The Development of the Solution Search Method Based on the Improved Bee Colony Algorithm

Andrii Shyshatskyi¹, Tetiana Stasiuk², Oleg Kuzmenko²

¹ Educational and Scientific Institute of Public Administration and Civil Service, Taras Shevchenko Kyiv National University, Kyiv, Ukraine
² Military College of the Sergeant Staff Military Institute of Telecommunications and Informatization named after the Heroes Krut., Poltava, Ukraine

Abstract

Active digitization of people's daily life leads to the use of the decision making support systems (DMSS). DMSS is actively used in data processing, forecasting the course of various processes, providing informational support for the decision making process by decision makers. However, a number of problems arise while evaluating monitoring objects, namely: a large number of destabilizing factors affecting the efficiency of the processes of information collection, processing and transmission; high dynamism of changes in the state and composition of heterogeneous monitoring objects during the conduct of hostilities (operations); high dynamism of conducting hostilities (operations); the uncertainty of the initial situation and the noise of the initial data. In this article, a method of finding solutions based on an improved bee colony algorithm was developed.

The efficiency of information processing is achieved by learning the architecture of artificial neural networks; taking into account the type of uncertainty of the information to be evaluated; the use of an improved algorithm of the bee colony, the use of an unordered linguistic scale of measurements with adjustment coefficients for the degree of awareness and the degree of noise of the initial data. An approbation of the use of the proposed method was carried out on the example of assessing the state of the operational grouping of troops (forces). The method is proposed to be used in the development of software for automated systems of control of troops and weapons, namely, in the modernization of existing and development of new automated systems of control of troops and weapons. The evaluation of the effectiveness of the proposed method showed an increase in the efficiency of the evaluation at the level of 21–28% in terms of the efficiency of information processing.

Keywords: bee algorithm, heterogeneous intelligence objects, intelligent systems, decision making support systems

1. Introduction

The active introduction of information systems into people's lives becomes a catalyst for the development of decision making support systems (DMSS). Intelligent DMSS are a natural development of classical DMSS. Intelligent DMSS are also used to solve specialized tasks [1–4]:

- an analysis of heterogeneous monitoring objects;
- collection, processing and generalization of information from various sources of information;
- collection, processing and generalization of data, etc.

The work [1] presents the method of routing in special purpose networks. The method basically uses the ant colony algorithm. The disadvantage of the specified method is the insufficient speed of decision making, failure to take into account the type of uncertainty and noisy input data.

The disadvantage of the specified method is the insufficient speed of decision making due to unidirectional search, failure to take into account the type of uncertainty and noisy input data.

The work [3] assessed the stability of the communication system under the influence of destabilizing factors. The list of them is determined, the need to increase the effectiveness of the assessment under the influence of their multitude is determined.

The work [4] proposed to use a generalized decision making metric while working with multidimensional data. The disadvantage of the mentioned approach is insufficient decision making speed, failure to take into account the type of uncertainty and noisy input data.

The works [5, 6] present the method of signal processing in unmanned aerial vehicles. The method is based on convolutional artificial neural networks to obtain a generalized estimate of the signal from a set of parameters. The disadvantage of this method is learning by the method of backpropagation of the error, which leads to the accumulation of the estimation error.

The work [7] presents an approach to assess the impact of radio-electronic warfare on MIMO systems. The approach is based on convolutional artificial neural networks to obtain a generalized signal estimate from a set of parameters. The disadvantage of this approach is learning by the method of backpropagation of the error, which leads to the accumulation of assessment errors and the subsequent development of management decisions.

The work [8] presents an approach to compensation of phase noise using convolutional artificial neural networks to obtain a generalized estimate of phase noise from a set of parameters. The disadvantage of this approach is that it does not take into account the degree of uncertainty and noise of the initial data.

In work [9] presents an approach to solve tasks for assessing the state of the object based on production-logical expressions. The disadvantage of this approach is that it does not take into account the degree of uncertainty and noise of the initial data.

The work [10] proposed a structural-semantic model for assessing the state of the channels of the MIMO system. In its basis, the proposed model uses the apparatus of fuzzy logic. The disadvantage of this approach is that it does not take into account the degree of uncertainty and noise of the initial data.

