Xiaomeng Pan Henan Vocational and Technical Institute, Zhengzhou, Henan, China

GENETIC-BEE COLONY DUAL-POPULATION SELF-ADAPTIVE HYBRID ALGORITHM BASED ON INFORMATION ENTROPY

Сяомен Пань Хенаньський професійно-технічний інститут, м. Чженчжоу, провінція Хенань, КНР

ДВОПОПУЛЯЦІЙНИЙ САМОНАЛАГОДЖУВАЛЬНИЙ ГІБРИДНИЙ ГЕНЕТИКО-БДЖОЛИНИЙ АЛГОРИТМ НА ОСНОВІ ЕНТРОПІЇ ІНФОРМАЦІЇ

Purpose. Swarm intelligence is the intelligent behaviour represented by a kind of individuals with no or simple intelligence through any form of cluster and collaboration. The research conserns the dual-population self-adaptive hybrid algorithm based on genetic algorithm (GA) and artificial bee colony (ABC). We have obtained some important performance measures, which are helpful for swarm intelligence algorithms.

Methodology. We proposed a genetic-bee colony dual-population self-adaptive hybrid algorithm based on information entropy, which uses dual-population structure and independent evolution and which conducts information exchange through information entropy to maintain population diversity and accelerate the evolution process between the two populations when appropriate.

Findings. We first analysed the basic structure and characteristics of GA and ABC, and then the dual-population based on GA and ABC, which joined the information entropy, was presented, in the parallel operation of two relatively independent populations to accelerate the emergence of a new individual by competition between the populations; it has better effects in complex function optimization problems.

Originality. We made a combinational study of GA and ABC. Although the current biological intelligent evolutionary algorithm has greatly improved its convergence speed, it is not ideal when optimizing complicated functions. This aspect of research is still relatively few at present.

Practical value. We researched the optimization algorithm, which is applied to various research fields. Nowadays, it is a development trend to improve the original algorithm by integrating the intelligent algorithm. Dual-population algorithm can overcome the shortage of separate algorithm, and become more suitable for complex optimization problems. We provided the foundation to search for complex distributed problems without centralized control or global model.

Keywords: *optimization problem, genetic algorithm, artificial bee colony, dual-population, information entropy, hybrid algorithm*

Introduction. There are numerous complex non-linear systems in real life, making our traditional optimization methods gradually trapped in a dilemma. In nature, the amazing and complicated swarm intelligence represented by those unimpressive social and low-intelligent creatures such as ant colony, bird flock, been colony and fish swarm, give us the inspiration. Among the social creatures, single individual usually has simple intelligence, but the group formed by several individuals has shown wisdom exceeding that added by the individuals. The individuals in the group help them to fulfil complicated tasks with their cooperative self-organization ability [1].

Genetic algorithm uses randomization techniques to perform the highly-efficient search on an encoded parameter space with all the individuals in a population as the objects. In this algorithm, selection, crossover and mutation are its genetic operations. Five elements including parameter coding, setting of the initial population, design of fitness function, design of genetic operations and setting of control parameters have formed the core contents of the genetic algorithm, which has the advantages of global parallel search, simplicity, universality and strong robustness [2]. Artificial bee colony algorithm is an optimization method by simulating bee behaviours and it is a specific application of swarm intelligence. Its main feature is that it does not need to learn about the special information of the problem; instead, all it needs is to compare the advantages and disadvantages of the problem. Through local optimization behaviour of every artificial bee individual, it highlights its globally optimal value in the group with faster convergence speed. Bee colony algorithm is easy to be trapped in locally optimal solution and stagnation [3]. How can we avoid the premature convergence to improve the global search ability of the algorithm? With the increasing complexity and scale of the optimization problems, it is usually very difficult to obtain a satisfactory solution through single optimization calculation. Starting from solving the practical problems and according to the idea of complementary advantages, to make full use of the strengths of various computational intelligent techniques can achieve the optimization effect that can't be fulfilled by a single algorithm and it is also the inevitable development trend to resolve problem [4].

