Design and application of a specialized information system for monitoring and decision support

S V Smirnov
Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia

E-mail: sapr2006@bk.ru

Abstract. The design is considered, involving the use of instrumental and analytical apparatus of GIS (geographic information systems) in order to provide monitoring and support for decision-making by government authorities. For these purposes, the basic models and methods for solving problems that arise before the administration of governing bodies have been analyzed. The design of a graphical-analytical model of a thematic layer of an electronic GIS map using a quadratomic tree is also described. A graph of conceptual-meaningful features and a graph diagram of the formation of graphic means for the design of signs are considered. The advantages and disadvantages of the quadratic tree model are described. In addition, the assessment of the temporal and capacitive complexity of information extraction, which is the main component of clarifying a clear algorithm for setting the problem during the functioning of the system being developed, is analyzed.

1. Introduction
One of the solutions to the problems of information deficiency in the modern world for a wide range of actions in the management of territories can be the creation of reference information systems, endowed with the rights of advisor systems. This would be, for example, a specialized geographic information system (GIS) designed to provide management assistance. This system ensures the solution of problems, in particular, in the educational sphere of the region (district, district) with a geographic reference to the area.

The formulation and solution of tasks facing government bodies is difficult, even in the presence of a multifunctional specialized information system. These tasks must be translated into the language of the system itself, and for this it is necessary to investigate them. To conduct a study of problems of this type, it is necessary to model them.

The modeling process includes the following main functions: description, explanation and prediction of the behavior of a real system. The main typical goals of modeling tasks for an administrative employee will be the following areas: search for optimal or similar solutions, assessment of the effectiveness of the decision-making process, as well as monitoring and forecasting the development of management objects.

In addition, when creating an electronic map (GIS), an important place is occupied by the design of its design. The correct design of the card determines its readability and clarity, as well as aesthetic advantages. Designing a card design is a complex creative process that requires a lot of knowledge and experience, constant searches and experiments. In practice, it is necessary to take into account many requirements for the design of any electronic map, to know the specifics of the design of different types
of GIS, the advantages and disadvantages of visual means. Often, the value of its content is determined through the appearance of the designed system. Achievement of readability and visibility, as well as the correct semantic perception of the electronic map is achieved with the correct design of the graph-analytical model of the thematic layer of the map [1].

Also, when designing a GIS, it is necessary to find out a clear algorithm for solving the problems described above, the main parameters of which is an assessment of the temporal and capacitive complexity, which, despite the modern development of computing technology, can arise when graphical data (objects) are accumulated on a map, as well as measures to minimize them. All of the above will be investigated further.

2. Information models for solving problems in the sphere of control

Since there are many tasks in the management area, it would be advisable to create a common model. The generated model is an object that is similar in some respects to the prototype and serves as a means of describing, explaining and predicting. The advantages of creating a model are that a simplified image is presented that reflects the properties of the prototype that are essential for research.

Let us formulate the basic models with the help of which the tasks are solved. These will be the following models: optimal choice of the management object, decision-making by the administration, object control, development forecasting and operational management. These types of models constitute an analytical apparatus that forms an expert part for a specialized geographic information system (GIS), with the help of which these tasks are solved [2].

The decision-making process as a function of transforming the content of information is formalized in terms of decision-making theory. The main concepts in the theory of decision making will be the following designations:

- $F$ - a universal set of options, alternatives, plans, from which a choice is made;
- $\theta$ - presentation, many options presented for selection, ($\theta \subseteq F$);
- $\zeta$ - many selected options;
- $H$ - selection principle (selection function), $\zeta = H(\theta)$.

1) **Optimal object selection model.** The set of options $\{F\}$ is defined, and the selection principle $H$ is strictly formalized. To solve problems using this model, analytical research methods are applied, which are used for the optimal choice.

