

Розроблено алгоритм комплексної обробки геопросторових даних в геоінформаційних системах спеціального призначення в умовах різноманітності та невизначеності даних. Новизна алгоритму полягає в забезпеченні функціонування геоінформаційної системи в умовах дефіциту обчислювальних ресурсів з врахуванням невизначеності про стану об'єкту моніторингу (об'єкту розвідки). Зазначений алгоритм враховує при обробці геопросторових даних, що циркулюють в геоінформаційній системі спеціального призначення коефіцієнт відносної значущості подій, що виникають. Запропонований алгоритм використовує розроблений комплексний показник виникаючих подій, який характеризує ймовірність виконання геоінформаційною системою поставлених перед нею завдань, повноту їх вирішення за цикл управління і врахуванні значимості виникаючих подій. Запропоновано взаємодоповнюючі один одного підходи до управління ресурсами геоінформаційних систем спеціального призначення. Розробка запропонованого алгоритму обумовлена необхідністю підвищення оперативності обробки різноманітної інформації в геоінформаційних системах з прийнятною обчислювальною складністю. Запропонований алгоритм дозволяє підвищити ефективність функціонування геоінформаційних систем за рахунок комплексної обробки геопросторових даних, що в ній циркулюють. Зазначений алгоритм доцільно використовувати при розробці програмного забезпечення для геоінформаційних системах спеціального призначення для підвищення їх ефективності за рахунок підвищення оперативності обробки інформації в геоінформаційних системах спеціального призначення. Запропонований алгоритм дозволяє підвищити оперативність обробки інформації в геоінформаційних системах спеціального призначення від 16 до 20 % в залежності від кількості інформації про об'єкт моніторингу

Ключові слова: геоінформаційна система, об'єкт моніторингу, геопросторові дані, обчислювальна складність, обробка інформації, тип інформації

Received date 13.08.2019

Accepted date 01.10.2019

Published date 28.10.2019

1. Introduction

The rapid development of geospatial data sources in the world significantly complicates the process of extraction of

UDC 621.396

DOI: 10.15587/1729-4061.2019.180197

DEVELOPMENT OF AN ALGORITHM FOR COMPLEX PROCESSING OF GEOSPATIAL DATA IN THE SPECIAL-PURPOSE GEOINFORMATION SYSTEM IN CONDITIONS OF DIVERSITY AND UNCERTAINTY OF DATA

A. Koshlan
Adjunct*

O. Salmikova

Doctor of Public Administration Sciences, Senior Researcher,
Head of Educational and Research Center
Educational and Research Center of Strategic Communications
in the Sphere of National Security and Defense**

M. Chekhovska

Doctor of Economic Sciences, Professor, Head of Department
Department of National Security
National Academy of the Security Service of Ukraine
Mykhaila Maksymovycha str., 22, Kyiv, Ukraine, 03022

R. Zhyvotovskiy

PhD, Senior Researcher, Chief of Research Department
Research Department of the Development of
Anti-Aircraft Missile Systems and Complexes ***

Ye. Prokopenko
PhD*

T. Hurskiy

PhD, Associate Professor
Military Institute of Telecommunications and Information Technologies
named after Heroes of Kruty
Moskovska str., 45/1, Kyiv, Ukraine, 01011

O. Yefymenko

PhD, Associate Professor
Department of Highway Building Machines
Kharkiv National Automobile and Highway University
Yaroslava Mudroho str., 25, Kharkiv, Ukraine, 61002

Ye. Kalashnikov

PhD, Head of Research Laboratory
Research Laboratory of Problems of Development
of Combat use of Rocket Forces and Artillery**

S. Petruk

Deputy Chief of Research Department
Research Department of the Development
of Anti-Aircraft Missile Systems and Complexes ***

A. Shyshatskiy

PhD, Chief of Research Laboratory
Research Laboratory of Electronic Warfare Development***
E-mail: ierikon12@gmail.com

*Institute of Information Technology**

**Ivan Chernyakhovsky National Defense University of Ukraine
Povitrofliski ave., 28, Kyiv, Ukraine, 03049

***Central Scientific-Research Institute of Arming and
Military Equipment of the Armed Forces of Ukraine
Povitrofliski ave., 28, Kyiv, Ukraine, 03168

Copyright © 2019, A. Koshlan, O. Salmikova, M. Chekhovska, R. Zhyvotovskiy,

Ye. Prokopenko, T. Hurskiy, O. Yefymenko, Ye. Kalashnikov, S. Petruk, A. Shyshatskiy.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0>)

geospatial data using technical devices and necessitates their compilation and processing of geospatial data coming from various technical devices [1, 2]. The main advantage of geoinformation technologies over other information technologies is

the set of tools for creating and combining databases with the possibilities of their graphical analysis and visualization [1, 2].

Currently, geoinformation systems (GIS) for military purposes are increasingly being used to solve the problems of modeling the processes and situations of military use of troops. This allows us to talk about the appearance of a new class of GIS – intelligent geoinformation systems that use artificial intelligence to collect and process information.

Devices of remote sensing of the Earth (RSE), forces and devices of electronic intelligence are the main sources of information for solving various computational and analytical tasks.

The processing of different types of intelligence from various sources of information requires considerable computational operations, with strict restrictions on the time of calculations.

This leads to the search for new scientific approaches to the processing of heterogeneous geospatial information in order to increase the operational efficiency of special-purpose geoinformation systems.

Considering the above, it is necessary to solve an actual scientific problem, which consists in developing an algorithm for processing various types of information in order to reduce the time to make a decision on the status of monitoring objects (MO).

All this confirms the relevance of the chosen research direction.

2. Literature review and problem statement

In [3], the analysis of known methods of processing different types of information was carried out. The analysis revealed that the researchers highlight a number of problems associated with the process of data extraction for further analytical processing:

- data in sources, as a rule, are presented in various formats, coding and forms, while the solution of analytical problems involves the use of a single, universal format, which will be supported by the data warehouse and analytical applications;
- excessively detailed data contained in the sources must be cleaned and summarized. In this case, the methods and algorithms that are intended for this purpose are often more complex than the analysis algorithms themselves;
- the lack of integrated use of information processing and distribution methods.

In [4], a generalized metric was developed in the problem of multivariate data analysis with different types of features. The essence of the proposed metric is that the metric allows you to build clustering, classification, and association algorithms based on it using classical processing methods.

The proposed metric is intended to estimate the proximity of objects with specified features, which allows it to be reduced to scalar numerical values. This allows us to reduce the problem to the classical numerical form and provides a fundamental opportunity to apply known methods and algorithms for its solution. However, this metric does not allow effective functioning in a deficit of computing resources.

