения.

### Библиографические ссылки

1. Zaloga A. N., Yakimov I. S., Dubinin P. S. Multipopulation 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., Semenko 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., Semenko 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., Semenko 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.

## References

1. Zaloga A. N., Yakimov I. S., Dubinin P. S. Multipopulation 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., Semenkin 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., Semenkin 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., 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. Yakimov I., Zaloga A., Dubinin P., Bezrukova O., Samoilo A., Burakov S., Semenkin 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., Semenkin 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.

© Павленко А. А., 2019

**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.

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