2) **Decision-making model by the administration of governing bodies.** Many options $\{F\}$ can be supplemented and modified, and the principle of choice is not formalized. The problem solved using this model allows you to change the solution when a new option is found. The general decision problem model is as follows:

$$\gamma = \{\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6\}$$  \hspace{1cm} (1)

$\gamma$ – is the set of the model of the general decision-making problem; $\varphi_1$ – is the goal of making a decision, $\varphi_2$ - is the initial data for the formation of options, $\varphi_3$ – is the set of formed options, $\varphi_4$ – is the chosen option, $\varphi_5$ – is the rule for the formation of options, and $\varphi_6$ – is the rule for choosing the best (optimal option).

3) **Object management control model.** Particular tasks will be: observation and classification of control objects. To establish control, it is necessary to ensure the potential observation of the internal states of the control object according to the characteristics of the external environment. The solution to the object observation problem is to find such a mapping (2), which puts each observed realization of the output characteristics $D$ in a one-to-one correspondence with the internal state of the control object $\lambda$: 

$$\rho: D \rightarrow \lambda$$  \hspace{1cm} (2)
The solution to the object classification problem consists in finding such a mapping (3) that provides a partition of the entire set of possible realizations of the output characteristics D into a limited number of classes \( \lambda_2 \) with general properties:

\[
\lambda_1: \Gamma \rightarrow \lambda_2
\]  

(3)

The previously defined general properties are a kind of standards for recognizing the real states of an object in the process of its control.

4) **A model for forecasting the development of the management sphere.** Solving problems using this model is based on assumptions about future qualitative changes in the system and the preservation of the existing patterns of development. Problems are solved using expert methods, when it comes to long-term forecasts, as well as extrapolation methods, if the forecasts are short-term.

5) **Model of operational management of the social and educational sphere.** Solving problems using this model consists in performing a sequence of control actions \( \{ \nu_i \}, \ i=1,2,\ldots,m \), previously compiled and introduced into the system, set for the entire period of achieving the goal of making a decision \( \varphi_i \). These influences are given to the control object at time intervals \( \Delta t_i, \ i=1,2,\ldots,m \) without taking into account its state.

It is also advisable to solve this type of tasks using a specialized geographic information system (GIS), by means of an automated process for extracting the necessary graphic information.

The main technique for solving the set tasks is to select separate layers in the SGIS that represent homogeneous semantic information [3].

The second principle is to ensure the possibility of combining information located on several selected layers on a time layer used to solve a specific problem. It should be noted that the technology for solving problems facing the controls requires detailed study so that non-programmers do not have any difficulties in interpreting them.

3. **Construction of visual means of the analytical model of the SGIS thematic layer using a quadratonic tree**

When developing a geographic information system, an important place is occupied by the design of its design. The correct design of the card determines its readability and clarity, as well as aesthetic advantages. Designing a map is a complex creative process that requires a lot of knowledge and experience, constant research and experimentation. In practice, it is necessary to take into account many requirements for the design of any electronic map, to know the specifics of the design of different types of GIS, the advantages and disadvantages of visual means. Often, the value of its content is determined through the appearance of the designed system [4].

Designing a graphical-analytical model of a thematic layer of an electronic map for a geographic information system consists of the stages of developing signs for the transfer of spatial forms of certain qualitative and quantitative characteristics of objects, as well as establishing visual means in relation to the content, scale and purpose of maps.

It is most expedient to design signs for the transmission of spatial forms of certain qualitative and quantitative characteristics of objects using a machine-oriented formalized cartographic language (MFCL). For this, graph theory is used. The application of graph theory in the development of maps was first put forward and developed by the famous Russian cartographer A S Wasmuth [5].

The specificity of the MFCL is that, in addition to logical and semiotic aspects, it contains other elements that are focused on the possibility of processing them on a computer. Based on this, the means of interaction between the user and the geographic information system were developed. The language interop contains five operator classes. The first class contains operators for setting units of measurement and scaling; the second is the operators that form the geometric model; the third is the choice of the color of elements and objects or types of lines; the fourth - describes related calculations or geometric constructions and transformations; fifth - describes raster information presented in various data formats.
The graph of qualitative features of objects (point, linear and area) has a general form of a hierarchical tree built using graph theory (figure 1). Graph G includes subgraphs of conceptual levels. Each level is determined by a sign of commonality and the number of diversity of this sign: the number of vertices of the subgraph [6-7].

