MODEL OF INTELLECTUAL VISUALIZATION OF GEOINFORMATION SERVICE

Stanislav L. Belyakov
Scientific and Technical Center
"Information Technologies"
Southern Federal University
347900, Taganrog, Russia
E-mail: beliacov@sfedu.ru

Alexander V.Bozhenyuk
Scientific and Technical Center
"Information Technologies"
Southern Federal University
347900, Taganrog, Russia
E-mail: avb002@yandex.ru

Marina L.Belykova
Scientific and Technical Center
"Information Technologies"

Southern Federal University
347900, Taganrog, Russia
E-mail: beliacov@yandex.ru

Igor N.Rozenberg
Public Corporation "Research and Development Institute of Railway Engineers"
Department of Computer Science
109029, Moscow, Russia
E-mail: I.rozenberg@gismps.ru

KEYWORDS
Visualization, geographic intellectual system, automatic mapping, uncertainty, fuzzy function, granulation.

ABSTRACT

In this paper we investigate a model of intellectual visualization of cartographic images. The selection of customer of geoinformation service most informative materials during the session is modeled. The use of a customer of geoinformation service of fuzzy function usefulness is a feature of the model. Model of selection of informative cartographical objects in the workspace of analysis is described. Estimation of level of usefulness uses the knowledge about the growth and reducing usefulness, witch depending on the number of objects in the cartographic image. The usefulness function is presented in a granular form. Cartographic description of the utility function is considered. Image defects due to mapping of partially defined situations are analyzed. These situations appear on the map due to the imperfections of algorithms of automatic recognition of real world objects. Visual and operational defects are marked. The model of the map visualization with defects of displaying of uncertain situations is built. The proposed approach will reduce the risk of making wrong decisions due to the incomplete and irrelevant maps of geographic information systems.

INTRODUCTION

Geoinformation services are a powerful tool for information support of decision-making processes. Management of objects and processes of the real world is based on the use of spatial data from electronic maps, diagrams and plans. Traditionally, geoinformation systems (GIS) are used for the storage and use of cartographic information. Geoinformation Services (GS) is a geographic information system LAN or WAN. It continuously accumulates the map data and provides the access to this data through programming or interactive dialog interface. Google MAPS is a typical example of the geoinformation service.

Visualization is very important for map data when used in dialog mode. Visualization of map stimulates the intuitive creative thinking and promotes the use of deep knowledge of user-analyst. This mechanism is important for a decision making in difficult situations. A lot of complicated problems are solved in the process of visual studying of the maps (Egenhofer 2002, Erle et al. 2005, Cartwright et al. 2007).

Completeness, accuracy and relevance of map data are required for effective decision making. Low quality of maps generates the risk of damage due to an incorrect decision. However, the selection and study of useful data is problematic. The problem occurs for the following reasons:

- The map data redundant. None of application tasks needs all of the data simultaneously in the system. Analysts use data from local areas for solving the problems use;
- In a work session of analyst and GS the dynamics of change the operational set of data is present. Examining of images involves the addition and the removal of large volume of the map tiles;
- Analysts notion the quality of selected data is subjective.

As a consequence users spend a lot of efforts for management of the visual image. Quality of the basic problem solution is reduced inevitably. As any GS continuously accumulates the information about the outside world, this problem is escalating over the time.

Because of the uncertainty and ambiguity of information visualization is of interest to the selection of useful information modeling of methods of artificial

Proceedings 28th European Conference on Modelling and Simulation ©ECMS Flaminio Squazzoni, Fabio Baronio, Claudia Archetti, Marco Castellani (Editors)
ISBN: 978-0-9564944-8-1 / ISBN: 978-0-9564944-9-8 (CD)

intelligence. This question is little studied to date. Probably the reason is that the problem of selecting of useful map data is referred to the map construction (Li et al. 2013, Pettit 2008). It is assumed that any selected fragment of the map by the analyst automatically has maximum information. From our point of view, a selection mechanisms of a map data should be realized by GS. This will allow to newly solve the problem of timely updates for geographical maps. It is known that the period of their updates is standardly 5-7 years. In a number of applications this is totally unacceptable. An alternative option is to fill the GS database with «raw data» and use is in real time mode.(Konecný and Bandrova 2006).

