The problems of processing heterogeneous data are discontinuous, undifferentiated, and multimodal. The most common approaches to processing heterogeneous data are swarm intelligence algorithms (swarm algorithms). Given the above, classical gradient deterministic algorithms are inappropriate for solving the problems of processing heterogeneous data. The problem solved in the study is to increase the efficiency of processing heterogeneous data circulating in information systems, regardless of the number of data sources. The object of the study is hierarchical systems. A method for increasing the efficiency of processing heterogeneous data using a metaheuristic algorithm is proposed. The study is based on the reptile algorithm (RA) for processing heterogeneous data circulating in the system. For RA training, evolving artificial neural networks are used.

The originality of the proposed method lies in setting RA taking into account the uncertainty of the initial data, improved global and local search procedures. Also, the originality of the study lies in determining RA feeding locations, which allows prioritizing the search in a given direction. The next element in the originality of the study is the possibility of choosing an RA hunting strategy, which allows a rational use of available system computing resources. Another original element of the study is determining the initial velocity of each RA. This makes it possible to optimize the speed of exploration of each RA in a certain direction. The method provides a 15–19 % increase in data processing efficiency by using additional improved procedures. The proposed method should be used in processing large amounts of data

Keywords: unstructured data, artificial neural networks, swarm algorithms, unimodal and multimodal functions

DEVELOPMENT OF A METHOD FOR INCREASING THE EFFICIENCY OF PROCESSING HETEROGENEOUS DATA USING A METAHEURISTIC ALGORITHM

Vitaliy Ragulin
PhD, Associate Professor*
Salman Rasheed Owaid
PhD, Assosiate Professor, Lecturer
Department of Computer Engineering
Al-Taff University College
Karrada str., 3, Karbala, Iraq, 31001
Heorhii Kuchuk
Doctor of Technical Sciences, Professor
Department of Computer Engineering and Programming
National Technical University «Kharkiv Polytechnic Institute»
Kyrpychova str., 2, Kharkiv, Ukraine, 61002
Serhii Andriienko
Lecturer*
Oleksandr Lytvynenko
Corresponding author
PhD, Senior Researcher
Research Department
Research Center
Military Institute of Taras Shevchenko National University of Kyiv
Yuliyi Zdanovskoi str., 81, Kyiv, Ukraine, 03680
E-mail: s63010566s@gmail.com
Evgen Ivanov
PhD, Associate Professor*
Anna Lyashenko
Senior Researcher
Scientific Center**
Alexander Momit
Deputy Head of Research Department
Research Department of the Development
of Anti-aircraft Missile Systems and Complexes
Central Scientifically-Research Institute of Armaments and
Military Equipment of the Armed Forces of Ukraine
Povitrofloski ave., 28, Kyiv, Ukraine, 03049
Oleksandr Gaman
Adjunct
Scientific and Organizational Department**
Taras Hurskyi
PhD, Associate Professor
Head of Scientific-Research Department
Scientific-Research Institute of Military Intelligence
Yuriia Illienka str., 81, Kyiv, Ukraine, 04050
*Department of Computer Graphics
Kharkiv National Automobile and Highway University
Yaroslava Mudroho str., 25, Kharkiv, Ukraine, 61002
**Military Institute of Telecommunications and
Information Technologies named after Heroes of Kruty
Knyaziv Ostroz'kykh str., 45/1, Kyiv, Ukraine, 01011

1. Introduction

Processing heterogeneous data is a complex process of defining a set of solutions for a wide range of problems, including management decision-making [1–3].

The problems of processing heterogeneous data are discontinuous, undifferentiated, and multimodal. Given the above, classical gradient deterministic algorithms [4–6] are inappropriate for solving the problems of processing heterogeneous data.

The most common approaches to processing heterogeneous data are swarm intelligence algorithms (swarm algorithms). The most well-known swarm algorithms are the particle swarm optimization algorithm, artificial bee colony algorithm, firefly algorithm, ant colony optimization algorithm, wolf optimization algorithm and sparrow search algorithm [6–8].

However, most of the basic bio-inspired algorithms mentioned above are unable to maintain a balance between exploration and exploitation, resulting in poor performance for real-world complex optimization problems. This encourages the implementation of various strategies to improve the convergence rate and accuracy of the basic bio-inspired algorithms. Therefore, research on the development of new approaches to processing heterogeneous data is relevant.

2. Literature review and problem statement

The works [9–11] define the main advantages and disadvantages of cognitive algorithms. The shortcomings of these approaches include the lack of consideration of the type of uncertainty, the inability to conduct a search in different directions by several agents.

The work [12] presents an approach focused on the search for hidden information in large amounts of data. The method is based on analytical baselines, reducing variables, identifying sparse features and specifying rules. The disadvantages of this method include the inability to take into account various decision-making strategies, the lack of consideration of the type of uncertainty of the input data.