This paper firstly illustrates the basic principles and structures of genetic algorithm and artificial bee colony algorithm. Then it designs the genetic-bee colony self-adaptive hybrid algorithm based on information entropy. Finally, by verifying the experimental simulation, it proves that the algorithm of this paper significantly improves its convergence accuracy while maintaining excellent effectiveness and robustness.

The basic principle of the genetic algorithm. The genetic algorithm is a random search algorithm simulating the natural selection and natural genetic mechanism in the natural world. Since it directly operates on the structure object, it has no limitations on derivative and function continuity and it has inherent implicit parallelism and better global optimization ability. As an optimization method using randomization, it can automatically obtain and instruct the optimized search space and it can adjust the search direction in a self-adaptive manner with no specific rules. This algorithm directly uses fitness as the search information. When solving the problem, its search process is not limited by the continuity of the optimization function and it does not need the derivative or other supplementary information. Having high parallelism, the genetic algorithm can search multiple regions in the solution space for the information in order to lower the probability to be trapped into locally optimal solution. Its strong robustness makes it highly possible to converge to the optimal or the approximate optimal solution to the discontinuous, multi-modal and noisy problem to be resolved. With its high expandability, the genetic algorithm is easy to combine the knowledge or algorithm in other fields to solve the specific problem. As an iterative algorithm, the genetic algorithm has a group of solutions in every iteration, which is generated randomly; in the meanwhile, a group of new solutions will be generated by simulating the genetic operations of evolution and inheritance in every iteration [5]. The typical flowchart is indicated as Fig. 1.

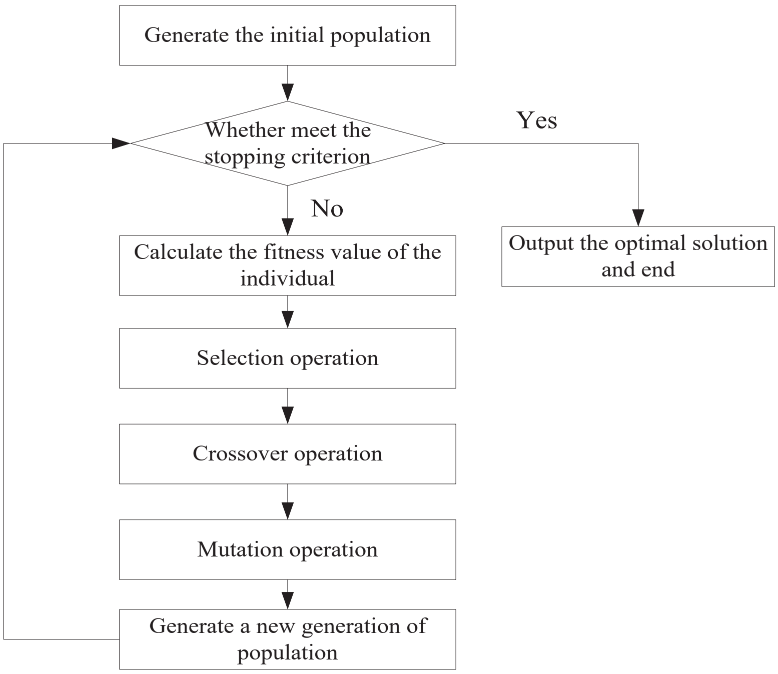

Fig. 1. Flowchart of genetic algorithm

Mechanism of artificial bee colony algorithm. The minimum search model of artificial bee colony algorithm includes the following three components: food source, employed bees and unemployed bees.

1. Food source: The value of the food source is determined by various factors such as its distance from the hive, the abundance of nectar and the difficulty to get the nectar. Abandoned food sources are determined and are replaced with the new food sources discovered by scouts. These factors are represented by a single parameter, namely the profitability of the food source.

2. Employed bees: They also called leaders and correspond to the food sources they have collected one by one. Each employed bee goes to a food source in her memory and determines a neighbour source, then evaluates its nectar amount. The employed bee has saved the relating information of a certain food source and it will share the information with other bees at certain probability.

