Deconvolution of complex spectra into components by the bee swarm algorithm

R R Yagfarov[1], M E Sibgatullin[2], D Z Galimullin[3], D I Kamalova[2] and M Kh Salakhov[2,4]

[1]Innopolis, Innopolis, University St. 1, 420500
[2]Kazan, Kazan Federal University, Kremlyovskaya St. 18, 420008
[3]Kazan, Institute of Economics, Management and Law, Moscow St. 42, 420008
[4]Kazan, Academy of Sciences of the Republic of Tatarstan, Bauman St. 20, 420008

E-mail: goodraff@gmail.com

Abstract. The bee swarm algorithm is adapted for the solution of the problem of deconvolution of complex spectral contours into components. Comparison of biological concepts relating to the behaviour of bees in a colony and mathematical concepts relating to the quality of the obtained solutions is carried out (mean square error, random solutions in the each iteration). Model experiments, which have been realized on the example of a signal representing a sum of three Lorentz contours of various intensity and half-width, confirm the efficiency of the offered approach.

1. Introduction

Complex contours with unresolved structure are often registered when carrying out spectroscopic research. The examples of these spectra are IR absorption spectra of polymers [1]. To obtain complete and reliable information about the structure and properties of a substance, methods of deconvolution of spectral contours into elementary components are necessary. The spectra belonging to a particular chemical compound or spectral contours arising from quantum transitions and broadened by different mechanisms are usually called elementary components.

The aim of this work is implementation and adaptation of the method of deconvolution of molecular FTIR spectra into components using the artificial bee colony algorithm (ABC). This algorithm is a stochastic method for global optimization of a residual function based on the patterns of the behavior of bees [2]. Solutions with the least mean square error are taken as "foragers", whereas randomly generated solutions at the each iteration play a role of “scout bees”. "Flying out from the hive to harvest nectar, return to the hive and data processing" occur in one iteration. In order to assess the effectiveness of the bee algorithm for solving the problem of deconvolution, modeling experiments were carried out.

2. Swarm intelligence algorithm for decomposition of spectral contours

In this work the method of a swarm of bees was adapted to the solution of the problem of deconvolution of spectral contours into components. This algorithm simulates the behavior of bees in the natural environment. The idea of the bee algorithm is that in search of nectar all bees inspect at each step both high profitable nectar sources and the nectar sources in the vicinity of the high profitable ones. This allows, firstly, varying the population of solutions in subsequent iterations and, secondly, increasing the probability of finding solutions close to the optimum [3]. At first, several scout bees fly out from the hive in random directions to find nectar sources. After some time, the bees return to the hive, and in a special way inform the whole hive, where and how much nectar they have found. After that, another type of bees (foragers) fly to the found sites. The more nectar is expected to be found in the given site, the more bees fly in this direction, while scouts fly out again to look for other sites. After that the process is repeated. The position of global extremum corresponds to the site with the greatest amount of nectar, this section being the only one. This means that other sites contain a smaller amount of nectar.

The basic concepts of the algorithm are nectar source (flower, site), foragers (employed bees), scout bees. The source of nectar is characterized by its profitability determined by various parameters. Foragers are assigned to the sources of nectar. The number of bees on these sites is greater than on the other ones. The average number of scouts amounts to 5 - 10% of the swarm. After returning to the hive, the bees exchange information on the so-called closed dance area [4].

Let us consider in more detail the algorithm applied to the problem of deconvolution of spectral contours. Table 1 shows the comparison of biological and mathematical notions.

Table 1. Matching of classical and mathematical notions.

Classical term	**Mathematical notion**
Nectar source (flower)	MSE (mean square error) $\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$
The most attractive sites	Solutions with the least MSE
Promising region	Solutions with a greater MSE than in attractive areas, but still quite promising
Bee	Solution as a sum of contours. In the case of a two-component Lorentz contour $P(\nu) = \frac{A_1}{\left(1 + \frac{(\nu - s_1)^2}{{w_1}^2}\right)} + \frac{A_2}{\left(1 + \frac{(\nu - s_2)^2}{{w_2}^2}\right)}$
Foragers	A set of solutions which calculates MSE on the promising and attractive areas
Scouts	Random solutions at each iteration
Flying out, harvesting nectar, returning to the hive	One iteration

The mean square error acts as a source of nectar. Search of the parameters of contour components is carried out in the space of dimension $N = n * m$, where n is the number of contour components, m is the number of unknown parameters of a component. Bees are the solutions obtained by adding up the estimated elementary components of a complex contour with certain parameters. Promising and most attractive sites are selected after ranking by the MSE value. Scouts are random solutions at each iteration of the algorithm. Foragers are the solutions calculated in the region of promising and attractive sites.