The works [13, 14] present an approach to transforming information models of objects to their equivalent structural models. This mechanism is designed to automate the necessary conversion, modification and addition operations during such information exchange. The disadvantages of the mentioned approach include the inability to assess the adequacy and reliability of the information transformation process, and make appropriate correction of the obtained models.

The work [15] proposes a method of fuzzy hierarchical assessment, which allows evaluating the quality of library services. The disadvantages of the specified method include the inability to assess the adequacy and reliability of the assessment and, accordingly, determine the assessment error.

The work [16] analyzes the 30 most common big data algorithms. It was found that the analysis of large amounts of data should be carried out in layers, take place in real time and have the opportunity for self-learning, search for solutions in different directions and take into account data noise.

The works [17, 18] present approaches to evaluating heterogeneous data for decision support systems based on clustering the basic set of input data, after which the system is trained based on the analysis. However, given the static architecture of artificial neural networks, errors are accumulated.

The work [19] carried out a comparative analysis of existing decision support technologies, namely: analytic hierarchy process, neural networks, fuzzy set theory, genetic algorithms and neuro-fuzzy modeling. The advantages and disadvantages of these approaches are indicated. For the problems of processing heterogeneous data under risk and uncertainty, using artificial neural networks and gradient algorithms is justified.

The works [20, 21] developed approaches to structural and objective analysis of the development of weakly structured systems. In this case, the problem is defined as the non-compliance of the existing state of a weakly structured system with the required one. At the same time, the disadvantages of the proposed approaches include the local optimum problem, the lack of consideration of system computing resources, and the inability to conduct a search in several directions.

The work [22] reviews evolutionary bio-inspired algorithms (Particle Swarm Optimization – PSO). These algorithms have proved effective in solving a number of rather complex problems and have already undergone a number of modifications. At the same time, these procedures are not without some drawbacks that worsen the properties of the global extremum search process.

An analysis of the works [9–22] showed that the common shortcomings of the above studies are:

– the lack of possibility of hierarchical processing of heterogeneous data;

– the lack of possibility of additional involvement of necessary system computing resources;

– the lack of consideration of the type of uncertainty and noise of data about information circulating in the system;

– the lack of deep learning mechanisms for knowledge bases;

– the lack of search priority in a certain direction.

3. The aim and objectives of the study

The aim of the study is to develop a method for increasing the efficiency of processing heterogeneous data using a metaheuristic algorithm. This will allow increasing the efficiency of processing heterogeneous data with a given reliability and developing subsequent management decisions. This will make it possible to develop software for intelligent decision support systems.

To achieve the aim, the following objectives were set:

– to determine the algorithm for implementing the method for increasing the efficiency of processing heterogeneous data;

– to give an example of applying the method for increasing the efficiency of processing heterogeneous data in analyzing the operational situation of a group of troops (forces).

4. Materials and methods

The problem solved in the study is to increase the efficiency of processing heterogeneous data circulating in information systems, regardless of the number of data sources. The object of the study is hierarchical systems. The subject of the study is the decision-making process in management problems using an advanced reptile algorithm (RA) and evolving artificial neural networks [23–31].

The hypothesis of the study is the possibility of increasing the efficiency of processing heterogeneous data while ensuring the processing reliability at 0.95, put forward to combat control systems.

Simulation of the proposed method for processing heterogeneous data was carried out in the MathCad 14 software environment (USA). The problem solved during the simulation was to assess the elements of the operational situation of a group of troops (forces). The hardware of the research process is AMD Ryzen 5.

The object of assessment was the operational group of troops (forces). The operational group of troops (forces) formed on the basis of an operational command with a standard composition of forces and means according to the wartime staff, and with a range of responsibilities under current regulations.

Initial data for determining the composition of the operational group of troops (forces) and elements of its operational structure using the method:

– the number of information sources about the state of the monitoring object – 3 (radio monitoring means, remote earth sensing tools and unmanned aerial vehicles). To simplify the simulation, the same number of each tool was taken – 4 tools each;

– the number of informational features for determining the state of the monitoring object – 12. These parameters include: affiliation, type of organizational and staff formation, priority, minimum width along the front, maximum width along the front. The number of personnel, the minimum depth along the flank, the maximum depth along the flank, the number of weapons and military equipment (WME) samples, the number of types of WME samples and the number of communication means, the type of operational structure are also taken into account;

– options for organizational and staff formations – company, battalion, brigade.

Parameters of the methodical approach:

– number of iterations – 100;

– number of individuals in the flock – 50;

– range of the feature space – [−150, 150].

The study is based on RA for processing heterogeneous data. For RA training, evolving artificial neural networks are used. The reptile optimizer was chosen due to the possibility of using different search strategies depending on available system computing resources. Evolving artificial neural networks allow learning not only parameters but also system architecture.

5. Development of a method for increasing the efficiency of processing heterogeneous data using a metaheuristic algorithm

5. 1. Algorithm of the method for increasing the efficiency of processing heterogeneous data using a metaheuristic algorithm

The study proposes an optimizer based on simulating reptile behavior (the case of crocodiles and alligators) – a population-based stochastic algorithm that uses reptile agents (RA) as search agents.