3. Unemployed bees: They mainly search and develop food sources. There are two kinds of unemployed bees, i.e. scout and follower. Each scout watches the dance of employed bees and chooses one of their sources depending on the dances, and then goes to that source. Normally, the average scouts are 5–20% of the bee colony.

Leader, follower and scout all play an important role at different stages. The leader and the follower update the position of food source according to (1) and (3).

$$v_{ij} = x_{ij} + r_{ij}(x_{ij} - x_{kj}) . \tag{1}$$

The selection probability P_i can be calculated through (1)

$$P_i = \frac{fitness_i}{\sum_{i=1}^{n} fitness_i} ; \tag{2}$$

$$x_i^1 = x_{\min}^j + rand(0,1)(x_{\max}^j - x_{mix}^j) . \tag{3}$$

The $fitness_i$ in (2) is the fitness value of the ith solution and n is the number of the solutions. Formula (3) randomly generates a new solution to replace the eliminated solution [6].

The main procedures of artificial bee colony algorithm can be introduced briefly as follows:

Step 1: Initialize the population $x_i, i = (1 \cdots\cdots n)$, calculate the fitness value of the solution to every population and assume that there are n_0 bees and n_1 observants.

Step 2: The leader searches the neighbourhood to generate new solution v_i according to (1), calculate its fitness value and decide to abandon the position of the nectar. If there is nectar, the bee here become the scout and the randomly generated nectar based on (1) replace this nectar.

Step 3: Use (3) to search a new nectar, calculate the fitness of that position, compare the fitness of v_i with that of x_i and consider to preserve x_i or not.

Step 4: Calculate the relating probability p_i according to (2) and sort the positions of the nectar.

Step 5: The follower selects a new nectar position according to p_i, generate new nectar (solution), perform neighbourhood search to generate new solution v_i according to (2) and calculate its fitness value.

Step 6: Select the nectar position from the population according to the sorted fitness, decide whether to generate a new solution x_i according to (3) to replace it and evaluate whether this position is superior to the position selected by the observant. If so, replace the selected position.

Step 7: Memorize the optimal nectar position in the search process, that is to say, preserve the current best solution.

Step 8: Judge whether meet the stopping condition or not. If not, return to *Step 2*, if so, stop the calculation [7].

Genetic-bee colony self-adaptive hybrid algorithm based on information entropy. Information entropy is used to solve the quantization measurement problem of information. After more than half a century's development and evolution, there appears probability measure entropy used in the common system, which is used to describe the overall uncertainty of the system. Information entropy is a measure of the information disorder degree, the greater the entropy, the higher the information disorder degree, the smaller the contribution of its information. On the contrary, the information entropy is smaller, the smaller information disorder degree, the bigger the contribution of its information. The follow of the artificial bee colony algorithm selects the food source at the uncertainty and the information entropy can measure the uncertainty of the occurrence of the event. That is why the proposed algorithm introduces information entropy, uses the value of the information entropy to measure the uncertainty of the selection process, controls the selection probability by controlling the value of the information entropy and realizes the self-adaptive adjustment of the algorithm [8].

The population initialization follows the maximum colony entropy principle. If the colony entropy of the existing individuals is big, the colony entropy is certainly bigger than the set threshold even though the newly-added individuals overlap the existing individuals. At the time, the population initialization strategy based on the entropy has no effect at all and it is still random initialization. The experiment shows that the colony entropy usually increases with the increased individuals and the randomly increased new individuals can make the colony entropy bigger than the set threshold. Assume that there are *NP* individuals in the colony after the initialization and every individual has the *D*-dimensional encoding. According to the definition of information entropy, the colony entropy H is equal to the sum of every encoding entropy $H_j\,(j = 1, 2, \ldots D)$, namely

$$H(x) = E\left[\log\frac{1}{p(x_i)}\right] = -\sum_{i=1}^{n} p(x_i)\log p(x_i);$$

$$H_j = \sum_{i=1}^{M}\sum_{k=i+1}^{M}(-P_{ik}\log P_{ik}).$$

Here, P_{ik} is similarity degree between the jth-dimensional encoding of the ith individual and the jth-dimensional encoding of the kth individual, namely

$$P_{ik} = 1 - \frac{|x_j(i) - x_j(k)|}{B_j - A_j}. \qquad (4)$$

Here, A_j, B_j are the jth-dimensional boundary of the search region D defined by (4).