This paper offers a model of visualization for useful map data selection in applied problem solving. Intellectual component of visualization process is defined by the accumulated knowledge (Luger 2004, Sowa 1999).

PROBLEM FORMULATION

The cartographic visualization task is seen as the process of map image of preset quality construction. This approach is different from the similar ones as it allows maximizing the efficiency of the cartographic images. It is expected that GS will generate the most efficient cartographic image. Herewith the generated output includes the uncoordinated objects and may be inconsistent with mapping standards. The survey shows that this kind of approach was never researched before.

The following tasks are to be solved when modeling the described process:

- To formalize the notion of cartographic image efficiency and to set objectives for visualization management;
- To fuzzy model the cartographic image efficiency;
- To model the visualization of uncoordinated and ill-defined objects for cartographic images.

MODEL OF VISUALISATION CONTROL

Solving the applied problem with the help of GS, the user implements a procedure of the cartographical analysis (Batty 2012, Dent 1990, Berlyant 1998).]. The creation of a local working area of global GS map is the basis of the procedure.

The working map area $\Omega = \{\omega_1, \omega_2, ...\omega_n\}$ that consists of map objects ω_i is a subset of objects $m_W \subseteq \Omega$ that describes the fragment of a map with boundaries

$$L_W = \{S_W, T_W, C_W, E_W\},$$

where S_W – spatial boundary, T_W – temporal boundary, C_W – semantic boundary, E_W – pragmatic boundary. The semantic boundary C_W is a set of classes of objects and relationships, which are described in the GS, E_W is a description of the limits of applicability of the map constructed. The specific work area with boundaries is generated by the sequence of request from the client to the GIS server:

$$Q_i(X_S, X_T, X_C, X_E), i = \overline{0, N},$$

where X_* are, accordingly, the spatial, temporal, semantic and pragmatic parameters of a single request.

$Q_i(X_S, X_T, X_C, X_E)$ is a predicate of the request which in modern GS is built by user through the system of dialogue menus or is directly set as an expression in SQL.

The structure of the workspace $m_W \subseteq \Omega$ can be represented as the union of two sets:

$$m_W = B \bigcup E,$$
$$B \subseteq \Omega : \forall \omega_i \in B \Rightarrow \exists Q_j(X_S, X_T, X_C, X_E) = true,$$
$$j = \overline{1, N}, i = \overline{1, |\Omega|},$$
$$B \bigcap E = \varnothing, E \subseteq \Omega.$$

The set B is a skeleton of the request which is determined by predicates of requests, E – environment of the skeleton, in other words, it is subset of map objects that provides the continuity of maps (Berlyant 1997), and semantic entirety of the work area. The entirety is formed by using the expert rules of construction of the image of expert response $K(B, \Omega)$ to the skeleton of request B :

$$\omega_i \in E \Rightarrow K(B, \Omega) = true, i = \overline{1, |m_W|}.$$

Expert rules $K(B, \Omega)$ represent knowledge about how to construct the mapping images needed for solving the problem. Application of the rules $K(B, \Omega)$ leads to a reduction in redundancy of cartographic images and improvement of their informational content.

Informational content $I(m_W)$ of any work area is a related, subject determined value. Its value is validly estimated only in the narrow scope of a particular class of applied problems that can be solved by a certain group of users. General restrictions following from the specification of a person's perception of graphic images are as follows:

$$|m_W| = 0 \Rightarrow I(m_W) = 0,$$
$$|m_W| = \infty \Rightarrow I(m_W) = 0.$$

The level of information content cannot be estimated by any other way except for a set of expert rules $K_I(m_W)$. Rules reflect the subjectivity and ambiguity that are inherent to the evaluation of information content.