The work [11] indicated that the most popular evolutionary bio-inspired algorithms are the so-called "swarm" procedures (Particle Swarm Optimization - PSO). Most of the above-mentioned classes of algorithms can be attributed to the so-called "bio-inspired algorithms". The obvious advantages of bio-inspired algorithms, in addition to the ability to solve large-scale problems, can also be highlighted by taking into account the experience of previous agents in solving the problem. The behavior of real animals and plants was used to develop bio-inspired algorithms. In particular, currently known bio-inspired algorithms are the Ant Colony Optimization (ACO) algorithm, the Artificial Bee Colony (ABC) algorithm, the Gray Wolf Optimization (GWO) algorithm, the Cat Swarm Optimization (CSO) algorithm, a method of imitating the behavior of frogs, a method of imitating the behavior of cuckoos, a method of imitating the behavior of fireflies and a method of imitating the spread of weeds.

The analysis of works [1–11] showed that while analyzing heterogeneous monitoring objects, a number of problematic issues arise, namely:

1. A large number of destabilizing factors at all stages of collection, processing and transmission of information from various sources.

2. Rapidity of hostilities and armed conflicts.

3. The uncertainty of the initial situation and the noise of the initial data.

It determines further development in terms of the search for new scientific approaches and technological solutions that will allow [7–11]:

- to provide an objective and operational analysis of the state of heterogeneous monitoring objects;
- to search for a solution in several directions at the same time;
- to take into account the noise of the initial data and the uncertainty of the initial situation.

The need for prompt decision making regarding the state of the object of analysis under the influence of destabilizing factors with various types of input data determines the use of artificial intelligence methods. Computational intelligence methods have become widespread for solving a variety of complex tasks, both purely scientific and in the field of technology, business, finance, medical and technical diagnostics and other fields. These include intelligent data analysis (Data Mining), dynamic data analysis (Dynamic Data Mining), analysis of data streams (Data Stream Mining), analysis of large data sets (Big
Data Mining), Web-Mining, Text Mining [1, 2].

Artificial bee colony algorithms (Simulated Bee Colony, SBC) simulate the behavior of honey bees and are used to find solutions to difficult or unsolvable combinatorial problems. SBC algorithms are classified as meta-heuristic algorithms because they provide a universal infrastructure and a set of rules for developing a solution to a task, rather than a detailed solution recipe. It usually includes the initial reconnaissance and subsequent work of the bees of the hive. At the end of its work, the best one is selected from the specified set of solutions, which is the result of the algorithm’s work.

The purpose of the research is to develop a method of finding solutions based on an improved bee colony algorithm. This will allow to increase the efficiency of the assessment with the given reliability and the development of subsequent management decisions. This will make it possible to develop software for intelligent decision making support systems in the interests of the combat management of the actions of troops (forces).

For this purpose, the authors set themselves the following tasks:

- to perform a mathematical formulation of the task of analyzing a heterogeneous monitoring object;
- to develop a method for finding solutions based on an improved bee colony algorithm;
- to evaluate the effectiveness of the proposed method.

### 2. Materials and methods

Evolutionary algorithms and evolving artificial neural networks were used in the research to solve the problem of analyzing the state of the monitoring object and training databases (knowledge bases). Also, in this research, the methods of multidimensional description of the radio-electronic environment, which were developed by the authors earlier, were used. The simulation was carried out using MathCad 2014 software and an Intel Core i3 personal computer.

### 3. Results and discussion

#### 3.1 The statement of the task of research on the development of the method of finding solutions

Let a set of elements be given

\[ A = \{a_j \mid j = 1, 2, \ldots, n\} \]

and multiple positions

\[ \Pi = \{\pi_i \mid i = 1, 2, \ldots, c\} \]

areas of search for solutions. The hypergraph \( H = (X, E) \) is used as a scheme model, where \( X = \{x_i \mid i = 1, 2, \ldots, n\} \) are the sets of vertices modeling elements, \( E = \{e_j \mid e_j \subseteq X, j = 1, 2, \ldots, m\} \) are the sets of hyperedges modeling chains connecting elements. To place all elements, the condition must be met \( c \geq n \). Some \( s \)-th solution to the problem of placing elements in positions is a permutation \( V_s = \{v_{s1}, \ldots, v_{sc}\} \), where \( v_{si} \) specifies the element number that is assigned to position \( \pi_i \). Depending on the selected criterion, the objective function \( F(V_s) \) is introduced to evaluate the placement results. The task is reduced to the search for each element of such positions in the area of the search for solutions, in which the selected quality indicator is optimized.