The initialization process based on the information entropy is as follows: firstly, randomly initialize several individuals as the initialized population. Then randomly initialize a new individual and calculate the information entropy according to the above-said formula. If the colony information entropy is bigger than the set threshold, add the new individual into the existing initialized population until *NP* individuals have been obtained. In this way, the initialized population may be uniformly distributed in the search space to improve the robustness for the algorithm to search the globally optimal solution [9]. The basic framework of the algorithm in this paper is indicated as Fig. 2.

The procedures of the algorithm are as follows:

1. Initialize the population with the information entropy and set the average to initialize the threshold as 0.085.

2. Generate the current self-adaptive mutation operator M_i and conduct mutation operation on the two populations with genetic algorithm and artificial bee colony algorithm.

3. Generate the current self-adaptive crossover operator C_i and perform crossover operation on the two populations.

4. Perform selection operation on the two populations.

5. Calculate the revised fitness value of the new offspring generations of the two populations according to the following formula.

$$F(x) = \frac{\sum_{x' \in X} F(x')P_i}{\sum_{x' \in X_i} F(x')} F(x).$$

Take the former *NP* individuals as the optimal individuals in the next generation of Population 1, in this way, Population 1 always integrates the optimal individuals of the two populations as its next generation, thus increasing the diversity and making the population converge towards a better direction while Population 2 evolves according to its original generation.

Fig. 2. Flowchart of genetic-bee colony self-adaptive hybrid algorithm based on information entropy

Experimental simulation and analysis. In order to verify the effectiveness of the above-mentioned analysis and the performance of the improved algorithm, the comparative experiment is made between GAABC of this paper and the basic GA and ABC. In the simulation experiment, this paper selects four test functions. The Table has listed the definition, value range and theoretical global optimal solution of these functions. Among them, Sphere is a single-modal function, which is mainly used to test the optimization accuracy and the implementation performance of the algorithm while the other three functions are complex non-linear multi-modal functions with numerous global extremums. It is difficult for the common algorithms to obtain the global optimal value; therefore, they can be used to examine the global search performance and the ability to avoid prematurity of the algorithm.

Table

Four benchmark functions

Function	Expression	Value range	Global Optimal Solution
Sphere	$f(x)=\sum_{i=1}^{n} x_i^2$	[-150,150]	0
Quartic	$f(x)=\sum_{i=1}^{n} ix_i^4+\text{random}[0,1)$	[-2,2]	0
Rastrigin	$f(x)=\sum_{i=1}^{n}[x_i^2-10\cos(2\pi x_i)+10]$	[-8,8]	0
Griewank	$f(x)=\frac{1}{4000}\sum_{i=1}^{n} x_i^2-\prod_{i=1}^{n}\cos(\frac{x_i}{\sqrt{i}})+1$	[-500,500]	0

In the experiment, randomly operate the algorithms 20 times on every test functions for the mean value. Fig. 3 is the evolution process curve of the mean fitness value of 30 dimensions of every function. The x-coordinate is the cyclic generation while the y-coordinate is the mean fitness value.

Due to their own characteristics, these test functions can easily be trapped in locally optimum, stagnating the search, however, because GAABC has integrated the supplementary advantages of GA and ABC, it makes dynamic adjustments of the parameters to make them jump out of the local optimum, thus improving the convergence accuracy of the algorithm. For Sphere function, it converges to the optimum in an approximate linearly decreasing manner and in terms of Quartic function; the convergence effect is even more obvious. Therefore, GAABC has excellent global search performance and faster search speed. Since this paper introduces information entropy in the colony algorithm, it realizes the self-adaptive adjustment of the algorithm and it can overcome such defects as prematurity in processing large-scale high-dimensional optimization problem. Although no optimal solution can be obtained, the solution it obtains is very close to the optimal solution. Besides, the algorithm of this paper has also improved its convergence speed.