Expert knowledge in the form of rules $K(B, \Omega)$ and $K_I(m_W)$ represent the "reasonable" rendering of map images strategy. The difference between the described intellectual visualization and the traditional

one is in support of subjective utility of map images. Not only the objects that meet the predicate of the request are visualized, but all those that fill a cartographic image with a meaning.

The management of visualization means solving the following problem:

$$\begin{cases} I(m_W) \to \max, \\ |m_W| \le m^*, \\ m_W : Q_i(Xs, X_T, X_C, X_E) = true, i = \overline{1, N}. \end{cases}$$

m^* stands for the image complexity limit. The complexity is evaluates by the number of primitives.

To solve the problem, the GS should be equipped with software mechanism of expert evaluation $I(m_W)$ and a mechanism of adding and removing mapping objects in the working area. The GS works in a following way:

- The user authenticates himself at the beginning of a session with GS. This allows us to classify and choose it from a database of expert rules $K(B,\Omega)$, $K_I(m_W)$ and necessary thematic layers of a map. The effectiveness of further work in the session depends on how adequate classification is. The user might not be satisfied with the system performance. Thus the user may start a new session with other authentication parameters;
- The GIS server builds workspace $m_W = B \bigcup E$ with maximal information content for any client request. After that the server synchronizes the workspace with the client. Accounting of the limited resources of the client $|m_W| \le m^*$ ensures the completeness of the result and security of interaction.

USABILITY ASSESSMENT APPROACH

Estimating the information content of the workspace $I(m_W)$ is the basis of the process of visualization management. In order to set the function of information content we have to take into account its dependency on the following factors:

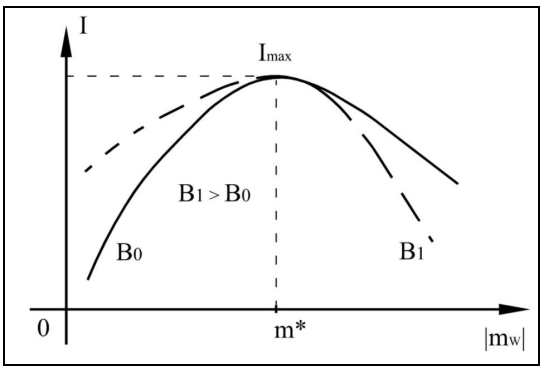
- the total number of map objects in the working area;
- the distribution of copies and classes of objects by the level of significance of the problem to be solved;
- the distribution by the level of importance of samples and classes of relations between objects;
- the degree of novelty in area studied for the user;
- the user's work experience with cartographic materials.

It can be said that information function depends on many variables $X = \{x_1, x_2, ..., x_M\}$:

$$I(m_W) = F(X).$$

The number of variables is not known beforehand. The multidimensionality of the space of factors $\{x_1, x_2, ..., x_M\}$ creates serious difficulties in finding the analytical relationship $I(m_W)$. Therefore it is proposed to evaluate the informative content by a fuzzy description. The description is based on expert data. This approach is explained by the high degree of incompleteness and uncertainty about the behavior of the function $F(x_1, x_2, ..., x_M)$. The model of granular representation of the information function should be reviewed (Zadeh 1997).

The analysis of the analysts' behavior shows that the function $I(m_W)$ can be characterized by the curve in Fig. 1. The shape of the curve reproduces the qualitative features of the information content changes. With a small number of graphic objects in the workspace m_W information content is not high and grows as you add new objects. The growth rate of information content should be related to the adding of objects from the environment (*E*), because the information content of skeleton *B* represents the minimum level that the working space gets under the user's request. The dotted line shows the information curve for skeleton B_1 with a large number of objects ($|B_1| > |B_0|$).



