

Table 3
The contribution of eigenvalues λ_i of matrix S , in percentage of their sum

i	1	2	3	4	5	6	7	8
$\lambda_i, \%$	2,50	2,49	2,85	2,14	1,94	3,29	4,00	4,73
i	9	10	11	12	13	14	15	
$\lambda_i, \%$	5,26	6,83	7,78	9,99	11,31	13,83	21,07	

Noise clearing with the 3 most significant components is $W = 21.8 \%$, for 4 it's $W = 29.2 \%$ and for 5 it's $W = 34.6 \%$.

The results for 3 selected components are shown below as graphs (fig. 3).

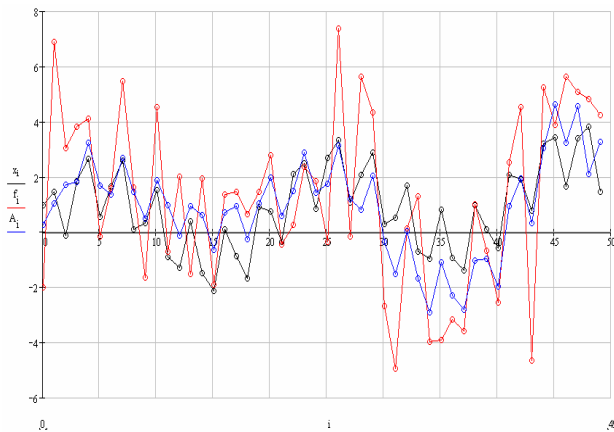


Fig. 3. Graphs of series: “clear”, with noise and restored

Concluding the given examples we can state that the basic algorithm of the “Caterpillar”-SSA method copes with the assigned task: for time series it separates trend and periodicals from interferences, reducing noise level down to 2–3 times; although the types of significant components aren't defined, whether they are linear, periodic, logarithmic or other. This is an advantage of the method, which will make possible to create a powerful mechanism of non-parametric analysis of time series in the future, including computer programs.

The disadvantage of the basic algorithm is the necessity of manual intervention for the divided components analysis; also there is a problem in selecting the length of period and the quality of additive components division, depending on that. Further research will be dedicated to the automation of analyzing processes and other methods, improving the quality of the algorithm work results and reducing the manual aspect in this process.

References

1. Golyandina N. E. The method of “Caterpillar”-SSA: the analysis of temporal aisles : textbook. Saint-Petersburg, 2004.
2. The main components of temporal aisles: the “Caterpillar” method / under the editorial of D. L. Danilov, A.A. Zhigliavski. Saint-Petersburg : Presscom, 1997.
3. Golyandina N., Nekrutkin V., Zhigljavsky A. Analysis of Time Series Structure: SSA and Related Techniques. London : Chapman& Hall/CRC, 2001.

© Vohmyanin S. V., 2010

Yu. Yu. Yakunin, A. A. Gorodilov
Siberian Federal University, Russia, Krasnoyarsk

ALGORITHMS FOR CALCULATING COMPLEX INDICATORS IN DYNAMIC STRUCTURES OF DATA REPRESENTATION

This paper presents algorithms for calculating complex indicators on set factual data, represented in dynamic structures with the application of the graph theory.

Keywords: dynamic structures of the data, tripartite graph, algorithm of graph's round, complex indicators.

The problem of rupture between scientific methods of representing (describing) real world objects and the storage of this information in information systems has existed for a long time and has not yet been solved satisfyingly. The database management system (DBMS) is the best system available today, which allows the storage of information in the form of objects [1; 2] (the object-oriented approach [3]) or globals [1] (hierarchical representation of the information in the form of a tree). However, even such an approach can capture only part of the variety represented in the information of modern scientific methods [4]. Such rupture substantially slows the development of science and engineering in the field of information technology.