Figure 1 shows the block diagram of the algorithm. At the first stage a certain number of random initial sets of parameters of input contour components are generated. The number can be varied. Next, MSE is calculated, and ranking by the MSE value is performed. The n solutions with the least objective function (MSE) are selected. So called promising sites m are selected, where the objective function (MSE) is greater than in the best sites, but these sections can also be used for the search. After selecting the best and promising solutions at the given iteration, MSE in the domain of these solutions is calculated. The smaller is the MSE value, the more calculations are carried out in the domain of the given solution. Thus, the foragers are sent not exactly to the points where the best and the promising solutions have been found, but to the vicinity of these points. Then, ranking by MSE value is repeated, but now it is carried out not only among the scouts (random solutions), but also among all the bees (foragers and scouts). The algorithm is repeated until one of the stopping criteria is reached. Moreover, the parameters of the search area are changed at each iteration. The number of iterations or achievement of the deconvolution permissible error can be chosen as stopping criteria.

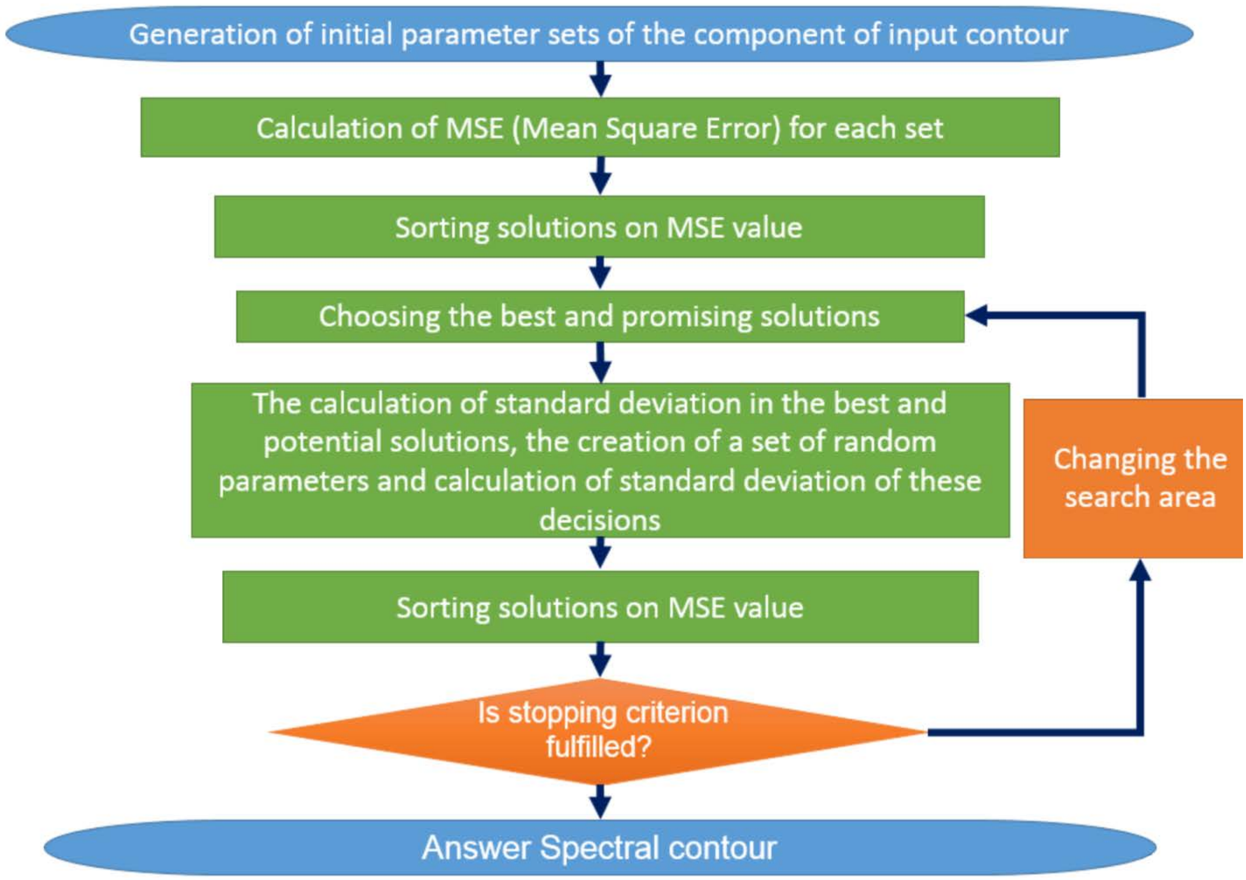

Figure 1. Block-diagram of the bee algorithm adapted to the problem.