Figures 1: The usefulness function

Where $I_{\max}$ is the highest possible level of workspace's m_W information content. Qualitative analysis of the curve in Figure 1 allows us to enter at least three granules $\{I_L, I_M, I_R\}$. Each granule defines the area in space $I \times N$ in a fuzzy way:

$$I_L = \{\mu_{Lk} / (i_{Lk}, N_{Lk})\}, I_L \subset I \times N,$$
$$I_M = \{\mu_{Mk} / (i_{Mk}, N_{Mk})\}, I_M \subset I \times N,$$
$$I_R = \{\mu_{Rk} / (i_{Rk}, N_{Rk})\}, I_R \subset I \times N.$$

Where μ_{Ak} is the degree of belonging of the point (i_{Ak}, N_{Ak}) with the number k to a granule with an index A, an each pair (i_{Ak}, N_{Ak}) is a value of information content i_{Ak} with a number of objects N_{Ak}.

Please note that the definition of function in the space $I \times N$ is a simplification. Simplification is used deliberately to implement controls.

Granule I_L represents an area of increasing information content for cartographic images of low complexity of perception. Granule I_M is the area of the most informative images with the utmost level of perception. Granule I_R is the area of falling informative content of images difficult to perceive.

For the use and description of the granules $\{I_L, I_M, I_R\}$, we suggest using the cartographic representation. The essence of method is in constructing figurative and symbolic model of granules. The model describes the distribution of the importance of classes and samples of objects and relations of the work area; it also describes the relationship between the grade of information content and the total number of objects in the working space. Demonstrativeness of the cartographic representation, from our point of view, is extremely important in obtaining reliable and coordinated knowledge from expert GS users (Harrie and Weibel 2007).

Cartographic representation has a form of atlas – a set of maps that shows the behavior of informative value. Formally, this means that the information function is a superposition of functions

$$I(m_W) = F(x_1, x_2, ..., x_M) = F(Y_1, Y_2, ..., Y_H),$$

where

$$H < M, Y_i = Y_i(X_i), X_i \subset X, i = \overline{1, M}.$$

Each map of atlas shows the dependence $Y_i(X_i)$. The coordinate axes may be connected with different degrees of complexity. Each map is a graphic "projection" of the multidimensional space of problem domain concepts. The problem of mapping $I(m_W)$ thus is reduceds to the determination of a set of "projections" with the required properties.

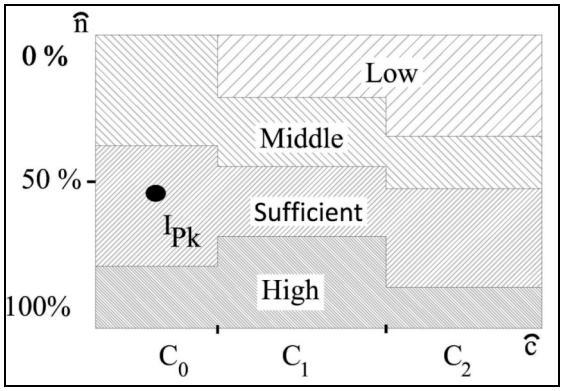
Let us consider the example of mapping. We assume that the atlas consists of two maps: the first allows determining a preliminary assessment of informative value, the second determines the granule, which contains the analyzed workspace of the map. Technically