The method for increasing the efficiency of processing heterogeneous data using a metaheuristic algorithm consists of the following sequence of actions:

Step 1. Entering initial data. At this stage, the main parameters of the algorithm are defined, such as:

– type of problem being solved;

– number of agents in the population;

– type of data (structured, unstructured), archived, real-time data;

– number of variables characterizing the problem being solved;

– available system computing resources;

– type of uncertainty about the hierarchical system (complete uncertainty, partial uncertainty, complete awareness);

– volume and type of the training sample;

– volume and type of the test sample;

– architecture of an artificial neural network, etc.

Step 2. Creating an RA flock. The RA population X_i ($i = 1, 2,..., n$) is initialized. The set of RA form a population described by the matrix X. The initial RA population in this algorithm is generated taking into account the uncertainty about the state of the hierarchical system based on the constraints of the problem under consideration. Members of the RA population are search agents in the solution space, providing candidate values for problem variables based on their positions in the search space. Mathematically, each member of the general population is a vector whose number of elements is equal to the number of problem variables.

RA are set taking into account the uncertainty about the data circulating in the system on the grounds of the basic system model and circulating data models [2, 19, 21] (1):

$$X = \begin{bmatrix} X_1 \\ \cdot \\ \cdot \\ \cdot \\ X_i \\ \cdot \\ \cdot \\ \cdot \\ X_N \end{bmatrix}_{N \times m} = $$

$$= \begin{bmatrix} x_{1,1} \times \iota_{1,1} & \cdots & x_{1,d} \times \iota_{1,d} & \cdots & x_{1,m} \times \iota_{1,m} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{i,1} \times \iota_{i,1} & \cdots & x_{i,d} \times \iota_{i,d} & \cdots & x_{i,m} \times \iota_{i,m} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{N,1} \times \iota_{N,1} & \cdots & x_{N,d} \times \iota_{N,d} & \cdots & x_{N,m} \times \iota_{N,m} \end{bmatrix}, \quad (1)$$

where X is the RA population matrix, X_i is the i-th member of the RA flock (solution candidate), $x_{i,d}$ is the d-th dimension in the search space (solution variable), N is the number of RA, m is the number of solution variables.

Step 3. Numbering RA in the flock i, $i \in [0, S]$. At this stage, each RA is assigned a serial number. This allows determining the parameters of finding a solution for each individual in the flock.

Step 4. Determining the RA initial velocity.

The initial velocity v_0 of each RA is determined by the following expression:

$$v_i = (v_1, v_2...v_S), v_i = v_0. \quad (2)$$

The RA population updating process is based on simulating two strategies: the exploration phase and the exploitation phase.

Step 5. Preliminary assessment of the RA search area. In this procedure, the natural language search area is determined precisely by the halo of RA existence. Given that RA food sources are food of animal origin, it is advisable to sort the fitness of food sources (Step 6).

Step 6. Classifying RA food sources.

The location of the best food source (minimum fitness) is considered to be (FS_{ht}) food of animal origin (carrion) that is nearby and requires the least energy to find and obtain it. Delicacy food of animal origin is denoted as FS_{at}.

Other non-priority food sources (food that is necessary for the survival of individuals) are designated as FS_{nt}:

$$FS_{ht}=FS(\text{sorte\_index}(1)), \tag{3}$$

$$FS_{at}(1\!:\!3)=FS(\text{sorte\_index}(2\!:\!4)), \tag{4}$$

$$FS_{nt}(1\!:\!NP\text{-}4)=FS(\text{sorte\_index}(5\!:\!NP)). \tag{5}$$

Step 7. Determining the amount of available system computing resources.

At this stage, the amount of computing resources available for calculations is determined. In accordance with the positions outlined in Step 4, the concept of RA position updating is chosen.

Step 8. Exploration (prey encirclement).

The encirclement (exploration) phase is the phase of global exploration of the RA space. The RA exploration strategy is related to the current number of iterations. If $t\leq0.25T$, the RA will enter the high speed strategy. The exploration phase is described by the following mathematical expression:

$$Xnew_i^j = \begin{cases} XG^j \times -\eta_i^j \times \beta - R_i^j \times (\varsigma\cdot rand), t\leq\dfrac{T}{4}, \\ XG^j \times X_{r1}^j \times (\varsigma\cdot rand)\times ES, t>\dfrac{T}{4}\,mat\leq\dfrac{T}{2}, \end{cases} \tag{6}$$

where $Xnew_i^j$ is the j-th dimension of the i-th new RA solution, XG^j is the j-th dimension of the optimal solution obtained at the moment, t is the current iteration number, T is the maximum number of iterations. η_i^j is the j-th hunting operator of the i-th RA, which is calculated by expression (7), β is a constant and is equal to 0.1, R_i^j is the value of the i-th solution option used to reduce the search area by the j-th size, which is calculated by equation (8). r_1 randomly takes values from 1 to n. ς is the noise degree of the data circulating in the system, ES is a random number in the range from -2 to 2, when the number of iterations decreases and equation (9) will be used for calculation:

$$\eta_i^j = XG^j \times P_i^j, \tag{7}$$

$$R_i^j = \frac{XG^j - X_{r2}^j}{XG^j - \varepsilon}, \tag{8}$$

$$ES = 2\times r_3 \times\left(1-\frac{1}{T}\right), \tag{9}$$

$$P_i^j = \alpha + \frac{X_i^j - Md_i}{XG^j \times\left(UB^j - LB^j\right)+\varepsilon}. \tag{10}$$

In equation (8), ε is the minimum, r_2 is a number randomly selected from the range $1-n$. In equation (9), r_3 is a random integer from -1 to 1. In equation (10), $\alpha=0.1$, Md_i is the average position of the i-th solution candidate, which is calculated by equation (11):

$$Md_i = \frac{1}{d}\sum_{j=1}^{d} X_i^j, \tag{11}$$

where d is the dimensionality of the problem to be solved.

Step 9. Verification of reaching the global optimum. At this stage, the condition for the algorithm to reach the global optimum is checked according to the specified optimization criterion.

Step 10. Global restart procedure.

The restart procedure can effectively improve the algorithm's ability to go beyond the current optimum and im-

prove the algorithm's exploration capabilities. If the optimal population of the algorithm remains unchanged after ke iterations, the population is likely to fall into the local optimum. Thus, the candidate solution will be randomly initialized to speed up the exit from the global optimum:

$$Xnew_i^j = rand \times(UB-LB)+LB, nt>ke. \tag{12}$$

Step 11. Hunting phase (exploitation).

Under the condition $t>0.5T$, the hunting phase begins, at $t>3/4T$ and $t\leq T$ – the RA hunting cooperation strategy. Its equation for position updating is as follows:

$$Xnew_i^j = \begin{cases} XG^j \times P_i^j \times(\varsigma\cdot rand), t>\dfrac{T}{2}\,mat\leq3\dfrac{T}{4}, \\ XG^j \times \eta_i^j \times\varepsilon - R_{r1}^j \times(\varsigma\cdot rand), t>3\dfrac{T}{4}\,mat\leq T. \end{cases} \tag{13}$$

Finally, if the position of the new RA candidate is closer to food than the current one, the RA will move to the new candidate position and proceed to the next iteration:

$$X_i^j\left(t+1\right)= Xnew_i^j\left(t\right), \text{if } F\left(X_i(t)\right)> F\left(Xnew_i(t)\right), \tag{14}$$

where $F()$ is a function for calculating the fitness value, X_i is the location of the i-th candidate solution and $Xnew_i$ is the location of the i-th new candidate solution.

Step 12. Checking the stop criterion. The algorithm terminates if the maximum number of iterations is completed. Otherwise, the behavior of generating new locations and checking conditions is repeated.

Step 13. Training RA knowledge bases.

This study uses a training method based on evolving artificial neural networks developed in [2] to train the knowledge bases of each RA. The method is used to change the movement nature of each RA, for more accurate analysis results in the future.

The end of the algorithm.

5. 2. Example of applying the method for processing heterogeneous data using a metaheuristic algorithm

A method for processing heterogeneous data using a metaheuristic algorithm is proposed. To determine the effectiveness of the proposed method, it was simulated to solve the problem of determining the composition of the operational group of troops (forces) and elements of its operational structure in order to determine the expediency of regrouping troops (forces).

The efficiency of the method for processing heterogeneous data is compared with the swarm optimization algorithms listed in Table 1 for 23 unimodal and multimodal functions.

As shown in Table 1, a 15–19 % increase in the efficiency of processing heterogeneous data is achieved by using additional procedures.

It can be seen that the heterogeneous data processing method is able to converge to the true value for most unimodal functions with the fastest convergence rate and the highest accuracy, while the convergence results of the ant colony optimization algorithm are far from satisfactory.

Table 1

Comparison of the proposed method with other swarm algorithms for unimodal and multimodal functions