Discussion and conclusion. Swarm intelligence resolves the practical optimization problem by simulating the activities of the species in the natural world. It treats every individual as an intelligent agent and several agents cooperate with each other to fulfill the search and optimization on the objectives in a highly efficient manner. This paper integrates the advantages of genetic algorithm and artificial bee colony algorithm, proposes a genetic-bee colony dual-population self-adaptive hybrid algorithm based on information entropy, and verifies that this algorithm has excellent effectiveness and robustness. Its specific hybrid strategies can be selected flexibly according to different application fields and problems.

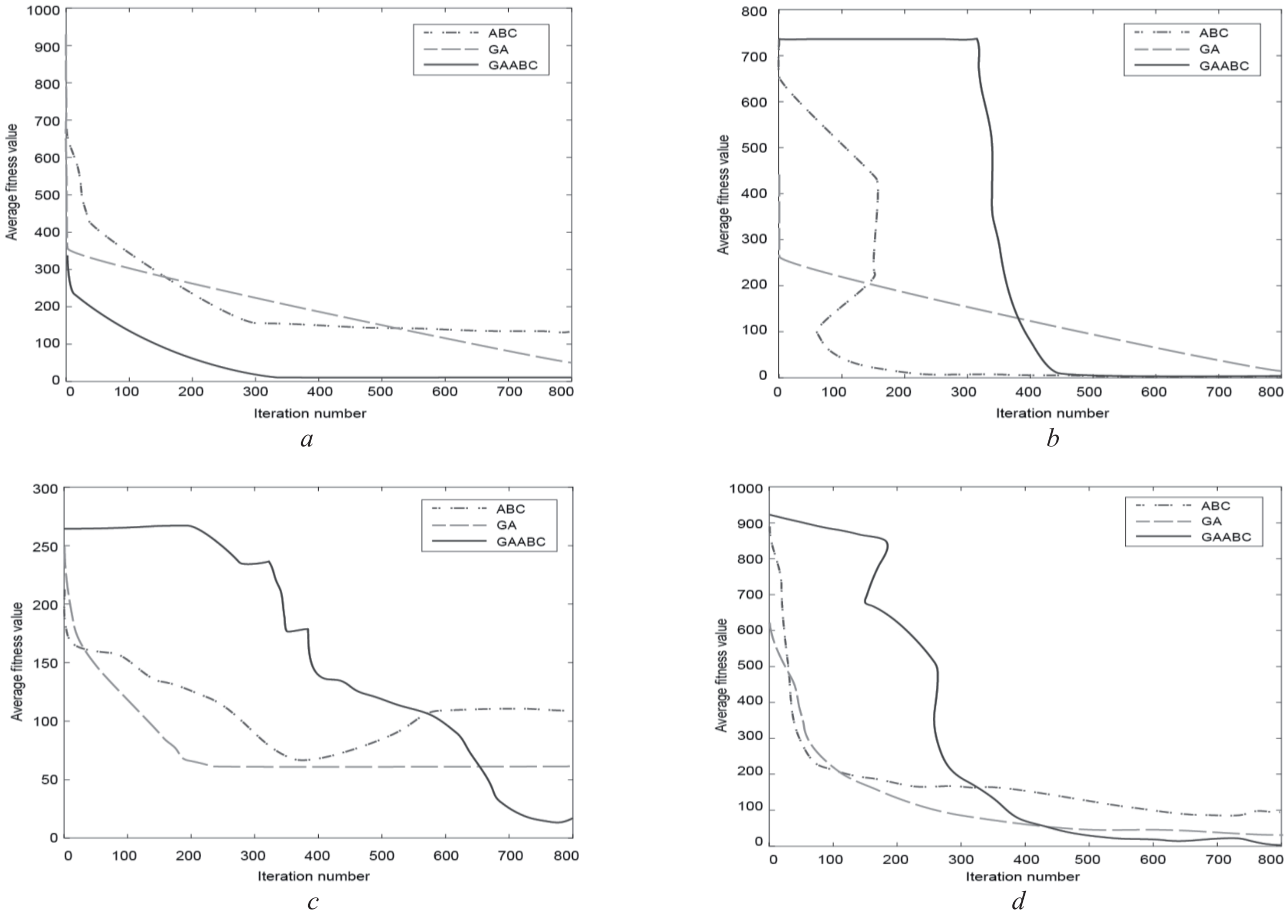

Fig. 3. Comparison of benchmark functions: a – sphere function; b – quartic function; c – rastrigin function; d – griewank function

Acknowledgements. This work was supported by Henan province Natural Science Foundation Research project Funding (Project Number: 142300410435).

References / Список літератури

1. Rahmani, A. and MirHassani, S.A., 2014. A hybrid firefly-genetic algorithm for the capacitated facility location problem. *Information Sciences*, vol. 283, no. 1, pp. 70–78.

2. Saeed Soltanali, Rouein Halladj, Shokoufe Tayyebi and Alimorad Rashidi, 2014. Neural Network and Genetic Algorithm for Modeling and Optimization of Effective Parameters on Synthesized ZSM-5 Particle Size. *Materials Letters*, vol. 136, no. 1, pp. 138–140.

3. Ergun Uzlu, Murat İhsan Kömürcü, Murat Kankal, Tayfun Dede and Hasan Tahsin Öztürk, 2014. Prediction of berm geometry using a set of laboratory tests combined with teaching–learning-based optimization and artificial bee colony algorithms. *Applied Ocean Research*, vol. 48, no. 10, pp. 103–113.