As shown in the figure, graph G at the first level includes several subgraphs c. Some subgraph c includes subgraphs d₁, d₂, ..., dₙ. The number of vertices of the subgraph c is equal to the sum of the vertices of the subgraphs dᵢ. In the subsequent division, the graph dᵢ has the following levels, which include the subgraphs qᵢ, qᵢ₋₁, ..., q₁. Moreover, dᵢ defines their commonality, and the signs of qi-differences. Based on this, a graph of visual means is drawn up.

The graph of visual means will have the form of a hierarchical tree, where at the vertices there will be parameters of signs that determine the signs of commonality and differences in the levels of characteristics of an object. In view of the fact that the objects of mapping are characterized by commonality, which is expressed in the unity of the structural organization of system-forming features, the types of graphs of qualitative features for different objects will differ in the following features: the number of levels, elements of diversity.

For all graphs of depicting means for the same (general) levels, a single means of depicting an object can be adopted. Imaging tools rely primarily on approaches that save the amount of storage required for raster data and help reduce the time it takes to access feature descriptions. One such approach is the quadtree method. The use of the quadratotic tree model as an object of development for graphic information systems is presented in the work [8].
In quadratomic structures, which divide the image into nested cells, precisely those requirements are achieved that are necessary for the convenience of the formation of pictorial means. This approach is used to represent raster and vector graphics data structures. There are examples where the quadtree approach has been used to represent thematic data in order to model the content of synthetic maps.

Consider this method in figure 2. The figure shows the representation of an object (a) in the form of nested square blocks (b) and a quadratomic tree derived from this (c). At each level of the tree, it is shown (with a light or dark square) whether or not the properties of this level belong to the object. Let us investigate why the quadtree model is suitable for development.

Figure 2. Representation of the object on the map.

The development of the necessary data model for the thematic layer is one of the most important tasks in the construction of the SIS. The main problems of processing, efficiency, storage and usability remain in connection with the presence of vector and raster data models.

The solution to these problems is achieved by creating mosaic hierarchical models. Mosaicity of models is obtained due to the division of the data plane into polygons, which can be squares, triangles, etc. The pyramidal and quadratomic tree belong to the mosaic hierarchical models.

The pyramidal model is based on the use of exponentially growing blocks of discrete arrays, each subsequent one representing one fourth of the size of the previous level. The pyramid model is most effective when it is necessary to increase the speed to obtain an image.

Now consider a quadtree whose model is based on a recursive decomposition of a square grid, where the resolution is determined by data that has spatial homogeneity. The difference between the quadratic tree model and the pyramid is that the first has a variable resolution, while the second has a composite one [9].

The advantages of the quadratic tree model are as follows: spatial relationships are encoded in the model itself, they carry out fast search, with an increase in the homogeneity of the data, they are formed in a more compact version.

Along with the advantages of such a data presentation model, there are, of course, disadvantages that also need to be considered in order to draw up a more objective opinion. In a generalized version, the
disadvantages look like this: quadratomic trees are laborious when creating, when any change in the data plane requires recalculation of the entire tree, these models are not always effective when storing satellite data and information about digitizing territories.

We will also give indicators of the advantages of using this particular model in organizing data that are not indicated above, which are essential for GIS design:

- the larger quadrant becomes a node of a higher hierarchical level of the quadratomic tree, and smaller quadrants appear at lower levels.
- calculations of the optimal areas for the placement of signs, centroid definitions, accelerated pattern recognition;
- conversion of distances for the most accurate transmission of metric data.
- splitting images in order to create separate graphic layers.