Function	Value	Bee colony algorithm [32]	Firefly algorithm [18]	Ant colony optimization algorithm [33]	Proposed method
1	2	3	4	5	6
F1	Average	5.1456×10^{-26}	1.3752×10^{-30}	2.0262×10^{-34}	2.2410×10^{-36}
	Standard	0	0	0	0
	Better	0	0	0	0
F2	Average	2.7982×10^{-25}	5.1533×10^{-30}	7.6024×10^{-34}	1.2270×10^{-35}
	Standard	0	0	0	0
	Better	0	0	0	0
F3	Average	1.2335×10^{-15}	1.3169×10^{-17}	2.0901×10^{-21}	2.9870×10^{-26}
	Standard	0	0	0	0
	Better	0	0	0	0
F4	Average	2.7063×10^{-15}	3.4440×10^{-17}	4.9966×10^{-21}	1.1231×10^{-25}
	Standard	0	0	0	0
	Better	0	0	0	0
F5	Average	3.1633×10^{-23}	8.5720×10^{-22}	9.3947×10^{-30}	1.0063×10^{-4}
	Standard	9.9721×10^{-23}	4.6942×10^{-21}	3.3014×10^{-29}	5.5042×10^{-4}
	Better	3.8292×10	4.5361×10	5.3971×10	9.2311×10^{1}
F6	Average	1.2740×10	2.9681×10	5.0036×10	1.6587×10
	Standard	9.8948×10^{1}	9.8953×10^{1}	9.8920×10^{1}	9.8708×10^{1}
	Better	5.2147×10^{-2}	1.0042×10^{-2}	6.5905×10^{-2}	1.9575×10^{-1}
F7	Average	0.0000×10	0.0000×10	0.0000×10	0.0000×10
	Standard	1.5136×10^{-4}	1.7352×10^{-4}	1.1218×10^{-4}	1.3047×10^{-3}
	Better	1.3928×10^{-4}	2.1698×10^{-4}	1.2010×10^{-4}	7.8270×10^{-4}
F9	Average	-1.4903×10^{4}	-1.4796×10^{4}	-1.4585×10^{4}	-2.1956×10^{4}
	Standard	4.5009×10^{2}	4.0984×10^{2}	3.3323×10^{2}	4.2788×10^{2}
	Better	8.4524×10^{-14}	3.5231×10^{-15}	1.2730×10^{-15}	1.9329×10^{1}
F10	Average	2.4061×10^{-13}	8.2258×10^{-15}	2.7492×10^{-15}	3.6507×10
	Standard	1.1622×10	1.1296×10	1.0991×10	3.2542×10^{-1}
	Better	6.1114×10^{-2}	4.8698×10^{-2}	6.1272×10^{-2}	7.2941×10^{-2}
F11	Average	5.8013×10^{-46}	5.1764×10^{-51}	4.6656×10^{-49}	5.0181×10^{-42}
	Standard	3.0141×10^{-45}	1.5408×10^{-50}	8.7556×10^{-49}	1.5187×10^{-41}
	Better	6.4561×10^{-31}	4.8992×10^{-33}	2.0915×10^{-35}	3.1866×10^{-30}
F12	Average	1.5124×10^{-30}	1.7674×10^{-32}	3.9040×10^{-35}	1.2116×10^{-29}
	Standard	5.7721×10^{3}	1.1202×10^{4}	2.6989×10^{4}	5.6676×10^{4}
	Better	1.5215×10^{4}	2.5396×10^{4}	2.9479×10^{4}	4.7540×10^{4}
F13	Average	9.2891×10^{1}	9.1533×10^{1}	9.1888×10^{1}	7.1597×10^{1}
	Standard	1.3709×10	1.7876×10	2.7794×10	2.6573×10
	Better	9.8629×10^{1}	9.8675×10^{1}	9.8851×10^{1}	9.8845×10^{1}
F14	Average	3.6026×10^{-1}	2.9301×10^{-1}	7.3941×10^{-2}	5.0345×10^{-2}
	Standard	9.7175×10^{-4}	1.2665×10^{-3}	1.3941×10^{-3}	1.5076×10^{-3}
	Better	9.3030×10^{-4}	9.5681×10^{-4}	1.1079×10^{-3}	1.8053×10^{-3}
F15	Average	-2.2532×10^{4}	-2.3085×10^{4}	-2.3315×10^{4}	-2.3156×10^{4}
	Standard	5.4506×10^{2}	4.3601×10^{2}	5.5671×10^{2}	4.7225×10^{2}
	Better	1.7974×10^{1}	1.9291×10^{1}	1.9948×10^{1}	1.9292×10^{1}
F16	Average	6.0939×10^{0}	3.6436×10^{0}	2.4914×10^{-2}	3.6437×10^{0}
	Standard	7.4214×10^{-1}	7.3596×10^{-1}	7.9325×10^{-1}	7.3495×10^{-1}
	Better	7.7846×10^{-2}	7.1764×10^{-2}	6.5839×10^{-2}	8.4424×10^{-2}
F18	Average	1.528×10^{-1}	1.688×10^{-1}	8.445×10^{-2}	8.433×10^{-2}
	Standard	8.364×10^{-2}	1.419×10^{-1}	9.550×10^{-2}	1.312×10^{-1}
	Better	1.461×10^{-1}	1.382×10^{-1}	1.033×10^{-1}	2.180×10^{-1}