According to this, the essential restriction for information system design is the standard way of data storage, which is based on static structures (i. e. for the description of a subjected field's objects in order to store information, a database of the data storage structure is created in advance). This results in the fact that such structures should be created by the designer of information system during a stage of its designing and it (this structure) cannot change during the development of this system and its maintenance. It is not necessary to speak about the expenses at which changes in the system come... It is obvious that if such changes are possible, even in an insignificant part of this structure, it would come at the same expense as the original production of

the system, possibly even exceeding it. Any attempts of creating dynamical structures of data storage lead to problems of data transformation in one of the existing ways of their storage in DBMS (relational, objective, or hierarchical). Any of the listed ways is not capable of functioning directly with its dynamical structures. Frequently, such problems are solved individually in the process of their occurrence. As a result there are no general approaches for the solution of these problems; it is realized in some software products in the form of separate modules with similar problems, but these solutions are commercial secrets. Thus, there is a research problem in the description of dynamic data structures, information processing in these structures, and its storage in DBMS. There is also a separate problem of calculating complex indicators on a set of factual data in the dynamic structure, the solution of which is mentioned below.

For the description of dynamical data structure we have offered to allocate the following categories of the information [5]:

- indicators – quantitative characteristics of objects;
- qualifiers – structural formation of the data consisting of interconnected classes and their instance (objects);
- factual data – values of indicators concerning one or several qualifiers; to represent each category of information we will use corresponding graphs with descriptions.

The graph of indicators consists of two parts: the first – which represents a tree (fig. 1), containing nodes-categories (classification of indicators), and the second – a simple graph with any number of parts on the graph, the nodes of which represent indicators.

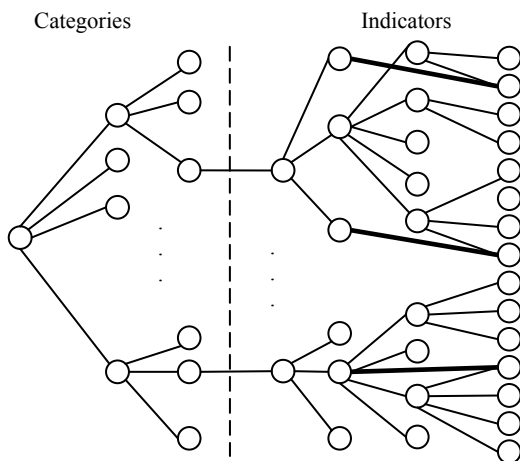


Fig. 1. The graph of indicators

In the rib section (from left to right) is the division of categories into subcategories. In the rib part indicator (from left to right) is the division of the aggregated indicators into more private and then to elementary indicators (indivisible). Fig. 1 shows that the aggregated indicators can consist not only of previous level indicators (k), but from indicators of the following level ($k + 1$), through level ($k + 2$), and so on.

The graph of qualifiers we shall consider is in the example of fig. 2. It can be clearly seen that the graph is divided into three levels, each of which carry own sense load. So level of categories is intended for splitting the qualifiers into integrated categories, i. e. the descriptions of a classification tree structure. At the qualifiers' level there are multiple nodes (qualifiers) with ribs between the category level, showing the qualifier attachment to a category; and ribs between qualifier nodes, showing the interaction between them. The node is characterized not only by name, but by a number of attributes.

The level of qualifier values, similarly to the previous level, contains multiple nodes; each of one has an obligatory rib depicting attachment of any qualifier's value to some of the unessential ribs, depicting the dependence between the values of different qualifiers. The nodes of the given level also are characterized by name and attribute values, corresponding to the nodes in the qualifier level with which they joined by ribs.

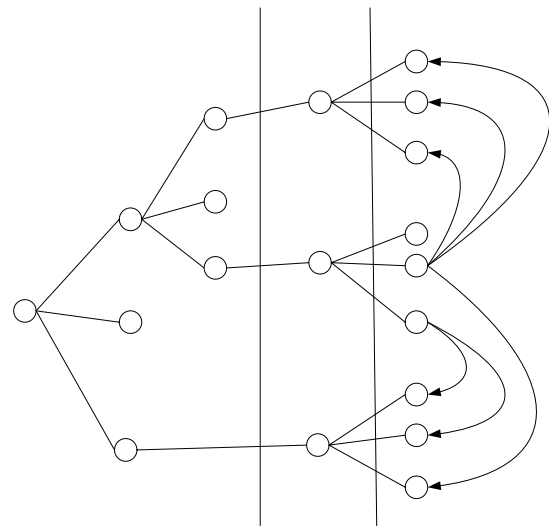


Fig. 2. Graph of qualifiers:
Cat – category of qualifiers; Q – qualifier; VQ – value of qualifiers

Let's draw an analogy graph representing the structure of qualifiers among with other ways of data presentation: for example, with the objective. As an example, we shall consider the organizational structure of some organization. We will allocate two essences in this example: management and department. The graph model of this example is presented in fig. 3, a, and the objective in fig. 3, b.