In [5], an approach to in-depth analysis of various types of data affecting the energy efficiency of buildings is proposed, which is based on the representation of a hierarchy of factors in the form of a multidimensional cube with different levels of abstraction. This approach allows for a multi-level description of the object, but does not take into account the uncertainty about the status of the monitoring object, which

does not allow a full assessment of its status. This approach is geared toward using sufficient computing resources.

In [6], an approach for processing various types of data obtained from an unmanned aerial vehicle implemented in the GRASS GIS software environment was presented. This approach is based on three-dimensional raster image processing techniques with a further reduction of their redundancy. However, this approach is intended only for processing graphic information and does not take into account the type of uncertainty about the status of the monitoring object.

In [7], a description of the work of spatial and temporal degradation of soil erosion was performed. The description of the soil erosion process is described using the erosion potential method. This method is based on the analytical processing of various types of data on the factors affecting the erosion process. The method is characterized by a high degree of reliability, less computational complexity, simplicity and adapted for the use in GIS. However, this approach is intended only for processing multiple mapping information with sufficient computing resources available. This feature limits its scope.

In [8], a method of binary classification of users of social networks, which is based on the method of logical regression was presented. This method allows to process various types of information about users of social networks. However, this method requires considerable computing resources and does not take into account the uncertainty about the state of the monitoring object (in this case, social network users).

In [9], the problem of processing information from heterogeneous technical monitoring tools is considered. As a possible solution to the problem, it is proposed to use a generalized methodology for information processing, which is based on the clustering methodology of territorially combined monitoring information sources and using the framework model of the knowledge base for the identification of monitoring objects. The clustering technique is based on the Lance-Williams hierarchical agglomerative procedure using the Ward metric. The framework of the knowledge base is built using object-oriented modeling tools. The disadvantages of the proposed generalized methodology include the neglect of the relative significance of the events that occur and the inability to operate in the face of a scarcity of computing resources. The disadvantages of this method also include the inability to redistribute computing resources between elements to increase the efficiency of information processing.

In [10], a method for processing different types of acoustic information from different sources of origin to identify the level of information security of unmanned autonomous objects was developed. This method is based on the use of a two-layer neural network with sigmoid hidden neurons. However, this method is intended only for use with acoustic information, requires considerable computational resources and does not take into account the degree of uncertainty about the condition of the monitoring object.

In [11], a method for determining the type of signal modulation based on a convolutional neural network by analyzing various signal parameters was proposed. This method is highly efficient, but it can only be used to solve radio monitoring tasks, requires considerable computing resources, and does not take into account the degree of uncertainty about the status of the monitoring object.

In [12], a method for identifying signals for unmanned aerial vehicles was proposed. The method is based on an artificial neural network and uses the knowledge base of radio wave propagation, taking into account geographical coordinates.

The disadvantages of this method are that the method is limited to solving radio monitoring problems, requires considerable computing resources and does not take into account the degree of uncertainty about the state of the monitoring object.

In [13], a methodology for correcting common geometric and topological errors in geoinformation systems was developed. This methodology is intended to correct errors that occur while converting heterogeneous data in geoinformation systems from analogue to digital. However, this methodology is not intended to handle information differently than geometric information, requires considerable computational resources and does not take into account the degree of uncertainty about the state of the monitoring object.

In [14], an approach for the transformation of aerial photographs and satellite images, which is based on the results of their geomorphological, geobotanical, reclamation, erosion and other surveys in geoinformation systems into a digital landscape model was proposed. However, this approach does not take into account other types of information circulating in geoinformation systems and does not take into account the uncertainty about the status of the monitoring object.

In [15], an intelligent system for processing various types of data circulating in geoinformation systems is proposed. This intelligent system is designed to solve the problems of geological exploration in geoinformation systems. The essence of this approach is that the components of the target mineral system that are mapped are converted into a set of cards, which leads to the automatic updating of the map. However, this intelligent system is only for geological exploration purposes and requires considerable computing resources.

In [16], an approach for monitoring the state of the grid using a geoinformation system was proposed. The essence of this approach consists in the fact that on the basis of complex processing of information about the state of the grid, its complexation is carried out. However, this intelligent system is intended solely for monitoring the state of the grid, requires considerable computing resources and does not take into account the uncertainty about the condition of the monitored object.

In [17], an integrated GIS platform architecture designed to meet the requirements of real-time processing of space-time data using cloud computing was proposed. This platform does not take into account the relative importance of the events to be analyzed and further processed and requires considerable computing resources.

In [18], an approach to the use of GIS for processing and submitting geotechnical data to formats that are useful to engineers, planners and land-use specialists was described. This approach significantly reduces the time for processing data circulating in the GIS. However, this approach is oriented solely for use in the field of land management without regard to the importance of information circulating in the system, which requires significant computing resources, is not able to distribute information to increase the efficiency of its processing.

In [19], the use of GIS for processing various spatial data for optimization of urban planning is considered. This approach allows us to increase the energy efficiency of urban development, to optimize transport connections and to propose city development strategies. The proposed approach requires considerable computing resources and does not take into account the degree of uncertainty of the monitoring object.

In [20], an approach for processing GIS-based data of different types was presented. The essence of the proposed approach is to analyze the energy efficiency of buildings by technical, economic, environmental and social criteria.

Energy balance and spatial analyses were performed on the basis of a geographical information system. This approach requires considerable computing resources and complete information on the status of the monitoring object.

The analysis showed that the known methods (techniques):

- do not allow high-quality processing of large arrays of different types of data of numerical and quantitative origin;
- have great computational complexity;
- do not take into account the level of awareness of the status of the monitoring object;
- do not allow in a complex to process and distribute information about the status of the monitoring object.

Therefore, it is necessary to develop an algorithm for processing various types of information in special-purpose geoinformation systems, which is able to efficiently process and distribute large arrays of data under uncertainty, as well as a shortage of computing resources.

3. The aim and objectives of the research

The purpose of the research is to develop an algorithm for integrated processing of geospatial data in special-purpose geoinformation systems in the conditions of data diversity and uncertainty about the status of the monitoring object. This algorithm allows for complex processing and distribution of information from multiple sources of information, takes into account the uncertainty about the state of the monitoring object and has a moderate computational complexity.

To achieve this goal, the following tasks were set:

- to formulate the task of complex processing of geospatial data;
- to develop an algorithm for complex processing of geospatial data;
- to identify the advantages and disadvantages of the proposed algorithm.

4. Formulation of the task of complex processing of geospatial data

Let us consider one of the main components of a single technological cycle (STC) of integrated processing of geospatial data (GSD) – management of the resources of the (system of complex processing of geospatial data) SCP GSD in order to maintain the stability of its functioning in all conditions.

The quality of functioning of the SCP GSD can be described by a set of the following interrelated criteria:

- 1) the number $\tilde{u}(\Delta t)$ of recorded events $\{\tilde{s}_i\}$ on the MO over a period of time Δt ;
- 2) the time τ required to develop a single information document;
- 3) the likelihood of the system performing tasks during the management cycle.