Continuation of the Table 1

1	2	3	4	5	6
F19	Average	1.043×10^{-1}	8.713×10^{-2}	9.100×10^{-2}	8.272×10^{-2}
	Standard	1.070×10^{-1}	9.662×10^{-2}	8.159×10^{-2}	1.306×10^{-1}
	Better	1.280×10^{-1}	1.978×10^{-1}	1.335×10^{-1}	2.230×10^{-1}
F20	Average	1.760×10^{-1}	1.373×10^{-1}	1.590×10^{-1}	1.210×10^{-1}
	Standard	1.449×10^{-1}	1.735×10^{-1}	1.327×10^{-1}	1.822×10^{-1}
	Better	1.324×10^{-1}	1.870×10^{-1}	1.685×10^{-1}	2.078×10^{-1}
F21	Average	2.481×10^{-1}	1.460×10^{-1}	2.100×10^{-1}	2.723×10^{-1}
	Standard	1.350×10^{-1}	1.449×10^{-1}	1.467×10^{-1}	1.375×10^{-1}
	Better	1.223×10^{-1}	1.145×10^{-1}	1.172×10^{-1}	1.386×10^{-1}
F22	Average	9.012×10^{-2}	1.000×10^{-1}	8.897×10^{-2}	5.730×10^{-2}
	Standard	1.944×10^{-1}	1.090×10^{-1}	5.980×10^{-2}	7.212×10^{-2}
	Better	1.000×10^{-1}	1.172×10^{-1}	9.750×10^{-2}	1.078×10^{-1}
F23	Average	6.104×10^{-2}	9.397×10^{-2}	9.970×10^{-2}	5.727×10^{-2}
	Standard	7.606×10^{-2}	6.572×10^{-2}	7.230×10^{-2}	4.779×10^{-2}
	Better	1.613×10^{-1}	1.657×10^{-1}	1.494×10^{-1}	1.570×10^{-1}

6. Discussion of the results of developing a method for processing heterogeneous data using a metaheuristic algorithm

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

– the initial RA setting is carried out taking into account the type of uncertainty (Step 2) using appropriate correction factors for the awareness degree of the data circulating in the system, compared to [9, 14, 21];

– the initial velocity of each RA is taken into account (Step 4), which allows prioritizing the search of each RA, compared to [9–15];

– the fitness of RA feeding locations is determined, which reduces the data processing time (Step 5), compared to [14, 16, 17];

– universality of RA food location search strategies, which allows classifying the type of data to be processed (Step 5, 6), compared to [14, 16, 17];

– the degree of data noise in the process of RA position updating (Steps 8–11) is taken into account, which reduces the data processing time, compared to [9–15];

– the degree of data noise during processing is taken into account, which increases the reliability of heterogeneous data (Steps 8–11), reducing the data processing time, compared to [9–15];

– universality of solving the problem of processing heterogeneous RA data circulating in the system as a whole, due to the hierarchical nature of their description (Steps 1–13, Table 1), compared to [9, 12, 13–18];

– the possibility of simultaneous search for a solution in different directions (Steps 1–13, Table 1);

– the adequacy of the results obtained (Steps 1–13), compared to [9–23];

– the ability to avoid the local extremum problem (Steps 1–13);

– the possibility of deep learning of RA knowledge bases (Step 13), compared to [9–23].

The disadvantages of the proposed method include:

– the loss of informativeness when processing heterogeneous data due to the construction of the membership function;

– lower accuracy of processing homogeneous data due to gradient search (the average value reached 4–7 %), expression (10);

– the loss of credibility of the obtained solutions when searching for a solution in several directions simultaneously;

– lower assessment accuracy compared to other assessment approaches.

This method will allow you:

– to process heterogeneous data circulating in the system;

– to determine effective measures for increasing the efficiency of processing heterogeneous data while maintaining the given reliability;

– to reduce the use of computing resources of decision support systems.

The limitations of the study are the need for an initial database on the state of hierarchical systems, the need to take into account the delay time for collecting and communicating information from intelligence sources.

The proposed method should be used to solve the problems of processing heterogeneous data characterized by a high degree of complexity.

This study is a further development of research aimed at developing methodical principles for increasing the efficiency of processing heterogeneous data, published earlier [2, 4–6, 21–23].

Areas of further research should be aimed at reducing computing costs while processing heterogeneous data in special-purpose systems.

7. Conclusions

1. The algorithm for implementing the method is defined, which differs from the canonical one due to additional and improved procedures:

– the type of uncertainty and noise of data circulating in the system is taken into account;

– RA food sources are sorted;

– the available system computing resources are taken into account, which allows selecting the RA hunting strategy;

– initial RA setting, taking into account the type of uncertainty;

– training of knowledge bases, which is carried out by training the synaptic weights of an artificial neural network, the type and parameters of the membership function, as well as the architecture of individual elements and the architecture of the artificial neural network as a whole.

2. An example of applying the proposed methical approach in solving the problem of determining the composition of an operational group of troops (forces) and elements of its operational structure is given. This example showed a 15–19 % increase in data processing efficiency by using 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.