4. Assessment of the complexity of automating the extraction of graphical information

The main requirements for algorithmic support in the design of SGIS are as follows: the use of adaptive models of the technical and network infrastructure for object management, with subsequent integration into the regional level automation system, as well as spatial data referencing, solving optimization problems for the placement and selection of objects, determining the shortest path on an electronic map, analysis of complex structured data, etc.

One of the stages of clarifying a clear algorithm in the design of SGIS for solving the problems described above is the assessment of the time and capacity complexity, as well as measures to minimize them. [10]

The overall difficulty score is as follows:

1) Time spent: \( T = T_1 + T_2 \),
   - \( T_1 \) - estimation of task execution time in an application built using GIS (geographic information system) toolkit,
   - \( T_2 \) - estimation of task execution time in an auxiliary application.
2) Required memory: \( Q = Q_1 + Q_2 \),
   - \( Q_1 \) - an estimate of the required memory to complete a task in an application program,
   - \( Q_2 \) - an estimate of the required memory to complete the task in the auxiliary.

Let us consider the estimation of the complexity of extracting graphic information in the environment of the auxiliary program. The time spent or the time complexity of the algorithm reflects the time required for its operation. This is a function that assigns to each input length \( n \) the maximum (over all individual problems of length \( n \)) time spent by the algorithm to solve individual problems of this length [11].

It is necessary to estimate the processing time depending on the number of vertices and edges of the original graph \( G=(X;U) \)

\[
T_2 = c_1 n^a + c_2 n^b, \tag{4}
\]

\( c_1 \) and \( c_2 \) - experimentally determined coefficients of the complexity of the program implementation, \( a \) and \( b \) - estimation of the number of operations in the program related to the processing of vertices and edges of the original graph, \( n \) - number of vertices in a graph, where \( n=|X| \), \( m \) - number of graph edges and \( m=|U| \).

The essence of the experiment consists in measuring the processing time for several graphs and then using the polynomial function of the best approximation to determine the coefficients \( c_1 \) and \( c_2 \).

We will use the initial data for testing in order to determine the coefficients \( c_1 \) and \( c_2 \) in (4):

- Graph 1 (\( n_1, m_1 \)) - measure the processing time (\( t_1 \)) of graph 1
- Graph 2 (\( n_2, m_2 \)) - measure the processing time (\( t_2 \)) of graph 2
- Graph 3 (\( n_3, m_3 \)) - measure the processing time (\( t_3 \)) of graph 3
- Graph 4 (\( n_4, m_4 \)) - measure the processing time (\( t_4 \)) of graph 4
\[\begin{align*}
c_1 n_1 a + c_2 m_1 b &= t_1 \\
c_1 n_2 a + c_2 m_2 b &= t_2 \\
c_1 n_3 a + c_2 m_3 b &= t_3 \\
c_1 n_4 a + c_2 m_4 b &= t_4 
\end{align*}\]

(5)

Using programs for approximation (MathCAD is used) of functions by polynomials of the best approximation, the coefficients \(c_1\) and \(c_2\): \(c_1 = 4.12, c_2 = 3.17\).

The estimate of the required memory consists of four components of the placement: the initial information \(q_1\), intermediate data \(q_2\), result \(q_3\) and programs \(q_4\).

Consider this below.

RAM (Random Access Memory):

\[
Q'_2 = q_1 + q_2 + q_3,
\]

\(q_4 + q_2 + q_3\) – variable part;

(6)

DSD (disk storage device):

\[
Q''_2 = q_1 + q_4,
\]

(7)

\(Q''_2\) – archive reserve on hard disk, \(q_4\) - permanent part (for placing the application and auxiliary program).

To place the initial graphic information you will need:

\[
q_1 = q_{11} + q_{12},
\]

(8)

\(q_{11}\) – the required memory for placing the coverage of the layer with graphical information (for example, the road network), \(q_{12}\) - the required memory for placing the coverage (layer) with additional graphical data, and

\[
q_{12} = q_{121} + q_{122},
\]

(9)

\(q_{121}\) – is the amount of graphical data coverage information without text content, \(q_{122}\) is the amount of data coverage information with text content.