The factual data is the graph with a set of nodes, which do not have ribs between each other, but have ribs with qualifier and indicator nodes (fig. 4). Factual data nodes represent numbers, which are quantitative characteristics of one or several indicators in the relation with one or several qualifiers. For example, for the indicator "salary" and values of the qualifier "department of social work", the factual given value may equal X rubles a year.

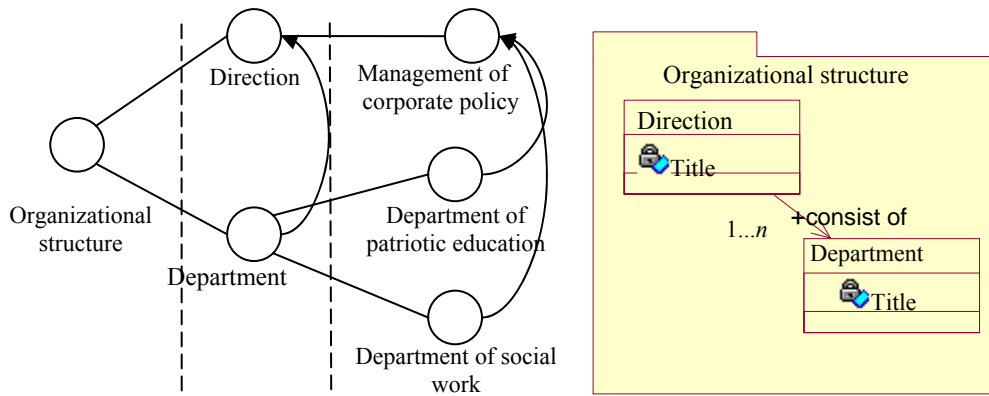


Fig. 3. Examples of representation structure:
a – graph view; b – objective view

Thus, the factual data nodes incorporating with nodes of qualifier and indicator graphs form a graph with difficult interrelations (fig. 5).

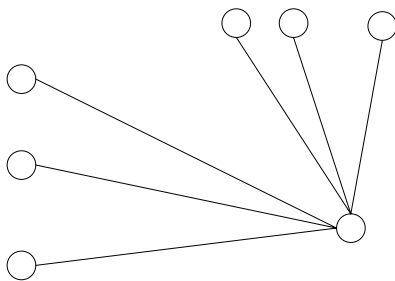


Fig. 4. Fragment from the factual data graph:
VQ – value of qualifier; I – indicator; FD – the factual data

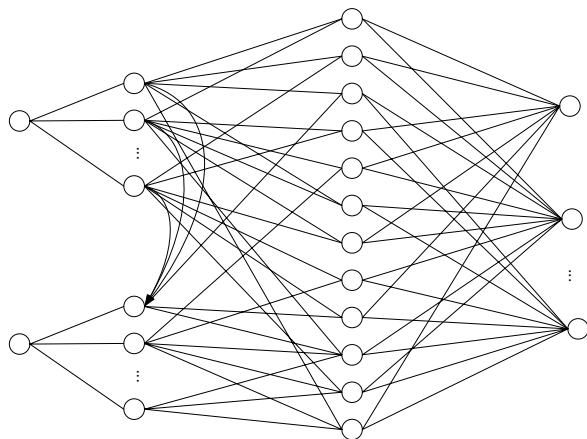


Fig. 5. The factual data graph:
Q – qualifier; VQ – value of qualifier; I – indicator;
FD – the factual data

In order to carry out analytical operations over the factual data for the purpose of calculating complex indicators from the presented graph (fig. 5), not all its nodes and ribs will need to be engaged. For example, in the organizational unit “management of corporate policy”, necessary for calculating the total annual salary, we need

the factual data sum values which are connected to nodes representing the management departments, and the node of the “salary” indicator.