References

1. Bashkyrov, O. M., Kostyna, O. M., Shyshatskyi, A. V. (2015). Rozvytok intehrovanykh system zviazku ta peredachi danykh dlia potreb Zbroinykh Syl. Ozbroiennia ta viyskova tekhnika, 1, 35–39. Available at: http://nbuv.gov.ua/UJRN/ovt_2015_1_7

2. Dudnyk, V., Sinenko, Y., Matsyk, M., Demchenko, Y., Zhyvotovskyi, R., Repilo, I. et al. (2020). Development of a method for training artificial neural networks for intelligent decision support systems. Eastern-European Journal of Enterprise Technologies, 3 (2 (105)), 37–47. https://doi.org/10.15587/1729-4061.2020.203301

3. Sova, O., Shyshatskyi, A., Salnikova, O., Zhuk, O., Trotsko, O., Hrokholskyi, Y. (2021). Development of a method for assessment and forecasting of the radio electronic environment. EUREKA: Physics and Engineering, 4, 30–40. https://doi.org/10.21303/2461-4262.2021.001940

4. Pievtsov, H., Turinskyi, O., Zhyvotovskyi, R., Sova, O., Zvieriev, O., Lanetskii, B., Shyshatskyi, A. (2020). Development of an advanced method of finding solutions for neuro-fuzzy expert systems of analysis of the radioelectronic situation. EUREKA: Physics and Engineering, 4, 78–89. https://doi.org/10.21303/2461-4262.2020.001353

5. Zuiev, P., Zhyvotovskyi, R., Zvieriev, O., Hatsenko, S., Kuprii, V., Nakonechnyi, O. et al. (2020). Development of complex methodology of processing heterogeneous data in intelligent decision support systems. Eastern-European Journal of Enterprise Technologies, 4 (9 (106)), 14–23. https://doi.org/10.15587/1729-4061.2020.208554

6. Shyshatskyi, A. (2020). Complex Methods of Processing Different Data in Intellectual Systems for Decision Support System. International Journal of Advanced Trends in Computer Science and Engineering, 9 (4), 5583–5590. https://doi.org/10.30534/ijatcse/2020/206942020

7. Yeromina, N., Kurban, V., Mykus, S., Peredrii, O., Voloshchenko, O., Kosenko, V. et al. (2021). The Creation of the Database for Mobile Robots Navigation under the Conditions of Flexible Change of Flight Assignment. International Journal of Emerging Technology and Advanced Engineering, 11 (5), 37–44. https://doi.org/10.46338/ijetae0521_05

8. Rotshteyn, A. P. (1999). Intellektual'nye tekhnologii identifikatsii: nechetkie mnozhestva, geneticheskie algoritmy, neyronnye seti. Vinnitsa: «UNIVERSUM», 320.

9. Ko, Y.-C., Fujita, H. (2019). An evidential analytics for buried information in big data samples: Case study of semiconductor manufacturing. Information Sciences, 486, 190–203. https://doi.org/10.1016/j.ins.2019.01.079

10. Ramaji, I. J., Memari, A. M. (2018). Interpretation of structural analytical models from the coordination view in building information models. Automation in Construction, 90, 117–133. https://doi.org/10.1016/j.autcon.2018.02.025

11. Pérez-González, C. J., Colebrook, M., Roda-García, J. L., Rosa-Remedios, C. B. (2019). Developing a data analytics platform to support decision making in emergency and security management. Expert Systems with Applications, 120, 167–184. https://doi.org/10.1016/j.eswa.2018.11.023

12. Chen, H. (2018). Evaluation of Personalized Service Level for Library Information Management Based on Fuzzy Analytic Hierarchy Process. Procedia Computer Science, 131, 952–958. https://doi.org/10.1016/j.procs.2018.04.233

13. Chan, H. K., Sun, X., Chung, S.-H. (2019). When should fuzzy analytic hierarchy process be used instead of analytic hierarchy process? Decision Support Systems, 125, 113114. https://doi.org/10.1016/j.dss.2019.113114

14. Osman, A. M. S. (2019). A novel big data analytics framework for smart cities. Future Generation Computer Systems, 91, 620–633. https://doi.org/10.1016/j.future.2018.06.046

15. Gödri, I., Kardos, C., Pfeiffer, A., Váncza, J. (2019). Data analytics-based decision support workflow for high-mix low-volume production systems. CIRP Annals, 68 (1), 471–474. https://doi.org/10.1016/j.cirp.2019.04.001

16. Harding, J. L. (2013). Data quality in the integration and analysis of data from multiple sources: some research challenges. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-2/W1, 59–63. https://doi.org/10.5194/isprsarchives-xl-2-w1-59-2013

17. Kosko, B. (1986). Fuzzy cognitive maps. International Journal of Man-Machine Studies, 24 (1), 65–75. https://doi.org/10.1016/s0020-7373(86)80040-2