In order to study the possibilities of applying the bee algorithm to the problem of deconvolution of complex contours, model experiments were carried out. The result of deconvolution of a three-component model using ABC is depicted in Figure 2. The results of deconvolution of the three-component model contour obtained by adding up three Lorentz contours are given in Table 2. These contours are characterized by the following parameters: $\boldsymbol{S_1 = 2200}$, $\boldsymbol{A_1 = 0.7}$, $\boldsymbol{W_1 = 77}$ (first contour); $\boldsymbol{S_2 = 2250}$, $\boldsymbol{A_2 = 2}$, $\boldsymbol{W_2 = 35}$ (second contour); $\boldsymbol{S_3 = 2350}$, $\boldsymbol{A_3 = 0.5}$, $\boldsymbol{W_3 = 80}$ (third contour), where $\boldsymbol{S}$ is the position of the center, $\boldsymbol{A}$ is the intensity, $\boldsymbol{W}$ is the half-width. Deconvolution of a complex contour into components was performed using the ABC method and the least square method (OLS). The position of the component centers was taken as a priori information.

Table 2. Numerical values

Method	A_1	A_2	A_3	W_1	W_2	W_3
OLS	-0,02	2,34	0,20	20681	57,93	215,39
ABC	0,69	1,99	0,50	76,99	35,07	80,22

As can see from figure 3, the use of OLS may lead to false results. The model contour is poorly resolved and represents the sum of three Lorentz contours with different intensities and half-widths. Each algorithm was run 2000 times. The number of false solutions for given model spectrum is amounted to 1% of all solutions in the case of ABC algorithm and to 35% in the case of OLS. Although the contour reconstituted by OLS is similar in the form to the original model contour, its components do not coincide with those of the original model contour. Therefore, in the case of multi-component spectral contours it is preferable to use the method of bee colonies adapted to solve the problem of deconvolution into components.

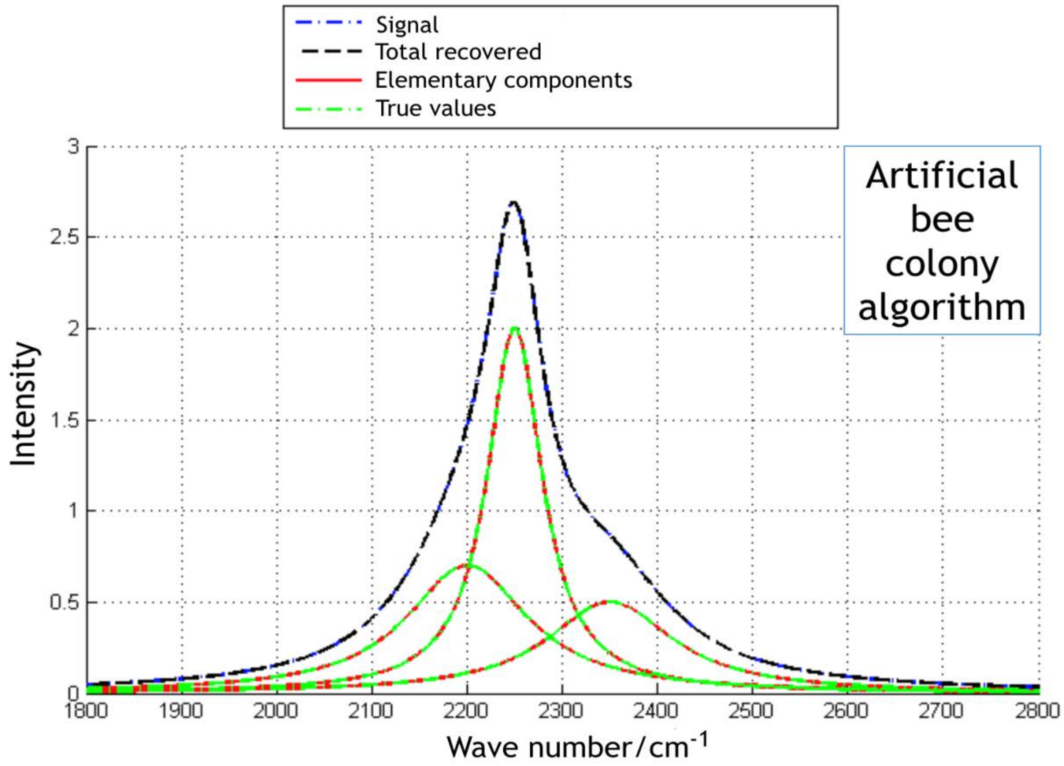

Figure 2. Deconvolution of the three-component model contour using ABC.