You should also enter a comprehensive indicator that characterizes the completeness of solving the tasks by the GSD complex processing system in a given period of time (per control cycle):

$$P = 1 - e^{-\frac{\tau_s}{\tau_0}} = 1 - p(\Delta t), \quad p(\Delta t) = \frac{\tilde{u}(\Delta t)}{u(\Delta t)} \times 100\%, \quad (1)$$

where $u(\Delta t)$ is the number of events that actually took place $\{s_i\}$, that happened during the period, Δt is the probability

of completing a control cycle task, τ_0 is the average time to complete the task on the control cycle, τ_s is the allowed time to solve problems for the control cycle.

Thus, the management of the SCP GSD resources is reduced to the problem of rationally allocating them between tasks that are solved in such a way that $p(\Delta t)$ is the maximum (tends to 100 %).

However, in reality, the indicator $p(\Delta t)$ in the formula (1) cannot fully characterize the quality of functioning of the SCP GSD, since it does not take into account such a characteristic as the relative significance of the recorded events.

Let events on the MO be divided into M types. The magnitude ζ_m of each event type can be determined by the number m of its type or by a value proportional to m . The proportionality factor is convenient to choose from the condition that the sum of the significance of events of all types is normalized by 1:

$$\sum_{m=1}^M (\zeta_m) = 1. \tag{2}$$

As a result, the quality of functioning of the SCP GSD can be characterized by a complex indicator $P(\Delta t)$, which takes into account the significance of the recorded events and is defined as follows:

$$P(\Delta t) = \frac{\tilde{U}(\Delta t)}{U(\Delta t)} \cdot 100\% = \frac{\sum_{m=1}^M (\zeta_m \tilde{u}_m)}{\sum_{m=1}^M (\zeta_m u_m)} \cdot 100\%, \tag{3}$$

where $\tilde{U}(\Delta t)$ is the number of recorded events on the MO for a given period of time, taking into account their relative importance; $U(\Delta t)$ is the number of events that actually occurred during the same period.

Herewith \tilde{u}_m is defined as the number of recorded $\{\tilde{s}_{mi}\}$, but u_m is defined as the number of events of a given type that occurred $\{s_{mi}\}$.

We formulate the problem of rational allocation of the SCP GSD resources as follows.

Let the SCP GSD entry receive tasks to respond to events of varying significance. The frequency of receipt of tasks can be described by the normal law of distribution. Each task requires its own resources to be connected. It is necessary to redistribute system resources in such a way as to ensure the maximum possible value of the integrated indicator $P(\Delta t) \rightarrow \max$.

The solution to this problem is possible while applying the following approaches:

1. Allocation of the SCP GSD resources to solve problems of response to events of different types, taking into account their relative importance.
2. Minimizing the total time τ , which is required to perform another m -type response.

Suppose there are N types of resources for solving the tasks in the SCP GSD.

The need for resources is specified by the matrix $R = \{r_{nm}\}$, $n=1, N, m=1, M$. In this case, the elements of the matrix determine the need for the m -type task in the n -th resource. The availability of resources is determined by the vector $\sigma = \{\sigma_n\}$, $n=1, N$.

The general scheme of functioning of such a system is presented in Fig. 1.

The physical content of the elements r_{nm} consists in the fact that the elements are binary, including time τ_{nm} , which

is spent by the resource to obtain the corresponding result at a predetermined stage of complex processing, and the degree χ_{nm} of contribution of the resource to the overall processing process:

$$r_{nm} = \{r_{nm}, \chi_{nm}\}; 0 \leq \chi_{nm} \leq 1; \sum_{n=1}^N \chi_{nm} = 1. \tag{4}$$

The introduction of the parameter χ_{nm} is justified because it eliminates the optional n -th resource from the process of solving the m -th problem in the case of a heavy load of the SCP GSD, when the process requires considerable time and its contribution is small.

Parameter τ_{nm} is determined according to the following formula:

$$\tau_{nm} = \tau_{nm1} + \tau_{nm2} + \tau_{nm3} + \tau_{nm4}, \tag{5}$$

where τ_{nm1} is the time of transfer of the n -th resource while solving the m -th problem of the output from the previous resource (taking into account the factor of their territorial distribution); τ_{nm2} is the time to search for the data that may be needed for the n -th resource at this stage of the solution of the m -th problem, in the databases of all known resources $\{r_n\}$ (also taking into account the factor of their territorial distribution and heterogeneity of the structures of the data processed in them); τ_{nm3} is the time, which is spent directly by the n -th resource for the procedure of performing the current stage of solving the m -th problem; τ_{nm4} is the time spent by the n -th resource to transfer the results of their work to the next resource in the chain.

Practice shows that these components τ_{nmi} in formula (5) have different "specific gravity". Time depends both on the degree of territorial distribution of the resources involved in solving the problem and on the level of automation of the subsystems that provide the retrieval and search of the necessary input data. Finding the information resource, you need in thematic databases involves not automatic, but automated or manual operation. In this case, "overhead", not directly related to the solution of the problem (times $\tau_{nm1}, \tau_{nm2}, \tau_{nm4}$) will be at least 50 % of all time spent by the resource to obtain the result at a given stage of complex processing (time τ_{nm}).

It should be noted that in order to calculate the time spent in general on solving the problem of responding to an event of the m -th type, the time τ_{nm} should be summed up. This is due to the fact that there is always a delay between the time when the event occurred and the time when the response of the SCP GSD began. Also, when resources are scarce, there is a waiting time for the availability of the required resource (waiting time for it).

The time, which is taken by the system to respond to the i -th event of the m -th type of MO can be determined by the formula

$$\tau_{m_i} = t_{i_1} + t_{i_2} + \sum_{n=1}^N (\omega_{nm} (\tau_{nm} + T_n)), \tag{6}$$

where t_{i_1} is the time of the start of the reaction to the i -th event; t_{i_2} is the time, which is taken to bring results to interested consumers; T_n is the waiting time for n -th resource availability (queuing);

$$\omega_{nm} = \begin{cases} 0, \chi_{nm} \leq X; \\ 1, \chi_{nm} > X. \end{cases} \tag{7}$$

In addition, it should be noted that time t_i and T_n , in the general case, are not constants and depend on the loading of the SCP GSD.

The physical meaning of the parameter ω_{nm} consists in the following. If the value of χ_{nm} is not more than some configured global configuration parameter X , the n -th resource is excluded from the process of solving the m -th problem, and the time τ_{nm} is not taken into account while calculating the total time spent to solve it.

The physical meaning of the parameter t_i consists in the fact that there is always a difference between the time when the event at the MO occurred and the moment when it was recorded by monitoring means and transmitted to the SCP GSD. This is due to the fact that, on the one hand, the monitoring of the facilities by a given object is not continuous, but discrete. On the other hand, the transfer of data from the points (devices) of extraction to the points of their reception occurs not immediately after fixation of a given monitoring object, but according to a given discrete regulation.