Thus, from graph presented in fig. 5, only the factual data nodes and the qualifier counts nodes, along with indicators connected to them will be used. Besides, if for the initial data of the algorithm for calculating complex indicators will be used the set of qualifier value nodes and the set of indicator nodes for which it is necessary to make calculations; ribs between the qualifiers’ value tops will be not necessary. Therefore, considering all the aforesaid, we shall simplify the factual data graph (fig. 5) and transform it into a tripartite graph (fig. 6). It represents three sets of nodes (X_1 , X_2 , and X_3); in each, the nodes have no ribs among each other, but have ribs between the nodes of the following sets [6].

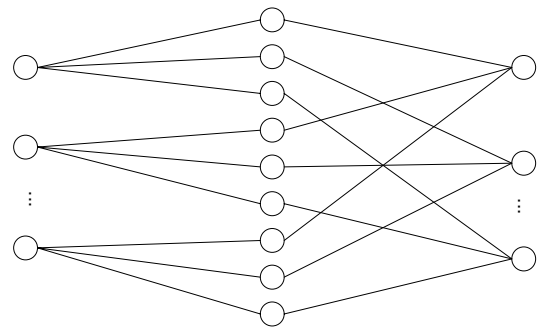


Fig. 6. Tripartite factual data graph:
VQ – value of qualifier; I – indicator;
 X – set of nodes

The tripartite graph is a particular case of the N -partite graph (fig. 7). It has its own properties, distinctive from those of the simple graph for which there are already some algorithms of its traversal [7].

Such algorithms poorly work on N -partite graphs as they are focused on achieving different targets (such as, finding the shortest way, searching for certain values, etc.), instead of solving the problem of selecting nodes in one partite with the set conditions through node subsets in the adjacent partite. With such properties the algorithm of traversal count is required to calculate the complex

indicators. In this article we have offered algorithms of traversal N -partite graph for the purpose of searching nodes of one share, having a full set* of ribs with set search conditions by node subsets in the following share of the graph.

*Conditions define subsets of nodes for $(k - 1)$ and $(k + 1)$ partite the graph if the search is carried out in k -th partite. If the node of k -th partite is connected to all nodes of the set subsets in $(k - 1)$ and $(k + 1)$ partite, such a node has a full set of ribs.

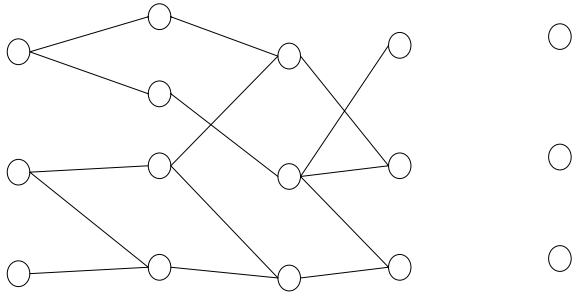


Fig. 7. N -partite graph
 x – node of the graph; X – set of nodes

To realize the search of nodes in one partite of the N -partite graph it is necessary to define two subsets of nodes X'_{k-1} and X'_{k+1} , united by multiple search conditions U_k (1), where k – is the partite number in which the search will be fulfilled. When realizing the search of nodes in the k -th partite, a subset will be generated, where the nodes having a full set of ribs with subsets from search conditions will be selected (for each node $x'_{k,i} \in X'_k$, condition (2) will be satisfied):

$$U_k = X'_{k-1} \cup X'_{k+1}, \quad (1)$$

$$(\Gamma(x'_{k,i}) \cap X'_{k-1}) \cup (\Gamma(x'_{k,i}) \cap X'_{k+1}) = X'_{k-1} \cup X'_{k+1}, \quad (2)$$

where $\Gamma(x'_{k,i})$ – is the set of all nodes having ribs with node $x'_{k,i}$.

Two new search algorithms of the graph N -partite are described further: “traversal of one node vicinities” and “the mark of nodes”.

Traversal of one node vicinities:

- select all nodes from set X_k , which have ribs with one element of set U_k and to place them in set X'_k ;
- check each node $x'_{k,i} \in X'_k$, whether it has a full set of ribs with nodes of set U_k . If the condition is not right, node $x'_{k,i}$ leaves set X'_k .