$$I(m_W) = F(I_P, |m_W|), I_P = Y_1(\hat{n}, \hat{c}).$$

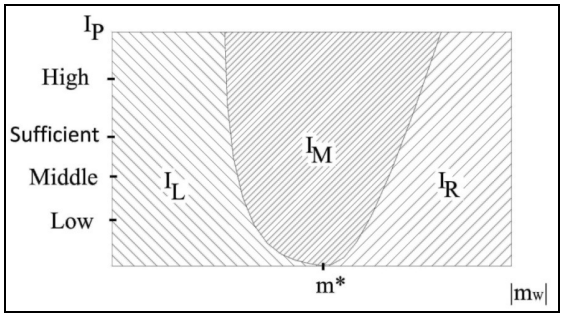
The value I_P is determined from the set {«Low», «Middle», «Sufficient», «High»}. Any value of I_P subjectively displays a preliminary evaluation of informative value. The value of I_P depends on the use of layers, objects and relations of the workspace map. Fig. 2 shows the example of a map that binds used layers, objects and relationships that are clustered into three clusters $\{C_0, C_1, C_2\}$. The coordinates of any point on the map is $(\hat{n}, \hat{c})$ where $\hat{n}$ is a fuzzy evaluation of the relative number of objects and workspace relations, $\hat{c}$ is the name of the most significant cluster of objects and relations in the work area. For example in Figure 3 point I_{Pk} displays workspace that contains some $\hat{n} = 40\%$ elements of the cluster $\hat{c} = C_0$. The cluster C_0 includes:

- examples of "automobile filling station";
- examples of the relation "nearby service stations";
- layers of "parking lots", "motorway", "restaurants", "service station", "hotels", "ATM";
- relations "crosses", "attached", "located in the danger zone."

The map of preset informative content is zoned. Each area corresponds to one of the values of the set {«Low», «Middle», «Sufficient», «High»}. Figure 3 shows a map of zoned granules of information content $F(I_P, |m_W|)$.



Figures 2: The map of preliminary informativeness



Figures 3: The map of the arrangement of information content granules

In this example of the map the fuzziness of granules is not displayed for the sake of simplicity. The coordinates of any point $(I_P, |m_W|)$ clearly belong to a single granule. . Having determined I_{Pk} on the map (Fig. 2), counting the total number of objects in the workspace m_W, you can accurately match the workspace to the granule of information content on the map (Fig. 3). All procedures are implemented in software.

VISUALIZATION FOR PARTIALLY DEFINED SITUATIONS

To get the operational spatial data one needs to access heterogeneous systems that record things and events of the real world. The Internet plays a special role in that situation. For example, the information about the natural anomalous situation can be got as a message from the news flow (RSS), published space or aerial photos, video streams from Web-cameras, reports from the electronic media, from personal blogs, from specialized communities in social networks, as well as from the map services. Quite often this information is not metric and does not contain an explicit gridding. However, its value in the case of responsible decision-making is very high. The information values may compensate the lack of accuracy and completeness.

Let us assume that GS has a search engine that can find information resources for later retrieval of spatial data. GS also has a set of programs for identification and mapping of situations. It is well known (Harrie and Weibel 2007) automatic mapping is not perfect. Newly created mapping objects can disrupt the logic of the map

Let us consider the model of visualization for the case when the objects inconsistent with the other map elements are added to the map.

Let us describe the set of mapping objects built by the recognition of situations outside the world as $S = \{ s_i \}$.

The information base of GIS is complemented by multiple map objects S in the process of real-time mapping:

$$\overline{\Omega} = \Omega \cup S .$$

Unlike any other object $\omega_i \in \Omega$, the space-time coordinates and semantic attributes $s_i \in S$ are not completely reliable because of the limited capacity of the recognizing subsystem.

Adding context to the working area and its further visualization with maximal information content

$$I(B \cup E \cup S)$$

cannot be performed by the method discussed above. The reason is that the visualization of situations $s_i \in S$ creates defects of cartographic displaying. The defect appears, for example, like an overlay of images of situations $V(s_i)$ and objects of cartographic basis:

$$V(s_i) \cap V(\omega_j) \neq \varnothing .$$

Analysis has shown that defects in the mapping image can be divided into two classes.

Defects of the first class are visual defects that create a misconception about the objects with overlapping images. But in this case any of the overlapping objects of the map would be perceived incorrectly under the visual analysis. Fig. 4 shows an example of a map with the representation of two situations: the both are displayed as areal objects of different colors. The situation depicted in gray overlapped with the line of the railway. However, the map is correctly recognized by the analyst. Contrary to that, the situation displayed by an area objects in black, hides an important object to analyze – the bend of the road. The visual recognition of the road would be incorrect.