The parameter T_n is defined as follows:

$$T_n = \varepsilon_n \cdot F(n), \tag{8}$$

where $F(n)$ is the procedure that determines the waiting time for the n -type resource to be available; ε_n determines whether there is a queue for the n -type resource.

The value for ε_n is set based on the analysis of the value of the corresponding element of the vector $\sigma = \{\sigma_n\}$, $n=1, N$, which is determined on the basis of such considerations.

The resources in the SCP GSD are divided into two classes. The first class includes resources, which are related to the participation of the human operator in the processing, providing automated or manual operation. For such resources, the value firstly is $\sigma > 0$. The value decreases by 1 when the n -type resource is assigned to solve a particular task and increases by 1 when the n -type resource completes its next task. If $\sigma = 0$, such a resource is considered unavailable.

The second class includes resources that operate without the participation of a human operator, automatic mode. It is believed that sometimes waiting for their availability can be neglected and capable of simultaneously serving solutions to many tasks at once. For such resources, $\sigma = -1$. Typical examples of such resources may be different thematic databases, as well as computing services.

Thereby, ε_n is given as follows:

$$\varepsilon_n = \begin{cases} 1, & \sigma_n = 0, \\ 0, & \sigma_n \neq 0. \end{cases} \tag{9}$$

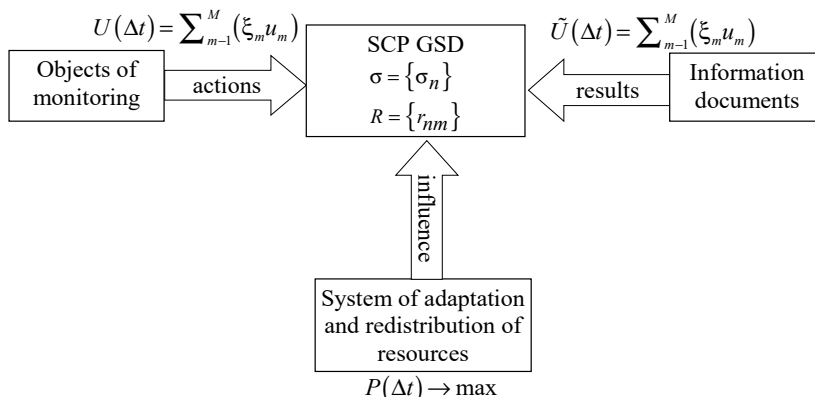


Fig. 1. Generalized scheme of SCP GSD functioning

5. Development of an algorithm for complex processing of geospatial data in special-purpose geoinformation systems

On the basis of the mathematical description of the functioning of the system of complex processing of geospatial data, it is possible to develop an algorithm of complex processing of geospatial data in geoinformation systems of special-purpose in the conditions of diversity and uncertainty of the data. The proposed algorithm is shown in Fig. 2.

1. Input of initial data (step 1 in the algorithm diagram).

The input of initial data (operational situation) about the monitoring object is made. The degree of uncertainty of the intelligence object (monitoring object) is determined.

The uncertainty is due to both the insufficient amount of information required to quantify the processes occurring in the system and the complexity of the monitoring object itself as a complex system.

Practice shows that most often there are a number of problems related, in particular, to the implementation of automated decision support systems for determining the status of the monitoring object in different conditions of a priori uncertainty. In other words, there is a contradiction between the current level of technical support for the processing of different types of geospatial information and the level of mathematical, software, and analytical support for decision-making (DM) regarding the state of the monitoring object in different uncertainty conditions.

In the conditions of certainty, the decision on the condition of the monitoring object with the best value of the quality criterion is generally chosen. Often, these conditions and decisions are called deterministic. In this case, it is assumed that the implementation of the decision on the status of the monitoring object will lead to a previously known, moreover, single result, which does not have any ambiguity about the prospects of the future status. In these circumstances, full awareness of current and future changes in the external and internal environment dominates. In its turn, it guarantees the achievement of the calculated (planned) quality indicators and the efficiency of determining the status of the monitoring object or its elements.

In practice, the conditions of full determination in the decision-making regarding the condition of the monitoring object are extremely rare, but this coarsening is occasionally allowed to substantiate the advantages and simplicity of decision-finding algorithms [21–25].

In the context of risk, the action taken within the framework of both the external and the internal environment of the operation of the monitoring object is taken into account, as a result of which the actual decision on the condition of the monitoring facility in the real conditions may not coincide with the calculated unambiguous value. That is why in risk conditions it is accepted that any decision to determine the status of a monitoring object can lead to many possible outcomes, the probabilities of which can be estimated or known in advance. In practice, this leads to the fact that the algorithm for choosing the decision on the condition of the monitoring object is based on the probability of obtaining certain results and their usefulness.

The degree of information uncertainty varies from complete ignorance of the future situation to the possible knowledge of

the lower and upper boundaries within which the indicator characterizing the status of the monitoring object may vary.

It is customary to distinguish between degrees [21–25]:

– full awareness, it is the predictability of G_τ events close to the unit:

$$\lim_{\tau \rightarrow \tau_k} G_\tau = 1, \tag{10}$$

where τ is the time, τ_k is the event prediction end time;

– complete uncertainty is the predictability of the event close to zero, which is mathematically expressed by the relation:

$$\lim_{\tau \rightarrow \tau_k} G_\tau = 0; \tag{11}$$

– partial uncertainty corresponds to the predictability of events lying in the range from 0 to 1, which can be expressed by inequality:

$$0 < G_\tau < 1; \tag{12}$$

– “hopeless” uncertainty corresponds to the lack of information about the state of the environment within which the monitoring object operates.

In these cases, the criteria of a guaranteed result, optimism, pessimism, minimum risk and generalized maxima are used to predict the status of the monitoring object for determining the best decisions [21–25].

The uncertainties that arise while making decisions about the state of the monitoring object are objectively due to the fact that complex systems of information collection, storage, processing and transmission in the course of their functioning depend on a number of factors.

It is important to take into account the factor of time in the analysis of the quality of the decisions made because always the quality of the decisions and the costs of their implementation are distributed over time [21–25].

It is known that differently distributed over time, but equal in magnitude costs give unequal useful result (the quality of the decisions regarding the state of the intelligence object (monitoring object)). That is why retrospective, current and prospective uncertainties are commonly distinguished.

In uncertain conditions, which are mostly characteristic for the environment of the monitoring object, the characteristics of the system may also be random, but unlike probabilistic conditions, the law of their distribution is unknown.

Changes in the parameters of the decision-making process can be explained by a number of factors [21–25]:

1. Changes are made to the baseline data, which are used to make the decision to organize collection, storage, processing and transmission of information on the status of the monitoring object.