As an estimate \( l_j \) of the length of the chain \( t_j \) modeled by the hyperedge \( e_j \), the following are used: the length of the minimum spanning tree built on the set of vertices \( e_j \subseteq X \); the length of the graph whose edges are incident to the vertices \( e_j \subseteq X \) and the root vertex is placed in the center of "weight" of the set of vertices \( e_j \); the length of the semi-perimeter of the rectangle describing the set of vertices \( e_j \); the total length of the edges of the complete graph built on the set \( e_j \).

With this in mind, the optimization criterion looks like this:

\[
F = \sum_{j=1}^{m} l_j, \quad (1)
\]

The placement task consists in finding the optimal value of the function \( F \) on a set of permutations \( V \). To fully account for the relationships between the placement and tracing tasks, a criterion based on estimates of the
number of chains crossing the given lines of the solution search section is used. These lines can be either straight, crossing all sections of the search for solutions or closed and limiting a certain section [1]. In the research, the main focus is on the construction of the placement search procedure that optimizes the selected quality indicator. Let’s represent the set of positions of the search area in the form of an orthogonal grid. Positions are numbered. Let \( \pi_i \) be the position of the search area. Each \( \pi_i \) has its own coordinates – row number and column number of the grid. Information about position coordinates is stored in the form of a two-dimensional matrix \( K = k_{ij} \in \mathbb{R} \times 2 \), where \( k_{i1} \) and \( k_{i2} \) are the position coordinates \( \pi_i \) in the grid, \( k_{i1} \) is the row number, \( k_{i2} \) is the column number. Position \( \pi_i = (k_{i1}, k_{i2}) \) located at the intersection of row \( k_{i1} \) and column \( k_{i2} \). In the future, we will consider the pair of coordinates \( (k_{i1}, k_{i2}) \) as equivalent to the position \( \pi_i \). Thus, \( \pi_i = (k_{i1}, k_{i2}) \). Some \( s \)-th solution to the task of placing elements in the positions of the search area is a permutation \( V_s = \{v_{s1}, \ldots, v_{si}, \ldots, v_{sm}\} \), where \( v_{si} \) specifies the element number that is assigned to position \( \pi_i \).

Now we consider the solution space. In contrast to the standard paradigm, instead of a metric (numerical) scale, the work uses an unordered linguistic scale with correction coefficients proposed in the work [2]. The solution space contains one axis. In our case, the position of the decision space \( x_i \) corresponds to the solution given by the permutation \( V_i \). The reference points on the axis are the values of the linguistic variable \( x_i \) disordered relative to each other. By value \( x_i \), there is a combination of the relative location of the elements in the vector \( V_i \).

We note that the reference points \( x_i \) on the axis are determined after the agents find a swarm of solutions \( V_i \) and are postponed in the order of finding these decisions.

The key operation of the algorithm is the research of promising positions and their neighborhoods in the solution space. Let’s call it the \( \delta \)-neighborhood of the position \( \pi_i \) multiple positions \( \Pi_i \) such that the coordinates of any of the positions of the set \( \Pi_i \) differ from the coordinates of the position \( \pi_i \) by no more than \( \delta \).

Let \( \delta = 1 \), then the \( \delta \)-neighborhood of the position \( \pi_i = (k_{i1}, k_{i2}) \) is a set of positions (2):

\[
\Pi_i = \{(k_{i1}, k_{i2} + 1), (k_{i1}, k_{i2} - 1), (k_{i1} + 1, k_{i2}), (k_{i1} - 1, k_{i2})\}.
\]

Let be the set of elements \( A = \{a_j \mid j = 1, 2, \ldots, n\} \) placed in the search area pursuant to the decision of \( V_o \). After any mutual permutation of a pair of elements \( (a_{i1}, a_{i2}) \), which is located in the position \( \pi_i \) and the other in one of the positions \( \pi_i \) belonging to the neighborhood of \( \Pi_i \), a new solution will be obtained. Let the elements \( (v_{si}, v_{si}) \) of the vector \( V_i \) correspond to a pair of elements \( (a_{i1}, a_{i2}) \), which are placed in adjacent positions \( (\pi_i, \pi_i) \). The permutation in the solution section of a pair of elements \( (a_{i1}, a_{i2}) \) corresponds to the permutation of a pair of elements \( (v_{si}, v_{si}) \) in the vector \( V_i \). As a result, the vector \( V_o \) is formed. We will assume that the solution of the vector \( V_o \) obtained in this way lies in the neighborhood of the solution \( V_s \). In the case of the neighborhood of solutions \( V_i \) are solutions obtained by pairwise permutation of elements located in two adjacent positions of the search area. We note that the permutation of a pair of adjacent elements in the vector \( V_i \) does not always result in a solution \( V_o \) lying on the edge of the solution \( V_s \).