Operational defects form a second class. Defects of this class produce incorrect results when the procedures of cartographic analysis are performed. The example in Figure 4 shows how an error of estimation of the topological relation "to be around" occurs. If you request "to find all objects that are close to the railway", the result will be incorrect. The object of situation shown in gray is for "overlap". The performance of a particular GS determines the possibility of operational defects. It may happen that the GIS has no evaluation procedures of topological relations. Then the operational defects are not possible.

Despite the fundamental differences in the causes of occurring and influence of defects, we can claim that any defect reduces the productivity of the cartographic image. In this case the procedure of productivity maximization can make one of two actions:

- remove the object $s_i \in S$ that raises the defect;
- correct the defect by removing the objects $\omega_i \in \Omega$.

Figures 4: The example of image with situation $s_i \in S$

Let us estimate the relevance of each action. Suppose that after the elimination of defects the set $D = \{d_i\}$ of objects removed from the workspace is formed. Then the modified working area described by the set

$$\tilde{m}_W = B \cup E \cup S \setminus D .$$

Adding indefinite situations to the workspace has the effect, if the following inequality is right:

$$I(B \cup E \cup S \setminus D) > I(B \cup E).$$

Finally, the procedure of visualization control can be described as follows:

- to build for the next request $Q_i(X_S, X_T, X_C, X_E)$ skeleton B and the environment E;
- to assess the information content of the constructed workspace;
- to construct objects $S = \{s_i\}$ that display information from external sources;
- to identify the visual and operational defects and to find objects that can be removed and thus that removal will increase the information content of cartographic images by a linear search;
- to synchronize the workspace with the GS client.

CONCLUSION

The approach described above results in the following.
The main factor for smart visualization quality improvement is the adequate balance between the expert description and the subjective perception of cartographic images. The productivity function when displayed in granules represents the unclearness and incompleteness inherent to the expert knowledge on productivity on the whole.
An important factor in the quality of smart visualization is the presence of a GS engine that allows extracting spatial data from data sources of different nature. The more important the data is the greater deviation from the standards of cartographic representation comes. The proposed approach puts the visual analysis of maps in the first place. A powerful mechanism inherent to men – an intuitive understanding of the meaning of the map is activated with the support of GS.
This approach gives the following positive results.
We assume that the wrong decision leads to damage U. The damage occurs when the map contains no objects $S = \{s_i\}$. The damage is reduced if such objects appear on the map and analyst successfully identifies the state of the outside world. The information content is maximized and it can be assumed that an absolutely accurate identification is determined by the effectiveness of recognition programs. We denote p by the probability of correct recognition of a given situation. Then, using the formula of the binomial distribution

$$P_N(k) = \binom{N}{k} p^k (1-p)^{N-k}$$

we can calculate the probability of that k situation would be recognized from N actually existing ones. Mean cost in this case is calculated as

$$\overline{U} = (1 - P_N(k))U .$$

This expression allows you to evaluate the effect depending on the parameters of the recognition system (probability p), and the specific of the decision making process (values N and k).

Acknowledgements. This work was partially supported by the Russian Foundation for Basic Research, projects N 13-07-13103 ofi_m_RzhD, and N 12-07-00032.

REFERENCES

Batty, M. 2012 "A Generic Framework for Computational Spatial Modelling". In*: Agent-Based Models of Geographical Systems.* Springer Science+Business Media B.V.

Berlyant, A.M.1998." 'Not For Use To Foresee…', Or About Future Maps". *Mapping Sciences and Remote Sensing.* V.H. Winston & Sons, Inc. No.3(35),166-172.

Cartwright, W.; M.P Peterson; and G. Gartner. 2007. "Multimedia cartography". Springer Verlag, Berlin.

Dent, B.D. 1990. "Cartography: Thematic map design, 2nd edn". Wm. C. Brown Publishers, Dubuque, Indiana.