2. Changes in external conditions and current requirements (quality criteria) of the DM for the organization of collection, storage, processing and transmission of information about the state of the monitoring object (cause – environment, changing of the management tasks).

3. Changing of decision-making methods for organizing collection, storage, processing and transmission of information on the status of the monitoring object.

Thus, there is objective uncertainty about the conditions for deciding on the status of the monitoring object. This is caused by the non-stationary functioning of the systems, effects of destabilizing factors, external environment, fuzzy goals and inconsistency of management tasks, which necessi-

tates the adoption not only of the multicriteria condition but also of the uncertainty.

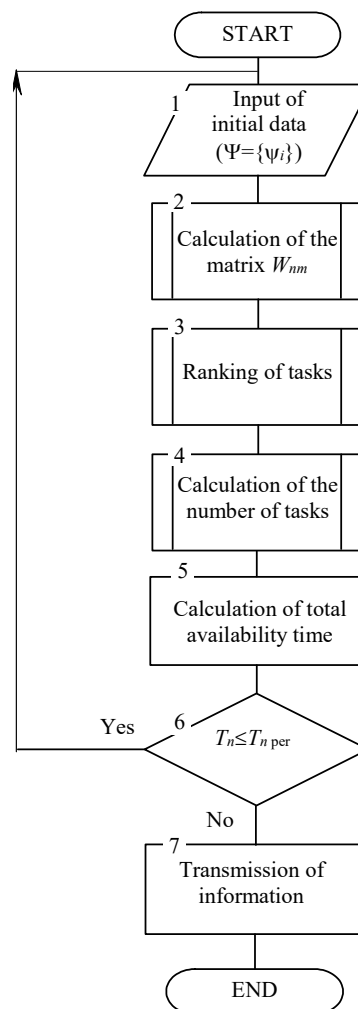


Fig. 2. Algorithm of complex processing of geospatial data in special-purpose geoinformation systems in conditions of diversity and uncertainty of data

Therefore, the task of evaluating the monitoring object is formulated as follows: it is necessary to have a vector of intelligence features, (x_1, \dots, x_n) to evaluate the condition of the monitoring object (make a decision D) for the admissible time with maximum certainty, and to present in a formalized form

$$D = \{D_j\} \rightarrow D_j^0 : \max P(x_1, \dots, x_n) / t_A \geq t_{AD}; \tag{13}$$

$$x_i \in X, i = \overline{1, n},$$

where $P(x_1, \dots, x_n)$ is the likelihood of evaluating the monitoring object correctly; x_i is the set of intelligence indicators of a certain set X ; $j = \overline{1, m}$ are the options for assessing the status of the monitoring object; t_A, t_{AD} is the time and allowable analysis time, respectively.

2. Calculation of the matrix W_{nm} (action 2 on the diagram of the algorithm).

There is a calculation for solving the problems of what type is assigned to each of the available n -type resources (matrix W_{nm} calculation)

$$W_{nm} = \{r_{nm}^{k_i}, \Delta\tau^{k_i}\}, \tag{14}$$

where $r_{nm}^{k_n}$ is the R -th resource of the n -th type, $R_n = 1$, K_n is the number of n -type resources; $\Delta\tau^{k_n}$ is the time that has elapsed since the start of solving the current task by the R_n -th resource.

Also, at that stage, the order of the autoregressive model and the volume of the training sample for forecasting are determined. To predict the status of the monitoring object, the method of forecasting the condition of the monitoring object, which was developed in [26] is used.

3. Ranking of tasks (action 3 on the scheme of the algorithm).

Based on the analysis of the setting regulations (r_{nm}), they are ranked for the estimated waiting time of their availability and the minimum waiting time for the release of any n -type resource is calculated:

$$T_n' = \min_{1 \leq k_n \leq K_n} (\tau_{nm} - \Delta\tau^{k_n}). \quad (15)$$

4. Calculation of the number of tasks (action 4 on the scheme of the algorithm).

At this stage, the algorithm calculates the number of S_i tasks that are already in the queue for the n -type resource, and calculates the time, which is required to complete this queue:

$$T_n'' = \sum_i \tau_{nm}^i. \quad (16)$$

During the implementation of steps 3 and 4 of the specified algorithm, the order of the autoregressive model is corrected, the variance is estimated and the magnitude of information mismatch is calculated. This is done to predict the status of the monitoring object using the developed forecasting method [26].

5. Determination of the total time of waiting for the resource availability (action 5 on the diagram of the algorithm).

In step 5 of the algorithm, the total waiting time T_n for the availability of the n -type resource is calculated:

$$T_n = T_n' + T_n''. \quad (17)$$

These models allow you to estimate the time spent on the main stages of the operation of the SCP GSD.

According to the purpose, it is advisable to take the probability of system tasks as an indicator of the efficiency of the process of complex processing of geospatial data in geoinformation systems. The basis for efficiency evaluation is the probability of fulfilling the GIS task during the control cycle.

In the general case, the control cycle includes the following steps:

1. Collection of information about tasks, the status of facilities and objects, for which the system is used, the results of tasks, the status of management and communication facilities.

2. Processing of information about item 1, entering databases, situation assessment. Planning and operational management of actions, use of common funds, management system tools and measures for their provision.

3. Preparation of decision options, calculations and situation assessment. Planning and operational management of actions, use of general facilities, management system means, measures for their provision.

4. Decision making and approval.

5. Generation of signals, commands, orders and control effects.

6. Bringing signals, commands, orders, control effects to management objects.

7. Acceptance of management influences and preparation of devices for the performance of tasks.

The first stage is the collection of information. This process is characterized by the rate of receipt of information λ , which is measured by the number (such as words, bytes or bits per unit time). The speed λ is determined by the intensity of situation change. Another feature of the stage is the amount of information V that must be collected and transmitted accordingly for a sufficiently complete assessment of the situation and decision. The average duration of the stage is determined by:

$$\tau_c = \max \left\{ \frac{V\alpha}{\mu}, \frac{V\alpha}{\lambda} \right\}, \quad (18)$$

where α is the redundancy factor, which is needed to encode and enhance the information; μ is the speed of information transfer. Given the random nature of the transfer and reception process, the probability of termination of the process after time τ_c is defined as follows:

$$P = 1 - e^{-\frac{\tau_c}{\tau_0}}. \quad (19)$$

The second stage is information processing. At this stage, information problems are solved. The average solution time τ_z is determined by the average number of machine operations N_i that must be performed for the information processing and PC performance:

$$\tau_z = \frac{N_i}{\nu}. \quad (20)$$

The N_i value is proportional to the volume V of transmitted initial data. The proportionality factor C_i is determined by the average number of machine operations transmitted per unit of initial data.