4. Y. Volkan Pehlivanoglu, 2014. Direct and indirect design prediction in genetic algorithm for inverse design problems. *Applied Soft Computing*, vol. 24, no. 11, pp. 781–793.

5. Muhammet Unal, Mustafa Onat, Mustafa Demetgul and Ha-luk Kucuk, 2014. Fault diagnosis of rolling bearings using a genetic algorithm optimized neural network. *Measurement*, vol. 58, no. 12, pp. 187–196.

6. Nafiseh Imanian, Mohammad Ebrahim Shiri, Parham Moradi, 2014. Velocity based artificial bee colony algorithm for high dimensional continuous optimization problems. *Engineering Applications of Artificial Intelligence*, vol. 36, no.11, pp. 148–163.

7. Shuzhu Zhang, C.K.M. Lee, K.L. Choy, William Ho and W.H. Ip. 2014. Design and development of a hybrid artificial bee colony algorithm for the environmental vehicle routing problem. *Transportation Research Part D: Transport and Environment*, vol. 31, no. 8, pp. 85–99.

8. A.J. Umbarkar, M.S. Joshi, and Wei-Chiang Hong, 2014. Multithreaded parallel dual population genetic algorithm (MPDPGA) for unconstrained function optimizations on multi-core system. *Applied Mathematics and Computation*, vol. 243, no. 15, pp. 936–949.

9. Dervis Karaboga and Beyza Gorkemli, 2014. A quick artificial bee colony algorithm and its performance on optimization problems. *Applied Soft Computing*, vol. 23, no. 10, pp. 227–238.

Мета. Ройовий інтелект є розумною поведінкою індивідуумів нерозумного виду або виду, що володіє простим інтелектом, що з'являється завдяки якій-небудь формі об'єднання та взаємодії. Досліджено двопопуляційний самоналагоджувальний гібридний генетико-бджолиний алгоритм, що поєднує в собі генетичний алгоритм (GA) і бджолиний алгоритм (ABC). Отримані оцінки робочих характеристик алгоритмів ройового інтелекту.