Since the conceptual apparatus of the program does not recognize letter designations, \(q_{121}\) is required to place the initial information. Thus, transforming (9) it turns out:

\[
q_{12} = q_{121}
\]

(10)

For example: this can be the import of a coverage of graphic information (road network) and additional graphic data from an application program into an auxiliary program, and then combining them into one coverage provides an opportunity to solve the problem of finding the shortest path.

5. Conclusion

According to the main results, the developed specialized geographic information system (SGIS) can be considered as an information advisory system, with the help of which analysts get the opportunity to relatively quickly answer the tasks (questions) of interest to them. The functional features of the system should enable the analyst from the controls to simply write down the problem statement (without programming).

By varying the parameters in emerging issues of geographic referencing, existing information, the analyst gets the opportunity to determine development trends and dependencies, which, in the end, provides an opportunity to formulate an optimization problem on the issue of interest. It should be noted that an analyst is a specialist who studies analytical research and generalizations in a particular field of activity, who is fluent in methods of analysis, is usually able to predict processes and develop promising development programs.

The system is equipped with an information retrieval module, which, in a short time, will search and visualize graphical and textual information necessary to support decision-making for a manager of the
middle administrative apparatus. The system must be based on an interactive procedure that sets a search command after receiving a request.

In conclusion, we note that the search time for the necessary information in a specialized geoinformation system directly depends on the development of a graphical-analytical model of the thematic layer, which in turn depends on the choice of the method of organizing information data in the system. Therefore, due to the advantages described above, it can be concluded that the use of the quadratomic tree method for organizing a graph-analytical model for storing geographic data is promising from the point of view of the efficiency of information presentation.

References

[1] Smirnov S V 2019 Design of conventional signs for the development of a geographic information system. Materials of the International Scientific and Practical Conference "Fundamental and Applied Research in Science and Education" (Taganrog, 2019) (Sterlitamak: AMI, Part 1) pp 199-205

[2] Smirnov S V 2020 Application of the information system for monitoring and decision-making support. Collection of selected articles of the International Scientific and Methodological Conference "Problems of Education Quality Management" (SPb: GNII "National Development") pp 48-52

[3] Smirnov S V and Tyukavkin D V 2003 Geoinformation system for decision support in the governing bodies of the social and educational sphere Problems of management 3 54-60

[4] Smirnov S V 2011 Designing Graphic Systems with Complex Data Structures (Saarbrücken: LAP Lambert Academic Publishing)

[5] Vasmut A S 1983 Computer Modeling in Cartography (Moscow: Nedra)

[6] Christofides N 1978 Graph Theory: Algorithmic Approach (Moscow: Mir)

[7] Akimov O E 2001 Discrete Mathematics: Logic, Groups, Graphs (Moscow: Lab. basic knowledge)

[8] Smirnov S V 2020 Construction of visual means of the analytical model of the thematic layer of a specialized geographic information system using a quadratomic tree. Collection of articles of the International Scientific and Practical Conference "Improving the Methodology and Organization of Scientific Research for the Development of Society" (Novosibirsk, December 29, 2020) (Sterlitamak: AMI, 2020. Part 2) pp 182-6

[9] Samet H, Shaffer C, Nelson R, Huang Y, Fujimura K and Rosenfeld A 1986 Recent Developments in Quadree-Based Geographic Information Systems. Proceedings of the Second International Symposium on Spatial Data Handling (Seattle, Washington) pp 15-32

[10] Smirnov S V 2020 Information models for solving problems in the field of government administration and assessing the complexity of automating the extraction of graphic information when solving them using GIS Eurasian Union of Scientists 5(12.81) 50-3

[11] Time Complexity Retrieved from: https://en.wikipedia.org/wiki/Time_complexity