18. Koval, M., Sova, O., Shyshatskyi, A., Artabaiev, Y., Garashchuk, N., Yivzhenko, Y. et al. (2022). Improving the method for increasing the efficiency of decision-making based on bio-inspired algorithms. Eastern-European Journal of Enterprise Technologies, 6 (4 (120)), 6–13. https://doi.org/10.15587/1729-4061.2022.268621

19. Maccarone, A. D., Brzorad, J. N., Stone, H. M. (2008). Characteristics and Energetics of Great Egret and Snowy Egret Foraging Flights. Waterbirds, 31 (4), 541–549. https://doi.org/10.1675/1524-4695-31.4.541

20. Koshlan, A., Salnikova, O., Chekhovska, M., Zhyvotovskyi, R., Prokopenko, Y., Hurskyi, T. et al. (2019). Development of an algorithm for complex processing of geospatial data in the special-purpose geoinformation system in conditions of diversity and uncertainty of data. Eastern-European Journal of Enterprise Technologies, 5 (9 (101)), 35–45. https://doi.org/10.15587/1729-4061.2019.180197

21. Mahdi, Q. A., Shyshatskyi, A., Prokopenko, Y., Ivakhnenko, T., Kupriyenko, D., Golian, V. et al. (2021). Development of estimation and forecasting method in intelligent decision support systems. Eastern-European Journal of Enterprise Technologies, 3 (9 (111)), 51–62. https://doi.org/10.15587/1729-4061.2021.232718

22. Gorokhovatsky, V., Stiahlyk, N., Tsarevska, V. (2021). Combination method of accelerated metric data search in image classification problems. Advanced Information Systems, 5 (3), 5–12. https://doi.org/10.20998/2522-9052.2021.3.01

23. Braik, M., Ryalat, M. H., Al-Zoubi, H. (2021). A novel meta-heuristic algorithm for solving numerical optimization problems: Ali Baba and the forty thieves. Neural Computing and Applications, 34 (1), 409–455. https://doi.org/10.1007/s00521-021-06392-x

24. Meleshko, Y., Drieiev, O., Drieieva, H. (2020). Method of identification bot profiles based on neural networks in recommendation systems. Advanced Information Systems, 4 (2), 24–28. https://doi.org/10.20998/2522-9052.2020.2.05

25. Kuchuk, N., Merlak, V., Skorodelov, V. (2020). A method of reducing access time to poorly structured data. Advanced Information Systems, 4 (1), 97–102. https://doi.org/10.20998/2522-9052.2020.1.14

26. Shyshatskyi, A., Tiurnikov, M., Suhak, S., Bondar, O., Melnyk, A., Bokhno, T., Lyashenko, A. (2020). Method of assessment of the efficiency of the communication of operational troop grouping system. Advanced Information Systems, 4 (1), 107–112. https://doi.org/10.20998/2522-9052.2020.1.16

27. Raskin, L., Sira, O. (2016). Method of solving fuzzy problems of mathematical programming. Eastern-European Journal of Enterprise Technologies, 5 (4 (83)), 23–28. https://doi.org/10.15587/1729-4061.2016.81292

28. Lytvyn, V., Vysotska, V., Pukach, P., Brodyak, O., Ugryn, D. (2017). Development of a method for determining the keywords in the slavic language texts based on the technology of web mining. Eastern-European Journal of Enterprise Technologies, 2 (2 (86)), 14–23. https://doi.org/10.15587/1729-4061.2017.98750

29. Stepanenko, A., Oliinyk, A., Deineha, L., Zaiko, T. (2018). Development of the method for decomposition of superpositions of unknown pulsed signals using the secondorder adaptive spectral analysis. Eastern-European Journal of Enterprise Technologies, 2 (9 (92)), 48–54. https://doi.org/10.15587/1729-4061.2018.126578

30. Gorbenko, I., Ponomar, V. (2017). Examining a possibility to use and the benefits of post-quantum algorithms dependent on the conditions of their application. Eastern-European Journal of Enterprise Technologies, 2 (9 (86)), 21–32. https://doi.org/10.15587/1729-4061.2017.96321

31. Koval, M., Sova, O., Orlov, O., Shyshatskyi, A., Artabaiev, Y., Shknai, O. et al. (2022). Improvement of complex resource management of special-purpose communication systems. Eastern-European Journal of Enterprise Technologies, 5 (9 (119)), 34–44. https://doi.org/10.15587/1729-4061.2022.266009

32. Shyshatskyi, A., Stasiuk, T., Kuzmenko, O. (2023). The Development of the Solution Search Method Based on the Improved Bee Colony Algorithm. Theoretical and Applied Cybersecurity, 5 (2). https://doi.org/10.20535/tacs.2664-29132023.2.278199

33. Sova, O., Zhuravskyi, Y., Vakulenko, Y., Shyshatskyi, A., Salnikova, O., Nalapko, O. (2022). Development of methodological principles of routing in networks of special communication in conditions of fire storm and radio-electronic suppression. EUREKA: Physics and Engineering, 3, 159–166. https://doi.org/10.21303/2461-4262.2022.002434