The probability of completing the stage over time τ_i is equal to:

$$P = 1 - e^{-\frac{\tau_i}{\tau_i}}. \quad (21)$$

The third stage is the preparation of variants of the decision. At this stage, the calculation and optimization problems are solved. However, the latter requires a much larger number of N_p machine operations, which are usually proportional to the cube or square of the initial data volume at this stage.

$$N_p = C_p (V\gamma)^3, \quad (22)$$

where γ is the depth of V data use factor; C_p is the average number of machine operations, which are performed per cube of unit volume of initial data.

The average time to solve the problems of this stage is mainly determined by optimization problems:

$$\theta = \frac{1}{\nu} C_p V_p^3. \quad (23)$$

The probability of completing a task over time τ_p is equal to:

$$P = 1 - e^{-\frac{\tau_p}{\theta}}. \quad (24)$$

The fourth stage consists in making a decision and approving its decision-making object (DMO) at the appropriate level regarding the status of the intelligence object (monitoring object). The average duration of the given stage τ_z depends on the volume V_z of DMO information that is determined by the volume V_p .

$$V_z = \ell V_p, \tag{25}$$

where $\ell \ll 1$ is the degree of aggregation of the information provided by the DMO.

If u – the average speed of solving the problem over time τ_z is equal to:

$$\bar{\tau}_z = \frac{1}{u} \ell V_p. \tag{26}$$

The probability of completing a task over time τ_r is equal to:

$$P = 1 - e^{-\frac{\tau_r}{\tau_0}}. \tag{27}$$

The fifth stage is the formation of management influences. At that stage, the design tasks are solved to confirm the decisions, which are taken to predict the future state of the monitoring object. Average stage duration:

$$\bar{\tau}_\phi = \frac{1}{v} N_\phi, \tag{28}$$

where $N_\phi = v V_r$ is the number of machine operations performed in the process of forming control effects, where v is the number of computational operations per unit of transformation V_g respectively

$$\bar{\tau}_\phi = \frac{1}{v} v V_g. \tag{29}$$

The probability of completing the stage over time τ_ϕ is equal to:

$$P = 1 - e^{-\frac{\tau_\phi}{\tau_\phi}}. \tag{30}$$

The sixth step consists in bringing the management effects to the extraction object. Its average duration τ_d is calculated similarly to the first stage:

$$\bar{\tau}_d = \frac{1}{\omega} V_d, \tag{31}$$

where $V_d = \eta V_p$ is the amount of information that comes to the extraction facility; η is the factor that takes into account the proportion of information V_p brought to one tool. In a particular case, it may be equal to

$$\eta = \frac{1}{n},$$

where n is the number of devices.

The probability of completing the stage over time τ_d is equal to:

$$P = 1 - e^{-\frac{\tau_d}{\tau_d}}. \tag{32}$$

The seventh stage is the preparation of the tool for the tasks. The average duration depends on the characteristics of the preparation process, such as input of initial data and preparation of the control object for action:

$$\bar{\tau}_n = \tau_1 + \tau_2, \tag{33}$$

where τ_1 is the average input time of initial data; τ_2 is the average preparation time.

The probability of completing the stage over time $\bar{\tau}_n$ is equal to:

$$P = 1 - e^{-\frac{\tau_n}{\tau_n}}. \tag{34}$$

The control cycle does not include the stage of executing the control cycle directly. Provided the independence of the solution of problems in stages, the overall probability of executing the control cycle is determined by multiplying the probabilities of completing all the steps.

Therefore, the amount of information, which was received in the first stage of the management cycle depends on the average duration of the remaining stages.

6. Discussion of results on the development of the algorithm of complex processing of geospatial data in geoinformation systems

The algorithm of complex processing of geospatial data in special-purpose geoinformation systems in the conditions of diversity and uncertainty of the data is proposed. The proposed algorithm is simulated in MathCad 14.

The scheme of solving information and calculation problems using the proposed algorithm is shown in Fig. 3.

Fig. 3 shows the procedure for calculations using the proposed algorithm in the system of complex processing of geospatial data.

Taking into account the developed algorithm of SCP GSD, approaches to solve the problem of rational allocation of its resources can be formulated as follows:

- firstly, the minimization of the total time, which is required to solve the next problem of the m -th type;
- secondly, the redistribution of the SCP GSD resources in terms of their scarcity to solve problems, but responsiveness to events considering their relative importance.

As part of the first approach, it should be noted that the mathematical description of the functioning of the existing SCP GSD has a pronounced reactive nature. This consists in the fact that the activity of its resources is either stationary (planned) or is the result of the reaction to the events that occurred at the monitoring objects (and if they were managed to be recorded by observation devices). These mathematical relations do not take into account the nature of the development of situations in the space of monitoring objects. The fact that the SCP GSD operates under such a model leads to the fact that the SCP GSD will always be late in the development of the situation and be in the role of catching up.

Thus, the first approach is to intellectualize the process of functioning of the SCP GSD, which is based on the application of knowledge bases of intelligent systems, namely the transition from reactive models and monitoring methods to proactive ones that will be able:

- firstly, to anticipate possible variants of the situation development in the space of the monitoring objects, including in the conditions of uncertainty of their data;
- secondly, to limit the degree of human involvement in the management cycle of complex processing resources, as well as to automatically offer options for the formation of scenarios for solving monitoring tasks.

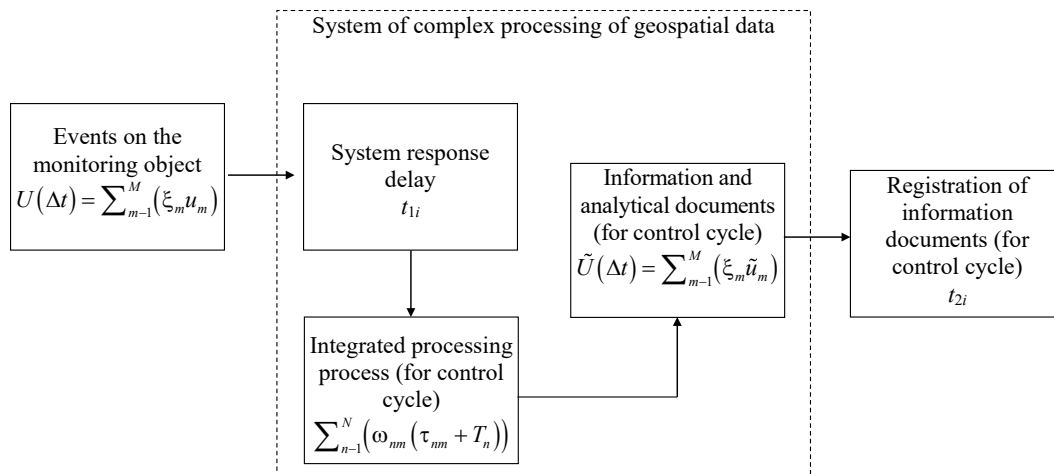


Fig. 3. Scheme of SCP GSD using the proposed algorithm

In the second approach, it is necessary to take into account the fact that an event that has a small value of the coefficient of relative importance, as a result, can cause events with a high value of this coefficient. Thus, taking into account the provisions of the first approach, it is necessary to dynamically change (increase) the coefficient of relative importance of such events in the process of functioning of the SCP GSD.