3.2. The development of a method for finding solutions based on an improved bee colony algorithm.

The method of finding solutions based on the improved bee colony algorithm is described below. The difference between the proposed method and the known ones is the use of an unordered linguistic scale of measurements with correction factors for the degree of awareness and the degree of noise of the initial data, adaptive management of the parameters of the placement procedure, the ratio between the number of forager bees and scout bees.

The following procedures are added to the classic algorithm:
• the initial display of bees, taking into account the type of uncertainty and noise of the initial data;
• training of bees using the method of learning artificial neural networks;
• taking into account the degree of uncertainty of the initial data and their noise.

The method for finding solutions based on the improved bee colony algorithm has the following sequence:

Step 1. The parameters of the bee colony algorithm are determined:

$L$ is the number of iterations of the algorithm; $n_s$ is the number of scout bees at the beginning of the algorithm; $n_b$ is the set of basic positions of the bee algorithm; $\delta$ is the limit value of the search area; $n_f$ is the basic number of forager bees; $n_{b1}$ is the number of base positions formed from the best positions $x^*_b(l)$, found by the swarm on the $l$-th iteration; $n_{i1}$ is the number of scout bees that choose new positions, taking into account the correction coefficients of the degree of data noise and the degree of awareness; $n_{i2}$ is the number of initial positions that are formed from new positions found by bees that perform reconnaissance on the $l$-th iteration.

Step 2. Generation of a set of solutions that differ from each other $V(1)=\{V_s(1)|s=1,2,\ldots,n_s\}$, to which a set of positions corresponds $X(1)=\{x_s(1)|s=1,2,\ldots,n_s\}$ taking into account correction coefficients for the degree of awareness and degree of data noise [2].

Step 3. To every decision $V_s(1)$, the value of the objective function $F_s(1)$ is calculated.

Step 4. A set of basic solutions is formed $V^b(1)\subseteq V(1)$ with the best values of the objective functions $F_s(1)$ and the corresponding set of basic positions $X^b(1)\subseteq X(1)$. $|V^b(1)|=|X^b(1)|=n_b$, $l=1$ ($l$ is the iteration of the algorithm), $z=1$. ($z$ is the forager bee number).

Step 5. The choice with probability $p(x^b(l))=F^b_f(l)/\sum_s(F^b_s(l))$ of basic position $x^b(l)\in X^b(1)$.

Step 6. Selection of position $x_s(l)$, which is located on the edge of the base position $x^b(l)$, with the solution $V_s(l)$.

Step 7. If the $x_s(l)$ position coincides with the previously selected positions, proceed to step 6, otherwise proceed to step 8.

Step 8. Position $x_s(l)$ is included in the plural $O_s(l)$.

Step 9. Calculation of the value of the objective function $F_z(l)$ of decision $V_z(l)$.

Step 10. If $z<n_f$, then $z=z+1$ and go to step 6, otherwise go to step 11.

Step 11. Formation of each basic position $x^b_s(l)$ of the search area $D_s(l)\in O_s(l)\cup x^b_s(l)$.

Step 12. In each region $D_s(l)$, the best position is chosen $x^*_s(l)$ with the best solution $V^*_s(l)$.

Step 13. Among $V^*_s(l)$, the best solution $V^*(l)$ is chosen.

Step 14. Bees are trained based on the method proposed in work [2]. If $V(l)$ is better than $V^*(l-1)$, it is kept, otherwise $V^*(l-1)$.

Step 15. If $l\leq L$, then $l=l+1$ and go to step 16, otherwise go to step 20.

Step 16. In the first part $X^{b1}(l)$ are included $n_{b1}$ best positions, among positions $x^*_b(l-1)$, found by bees in each of the $D_s(l-1)$ plots formed in the previous iteration.

Step 17. Randomly generating a set of solutions that differ from each other $V(l)\{V_s(l)|s=1,2,\ldots,n_{r1}\}$, to which a set of positions $X(l)\{x_s(l)|s=1,2,\ldots,n_{r1}\}$. $|V(l)|=n_{r1}$ corresponds.

Step 18. Inclusion in the plural $X^{b2}(l)n_{b2}$ obtained from the set $X(l)$ of new positions found by intelligence agents on the $l$-th iteration $n_{b1}+n_{b2}=n_b$.