Методика. Запропонований двопопуляційний самоналагоджувальний гібридний генетико-бджолиний алгоритм на основі ентропії інформації, що формує дві популяції з незалежною еволюцією, а інформаційна взаємодія між популяціями здійснюється через ентропію інформації, яка дозволяє підтримувати їх різноманітність і, за необхідності, прискорення процесу еволюції між ними.

Результати. Виконано аналіз базових структур і особливостей генетичного та бджолиного алгоритмів, на основі чого запропонована подвійна популяція з еволюцією за допомогою обміну інформацією між ними на основі ентропії. Паралельна дія двох відносно незалежних популяцій прискорює появу нового індивіда у процесі суперництва між популяціями і показує кращий результат у вирішенні завдань оптимізації складних функцій.

Наукова новизна. Проведено комбіноване дослідження генетичного та бджолиного алгоритмів. Визначено, що швидкість перетворення біологічно-інтелектуального еволюційного алгоритму значно покращала, проте, запропонований алгоритм не є ідеальним для завдань оптимізації складних функцій. Цей напрям досліджень вимагає подальшого вивчення.

Практична значимість. Проаналізовано оптимізаційний алгоритм, що застосовується в багатьох областях досліджень. Сучасною тенденцією розвитку є поліпшення вихідного алгоритму шляхом комбінування з інтелектуальним алгоритмом. Оскільки двопопуляційний алгоритм позбавлений недоліків однопопуляційного, він краще підходить для складних оптимізаційних завдань. Закладена основа для пошуку рішень складних розподілених завдань без централізованого контролю та глобальної моделі.

Ключові слова: *завдання оптимізації, генетичний алгоритм, бджолиний алгоритм, подвійна популяція, ентропія інформації, гібридний алгоритм*

Цель. Роевой интеллект представляет собой разумное поведение индивидуумов неразумного вида или вида, обладающего простейшим интеллектом, которое появляется благодаря какой-либо форме объединения и взаимодействия. Исследован двухпопуляционный самонастраивающийся гибридный генетико-пчелиный алгоритм, совмещающий в себе генетический алгоритм (GA) и пчелиный алгоритм (ABC). Получены оценки рабочих характеристик алгоритмов роевого интеллекта.

Методика. Предложен двухпопуляционный самонастраивающийся гибридный генетико-пчелиный алгоритм на основе энтропии информации, который формирует две популяции с независимой эволюцией, а информационное взаимодействие между популяциями осуществляется через энтропию информации, что позволяет поддерживать их разнообразие и, при необходимости, ускорение процесса эволюции между ними.

Результаты. Выполнен анализ базовых структур и особенностей генетического и пчелиного алгоритмов, на основе чего предложена двойная популяция с эволюцией посредством обмена информацией между ними на основе энтропии. Параллельное действие двух относительно независимых популяций ускоряет появление нового индивида в процессе соперничества между популяциями и показывает лучший результат в решении задач оптимизации сложных функций.

Научная новизна. Проведено комбинированное исследование генетического и пчелиного алгоритмов. Определено, что скорость преобразования биологически-интеллектуального эволюционного алгоритма значительно улучшилась, однако, предложенный алгоритм не является идеальным для задач оптимизации сложных функций. Это направление исследований требует дальнейшего изучения.

Практическая значимость. Проанализирован оптимизационный алгоритм, применяемый во многих областях исследования. Современной тенденцией развития является улучшение исходного алгоритма путем комбинирования с интеллектуальным алгоритмом. Поскольку двухпопуляционный алгоритм избавлен от недостатков однопопуляционного, он лучше подходит для сложных оптимизационных задач. Заложена основа для поиска решений сложных распределенных задач без централизованного контроля и глобальной модели.

Ключевые слова: *задача оптимизации, генетический алгоритм, пчелиный алгоритм, двойная популяция, энтропия информации, гибридный алгоритм*

Рекомендовано до публікації докт. техн. наук В.І. Корнієнком. Дата надходження рукопису 03.02.15.