Fig. 4 shows graphical estimates of the efficiency of the proposed algorithm.

For this purpose, a simulation of the algorithm was performed to determine the number of computational operations, which are required for the operation of the algorithm. Initial data to evaluate the condition of the monitoring object using the proposed algorithm:

- number of sources of information about the status of the monitoring object - 3 (radio monitoring, remote sensing and unmanned aerial vehicles);
- number of information signs, which determine the status of the monitoring object - 13 (affiliation, type of organizational and staff formation, priority, minimum width on the front, maximum width on the front, number of personnel, minimum depth on the flank, maximum depth on the flank, the total number of personnel, the number of samples of weapons and military equipment, the number of types of samples of weapons and military equipment, the number of communication devices, type of communication devices);
- variants of organizational and staff formations - company, battalion, brigade.

Comparison by criterion is the minimum computational complexity with restrictions on the reliability of the obtained estimate.

These graphical dependencies show that the proposed algorithm allows to increase the speed of information processing (reduce the number of computing operations) in special-purpose geoinformation systems from 16 to 20 %, depending on the amount of information about the monitoring object.

This gain is explained by a complex combination not only of different types of data on the status of the monitoring object but also of the procedure of redistribution of computational operations among system resources.

The aforementioned research is a further development of the authors' research, which is aimed at developing methodological foundations for improving the efficiency of data processing in special-purpose geoinformation systems, published previously [26, 27]. The areas of further research should be

aimed at reducing the computational cost of processing different types of data in special-purpose geoinformation systems.

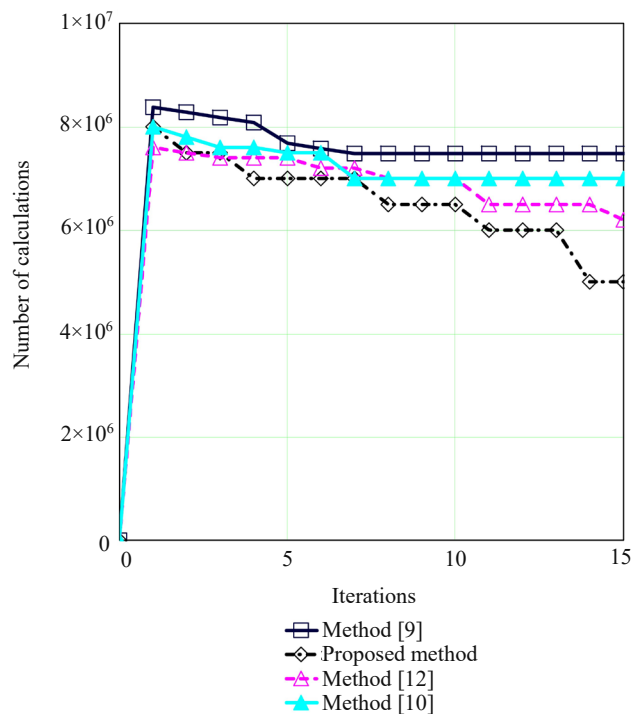


Fig. 4. Results of the evaluation of the efficiency of the proposed algorithm

7. Conclusions

1. In the course of the research, the problem of complex processing of geospatial data was formulated.

During formalization of the task of complex processing of geospatial data, it is found that complex processing can be solved by the simultaneous use of the following approaches:

- allocation of SCP GSD resources to solve the problems of response to events of different types, taking into account their relative importance.
- minimizing the total time, which is necessary to complete another *m*-type response.

2. The algorithm of complex processing of geospatial data in special-purpose geoinformation systems in the conditions of diversity and uncertainty of data is developed.

This algorithm allows to increase the speed of processing of different types of data in geoinformation systems due to the complex processing of geospatial data circulating in it. Complex processing consists not only in the possibility of working with different types of information but also in the redistribution of computational operations among the system resources.

The difference of the proposed algorithm, which determines its novelty, is as follows:

- allows high-quality processing of large arrays of various types of data, both numerical and quantitative;
- has less computational complexity due to the redistribution of computing operations;
- takes into account the degree of awareness of the condition of the monitoring object;
- allows complex processing and distribution of information about the status of the monitoring object.

The proposed algorithm improves the processing speed of information in special-purpose geoinformation systems from 16 to 20 %, depending on the amount of information about the monitoring object.

The practical significance of the developed algorithm consists in the fact that the proposed algorithm will significantly increase the efficiency of systems of complex processing of geospatial data in geoinformation systems.

The decision is formulated taking into account the results of forecasting the condition of the monitoring object on the basis of the forecasting method developed in [26].

3. The advantages of this algorithm include:

- the ability to rationally allocate the SCP GSD resources to solve the problems of response to events of different types, taking into account their relative importance;

- minimization of the total time, which is necessary to complete the task of responding to another m -type circumstance;

- the ability to predict the status of the monitoring object;
- limitation of the degree of human participation in the cycle of resource management of complex processing, as well as automatic determination of the variant of forming scenarios for solving the monitoring problems.

The disadvantages of this algorithm include:

- the need to process large amounts of data to determine the status of the monitoring object;
- lack of mechanisms for rejecting (sorting) unnecessary information during the object monitoring session.

Acknowledgments

The author team is grateful for the help:

- to the honored worker of science and technology of Ukraine, doctor of technical sciences, professor Slusar Vadim Ivanovich – chief scientific worker of the Central Research Institute of Armament and Military Equipment of the Armed Forces of Ukraine;
- to the doctor of technical sciences, senior scientific worker Sova Oleg Yaroslavovich – head of the Automated Control Systems Department of the Military Institute of Telecommunications and Information Technologies named after Heroes of Kruty;
- to the doctor of technical sciences, professor Rotshtein Alexander Petrovich – professor of the Jerusalem Polytechnic Institute – Mahon Lev.