Step 19. Formation of a set of basic positions. $X^b(l)=X^{b1}(l)\cup X^{b2}(l)$. Go to step 6.

Step 20. The end of the algorithm.

As the object of analysis, the operational grouping of troops (forces) according to the wartime staff was chosen. Initial data for
modeling and evaluating the effectiveness of the proposed method:

- the number of sources of information about the state of the monitoring object is 3 (radio monitoring tools, remote earth sensing tools and unmanned aerial vehicles). To simplify the modeling, the same number of each tool was taken – 4 tools each;
- the number of informational signs by which the state of the monitoring object is determined is 12. These parameters include: ownership, type of organizational and staff formation, priority, minimum width along the front, maximum width along the front. The number of personnel, minimum depth along the flank, maximum depth along the flank, the number of samples of weapons and military equipment (WME), the number of types of WME samples and the number of communication devices), the type of operational construction are also taken into account;
- the variants of organizational and personnel formations are company, battalion, brigade.

The total number of parameters by which the state of the operational grouping of troops (forces) is assessed is 1,574 parameters.

The results of the evaluation are given in the table 1. As it can be seen from the table 1, the gain of the mentioned method of searching for solutions is from 11 to 15% according to the criterion of efficiency of data processing.

To determine the effectiveness of the proposed method, researches were conducted in comparison with other swarm methods, namely the ant algorithm (AA), the particle swarm optimization (PSO) method and the classic bee swarm algorithm. The simulation results are presented in table 1.

From the table 1, it can be concluded that the proposed method has an acceptable computational complexity.

Table 2 provides an assessment of the efficiency of training of evolving artificial neural networks used in the 14-th step of the bee swarm algorithm.
The training procedure provides an average of 10–18% higher training efficiency of artificial neural networks and does not accumulate errors during training (table 2).

Based on the results of the comparative evaluation (tables 1, 2), it can be concluded that the total gain of the improved algorithm is from 21 to 28% due to the use of improved procedures.

The limitations of the research are the need for sufficient computing power, experienced service personnel and taking into account the time required for the entire cycle of information exchange.

The advantages of the proposed method are due to the following:

- during the initial display, the type of uncertainty and the degree of noise of the initial data are taken into account;
- universality of solving the task of analyzing the state of monitoring objects due to the hierarchical description of them;
- the possibility of rapid construction of models due to the simultaneous search for a solution by several individuals;
- an adequacy of the obtained results;
- the ability to avoid the problem of a local extremum;
- the possibility of in-depth learning of knowledge bases.

The disadvantages of the method are:

- reduced assessment accuracy while assessing one assessment indicator;
- lower assessment accuracy compared to other assessment methods designed for assessment by a separate indicator.

This method will allow:

- to evaluate a complex and heterogeneous object;
- to increase the efficiency of evaluation;
- to reduce the use of computing resources.

The proposed approach is expedient to use in information and automated systems for the management of troops and weapons, such as "Logistica-IT", "Oreanda-PS", "Dzvin-AS", namely:

- a comprehensive assessment of the air situation for further decision making on target engagement, planning and operational management of air defense of various basing options;
- an assessment of the level of provision of material and technical devices of groups of troops (forces) and individual units;
- an assessment of the operational (strategic) situation in the defined region of responsibility and planning of the use of troops (forces).

The directions of further research will be directed to the development of methods for increasing the efficiency of decision making using the bat algorithm.

4. Conclusions

1. A mathematical formulation of the research task was carried out. In contrast to the standard paradigm of the bee colony algorithm, instead of a metric (numerical) scale, an unordered linguistic scale of measurements with adjustment coefficients for the degree of awareness and the degree of noise of the output data is used. Improving the efficiency of the algorithm is achieved by adaptively controlling the parameters of the placement procedure, such as the size of the δ-neighborhood of the position \( \pi_i \), the ratio between the number of forager bees and scout bees, etc.

2. In the course of the research, a method for finding solutions was developed based on an improved bee colony algorithm, which allows:

- to conduct an initial display of bees taking into account the type of uncertainty and noise of the initial data;
- to conduct training of bees using the method of training artificial neural networks;
- during the generation of decisions about the condition of the monitoring object, the degree of uncertainty of the initial data and their noise are also taken into account.

3. Modeling of the proposed method was carried out. The simulation results showed an increase in the efficiency of data processing at the level of 21–28% due to the use of additional improved procedures.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this paper.
Financing
The study was performed without financial support.

Data availability
Manuscript has associated data in a data repository.

References