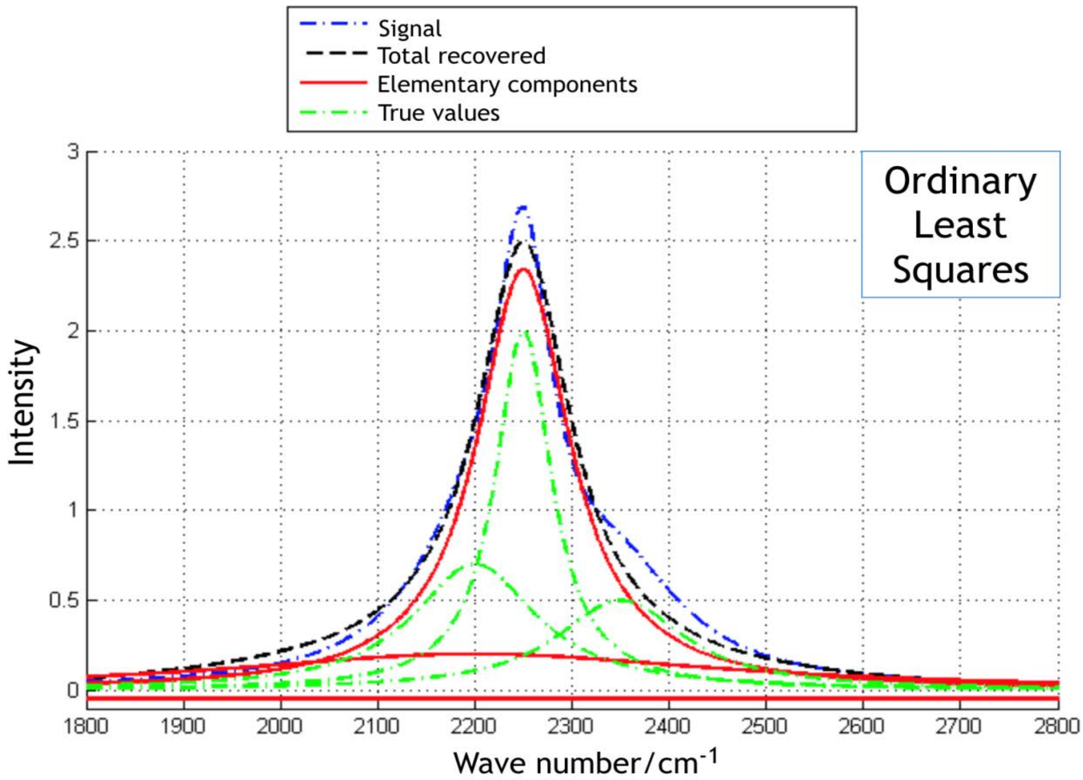

Figure 3. Deconvolution of the three-component model contour using OLS.

3. Conclusion

The artificial bee colony algorithm was implemented and adapted for processing model FT-IR absorption spectra. Model experiments on deconvolution of three-component spectral contours were performed. The possibility of using the artificial bee colony algorithm for solving the problem of deconvolution of poorly resolved spectral contours was shown. Also, comparison of the effectiveness of the proposed algorithm and least squares method was carried out. It is shown that in the case of a complex contour (in particular, of a three-component contour) using the least squares method increases the probability of obtaining false results when compared with the swarm intelligence algorithms.

References

[1] Kamalova D I, Remizov A B and Salakhov M Kh 2008 *Conformation probes in the study of local mobility of polymers* (Moscow: Fizmatkniga) (in Russian)
[2] Pham D T *et al* 2011 The Bees Algorithm–A Novel Tool for Complex Optimisation *Intelligent Production Machines and Systems - 2nd I* PROMS Virtual International Conference (3-14 July 2006)* ed D T Pham, E E Eldukhri and A J Soroka (Amsterdam: Elsevier) p 454
[3] Karaboga D 2005 An idea based on honey bee swarm for numerical optimization *Technical Report-tr06* (Erciyes University, Engineering faculty, Computer engineering department)
[4] Kureichik V M and Kazharov A A 2010 Graph coloring solution based on bee colony optimization *Izvestia SFedU. Engineering sciences* **113** (12) pp 30-36