References

1. Makridenko, L. A., Volkov, S. N., Khodnenko, V. P., Zolotoy, S. A. (2010). Conceptual problems on creation and application of small spacecraft. *Voprosy elektromekhaniki*, 114 (1), 15–26.
2. Shyshatskyi, A. V., Bashkyrov, O. M., Kostyna, O. M. (2015). Rozvytok intehrovanykh system zviazku ta peredachi danykh dlia potreb Zbroinykh Syl. *Ozbroiennia ta viyskova tekhnika*, 1 (5), 35–39.
3. Trotsenko, R. V., Bolotov, M. V. (2014). Data extraction process for heterogeneous sources. *Privolzhskiy nauchnyy vestnik*, 12-1 (40), 52–54.
4. Bodyanskiy, E., Strukov, V., Uzlov, D. (2017). Generalized metrics in the problem of analysis of multidimensional data with different scales. *Zbirnyk naukovykh prats Kharkivskoho natsionalnoho universytetu Povitrianykh Syl*, 3 (52), 98–101.
5. Noh, B., Son, J., Park, H., Chang, S. (2017). In-Depth Analysis of Energy Efficiency Related Factors in Commercial Buildings Using Data Cube and Association Rule Mining. *Sustainability*, 9 (11), 2119. doi: <https://doi.org/10.3390/su9112119>
6. Petras, V., Petrasova, A., Jeziorska, J., Mitasova, H. (2016). Processing UAV and lidar point clouds in Grass GIS. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B7, 945–952. doi: <https://doi.org/10.5194/isprs-archives-xli-b7-945-2016>
7. Polovina, S., Radic, B., Ristic, R., Milcanovic, V. (2016). Spatial and temporal analysis of natural resources degradation in the Likodra River watershed. *Glasnik Sumarskog Fakulteta*, 114, 169–188. doi: <https://doi.org/10.2298/gsf1614169p>
8. Poryadin, I., Smirnova, E. (2017). Binary Classification Method of Social Network Users. *Science and Education of the Bauman MSTU*, 17 (02), 121–137. doi: <https://doi.org/10.7463/0217.0000915>
9. Tymchuk, S. (2017). Methods of Complex Data Processing from Technical Means of Monitoring. *Traektorii Nauki. Path of Science*, 3 (3), 4.1–4.9. doi: <https://doi.org/10.22178/pos.20-4>
10. Semenov, V. V., Lebedev, I. S. (2019). Processing of signal information in problems of monitoring information security of unmanned autonomous objects. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 19 (3), 492–498. doi: <https://doi.org/10.17586/2226-1494-2019-19-3-492-498>
11. Zhou, S., Yin, Z., Wu, Z., Chen, Y., Zhao, N., Yang, Z. (2019). A robust modulation classification method using convolutional neural networks. *EURASIP Journal on Advances in Signal Processing*, 2019 (1). doi: <https://doi.org/10.1186/s13634-019-0616-6>
12. Zhang, D., Ding, W., Zhang, B., Xie, C., Li, H., Liu, C., Han, J. (2018). Automatic Modulation Classification Based on Deep Learning for Unmanned Aerial Vehicles. *Sensors*, 18 (3), 924. doi: <https://doi.org/10.3390/s18030924>

13. Kukulska-Kozieł, A., Salata, T., Cegielska, K., Szylar, M. (2018). Methodology of evaluation and correction of geometric data topology in QGIS software. *Acta Scientiarum Polonorum. Formatio Circumiectus*, 17 (1), 137–150.
14. Rulev, A., Yuferev, V. (2015). Theory of geoinformatic mapping of erosion geomorphological systems. *Vestnik Volgogradskogo Gosudarstvennogo Universiteta. Serija 11. Estestvennye Nauki*, 4, 62–67. doi: <https://doi.org/10.15688/jvolsu11.2015.4.7>
15. Yousefi, M., Kreuzer, O. P., Nykänen, V., Hronsky, J. M. A. (2019). Exploration information systems – A proposal for the future use of GIS in mineral exploration targeting. *Ore Geology Reviews*, 111, 103005. doi: <https://doi.org/10.1016/j.oregeorev.2019.103005>
16. Ashkezari, A. D., Hosseinzadeh, N., Chebli, A., Albadi, M. (2018). Development of an enterprise Geographic Information System (GIS) integrated with smart grid. *Sustainable Energy, Grids and Networks*, 14, 25–34. doi: <https://doi.org/10.1016/j.segan.2018.02.001>
17. Wang, S., Zhong, Y., Wang, E. (2019). An integrated GIS platform architecture for spatiotemporal big data. *Future Generation Computer Systems*, 94, 160–172. doi: <https://doi.org/10.1016/j.future.2018.10.034>
18. Wan-Mohamad, W. N. S., Abdul-Ghani, A. N. (2011). The Use of Geographic Information System (GIS) for Geotechnical Data Processing and Presentation. *Procedia Engineering*, 20, 397–406. doi: <https://doi.org/10.1016/j.proeng.2011.11.182>
19. Pedro, J., Silva, C., Pinheiro, M. D. (2019). Integrating GIS spatial dimension into BREEAM communities sustainability assessment to support urban planning policies, Lisbon case study. *Land Use Policy*, 83, 424–434. doi: <https://doi.org/10.1016/j.landusepol.2019.02.003>
20. Mokhtara, C., Negrou, B., Settou, N., Gouareh, A., Settou, B. (2019). Pathways to plus-energy buildings in Algeria: design optimization method based on GIS and multi-criteria decision-making. *Energy Procedia*, 162, 171–180. doi: <https://doi.org/10.1016/j.egypro.2019.04.019>
21. Karin, S. A. (2012). Integration in the single information space of heterogeneous geospatial data. *Informatsionno-upravlyayushchie sistemy*, 2, 89–94.
22. Karin, S. A. (2014). Developing a domain-specific ontology in spatial data processing systems. *Informatsionno-upravlyayushchie sistemy*, 4, 78–84.
23. Belousov, S. M. (2006). Matematicheskaya model' mnogopotchnoy sistemy massovogo obsluzhivaniya, upravlyaemoy planirovshchikom resursov. *Vestnik Novosibirskogo gosudarstvennogo universiteta. Seriya: Informatsionnye tehnologii*, 4 (1), 14–26.
24. Karin, S., Dudin, E. (2014). Methods for creating distribute systems of geospatial data collection, storage, and search. *Information & Cosmos*, 3, 48–53.
25. Gatsenko, S. S., Kal'nits'kiy, Yu. M., Gel'veychuk, O. M. (2014). The problem of distribution of information flows in automated control systems of troops (forces) of Armed Forces of Ukraine. *Zbirnyk naukovykh prats Tsentru voienno-stratehichnykh doslidzhen Natsionalnoho universytetu oborony Ukrainy imeni Ivana Cherniakhovskoho*, 2 (51), 107–111.
26. Koshlan, O. (2018). Method of parameters forecasting for the object of intelligence in geoinformation systems of the special purpose. XI nauково-praktychna konferentsiya "Priorytetni napriamky rozvytku telekomunikatsiynykh system ta merezh spetsialnoho pryznachennia. Zastosuvannia pidrozdiliv, kompleksiv, zasobiv zviazku ta avtomatyzatsiyi v operatsiyi Obiednanykh syl". Kyiv, 125–126.
27. Koshlan, A. (2018). Conceptual model of a specialized geoinformation system. *Advanced Information Systems*, 2 (3), 36–40. doi: <https://doi.org/10.20998/2522-9052.2018.3.06>