Egenhofer, M.J. 2002. "Toward the Semantic Geospatial Web". In *Proceedings of the* 10th *ACM International Symposium on Advances in Geographic Information Systems (ACM-GIS-2002)*. Virginia: ACM Press, 1–4.

Erle, S.; R. Gibson; and Walsh, J. 2005. "Mapping Hacks – Tips & Tools for Electronic Cartography". Sebastopol, CA, O'Reilly Press.

Harrie, L. and R. Weibel. 2007. "Modelling the overall process of generalization". *In: Ruas.A., Mackaness, W.A. and Sarjakoski, L.T. (eds.).: Generalization of Geographic Information:Cartographic Modelling and Applications, Series of International Cartographic Association*. Elsevier, 67–87.

Konecný, M. and T. Bandrova. 2006. "Proposal for a Standard in Cartographic Visualization of Natural Risks and Disasters". In *Proceedings of the Joint Symposium of Seoul Metropolitan Fora and Second International Workshop on Ubiquitous, Pervasive and Internet Mapping*. Seoul, Korea, 165–173.

Li, W.; L. Li; M.F. Goodchild; and L. Anselin. 2013."A geospatial cyberinfrastructure for urban economic analysis and spatial decision-making". *ISPRS International Journal of Geo-Information,* No.2, 413–431.

Luger, G.F. 2004. "Artificial Intelligence: Structures and Strategies for Complex Problem Solving". Addison Wesley.

Pettit, C.; W. Cartwright; I. Bishop; K. Lowell; D. Puller and D. Duncan.2008. "Landscape Analysis and Visualisation, Spatial Models for Natural Resource Management and Planning". Springer-Verlag, Berlin.

Sowa, J.F. 1999. "Knowledge representation: logical, philosophical, and computational foundations". Brooks Cole Publishing Co, Pacific Grove.

Sui, D.; S. Elwood; and M. Goodchild. 2012. "Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice". Springer.

Zadeh L.1997. "Towards a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic". *Fuzzy Sets and Systems*, No.90, 111-127.

AUTHOR BIOGRAPHIES

STANISLAV L. BELYAKOV was born in Ukraine. He graduated Leningrad Electrotechnical Institute in 1982. Professor of Scientific and Technical Center "Information Technologies", South Federal University, Russia. He holds a degree of Doctor of Technical Sciences in Theoretical Foundations of Informatics in 2003. His research interests include geographic information systems, modeling intelligence. Contact information: 44, Nekrasovsky Street, Taganrog, 347900, Russia, phone: +78634371743, e-mail: beliacov@yandex.ru.

ALEXANDER V. BOZHENYUK was born in Taganrog, Russia. Professor of Scientific and Technical Center "Information Technologies", South Federal University, Russia. He holds a degree of Doctor of Technical Sciences in Theoretical Foundations of Informatics. His research interests include fuzzy models, fuzzy decision making. He has more than 180 publications in the fields of Computer Sciences. Contact information: 44, Nekrasovsky Street, Taganrog, 347900, Russia, phone: +78634371743, e-mail: avb002@yandex.ru.

IGOR N. ROZENBERG was born in Taganrog, Russia and went to the Radio Engineering Institute. He holds a degree of Doctor of Technical Sciences in Geographic Information Systems in 2007. He is the First Deputy General Director of Public Corporation "Research and Development Institute of Railway Engineers". Contact information: 27/1, Nizhegorodskaya Street, Moscow, 109029, Russia, phone: +78634371743, e-mail: I.rozenberg@gismps.ru.

MARINA L. BELYAKOVA was born in Taganrog, Russia. She graduated Leningrad Electrotechnical Institute in 1984. Received a Ph.D. in 1989 in Taganrog Radio Engineering Institute. Contact information: 44, Nekrasovsky Street, Taganrog, 347900, Russia, phone: +78634371743, e-mail: beliacov@yandex.ru.