The analysis of the given algorithm has allowed us to reveal the dependence of the transitions C quantity, necessary for conducting the search, from the search conditions (3):

$$C = \sum_{i=1}^n (|\Gamma(x'_{k,i})| - 1), \quad (3)$$

where $|\Gamma(x'_{k,i})|$ – is the quantity of nodes connected to element $x'_{k,i}$.

Mark of nodes:

- choose node $u_{k,i}$ from set U_k . We will introduce a set of labels M the dimension of which is equal to $|X_k|$ and for each element of this set we will appropriate “0”. $i = 1$;
- for node $u_{k,i}$ we find set $\Gamma(u_{k,i}) \subset X_k$. For each $x_{k,j} \in X_k$, where conditions are realized $x_{k,j} \in \Gamma(u_{k,i})$, the corresponding label $m_j \in M$ will increase the value by “1”;
- $i = i + 1$. If $i = n$ we pass on to the following step. Otherwise we pass to step 2;
- find the node in set U_k with the least quantity of ribs and move the nodes connected to it from X_k to set X'_k ;

– for resultant set X'_{res} from set X'_k we select nodes having a full set of ribs with nodes from set U_k :

$$X'_{res} = \{ \forall x_{k,i} \in X'_k \mid m_i = |U_k| \}.$$

Let's define the given algorithm dependency of transitions C quantity, necessary for conducting search, from search conditions (4).

$$C = \sum_{i=1}^n |\Gamma(u_{k,i}) \cap X_k| + \min(A), \quad (4)$$

where $A = \{ |\Gamma(u_{k,i}) \cap X_k|, i = \overline{1, n} \}$.

The analysis of formulas (3), (4) has shown that at a small number of search conditions (from 1 to 4) is more effective (from the position of quantity transition between graph nodes) than the algorithm “the mark of nodes”; and works more effectively during the increase of search conditions in the algorithm “traversal of one node vicinities”.

The complex indicator is calculated by the factual data multiplier X_2 of the tripartite graph G_3 (fig. 6) by allocating the subset $X'_2 \subset X_2$, according to conditions U_2 (1), and performing operations over this subset. Thus, the calculation function of the complex indicator represents an algorithmic search function on a tripartite graph with the set conditions (5) and analytical transforming function over the received set:

$$CI = F(G_3, U_2). \quad (5)$$

Let's consider a calculating example of a complex indicator with the application of new search algorithms on the N -partite graph. In fig. 8 there is a factual data graph, in which it is required to calculate the quantity of turning graduates in 2007. In table there is a description of the account nodes (fig. 8) for the resulted example.

Symbols of qualifier and indicator values

Node	The node characteristic	Explanation
$x_{1,1}$	TC-1	Name of educational institution
$x_{1,2}$	TC-2	Name of educational institution
$x_{1,3}$	2007	Period (all 2007 year)
$x_{1,4}$	2008	Period (all 2008 year)
$x_{1,5}$	Turner	Occupation
$x_{1,6}$	Driver	Occupation
$x_{1,7}$	Cook	Occupation
$x_{2,1}$	10	The quantitative characteristic factual data
$x_{2,2}$	20	The quantitative characteristic factual data
$x_{2,3}$	5	The quantitative characteristic factual data
$x_{2,4}$	7	The quantitative characteristic factual data
$x_{2,5}$	13	The quantitative characteristic factual data
$x_{2,6}$	3	The quantitative characteristic factual data
$x_{2,7}$	7	The quantitative characteristic factual data
$x_{2,8}$	4	The quantitative characteristic factual data
$x_{2,9}$	11	The quantitative characteristic factual data
$x_{2,10}$	6	The quantitative characteristic factual data
$x_{3,1}$	Graduating students	Indicator name
$x_{3,1}$	Set of students on the first course	Indicator name

Before the search has begun, we will set a number of conditions, according to (1):

$$U_2 = \{x_{1,3}, x_{1,5}\} \cup \{x_{3,1}\} = \{x_{1,3}, x_{1,4}, x_{3,1}\}.$$

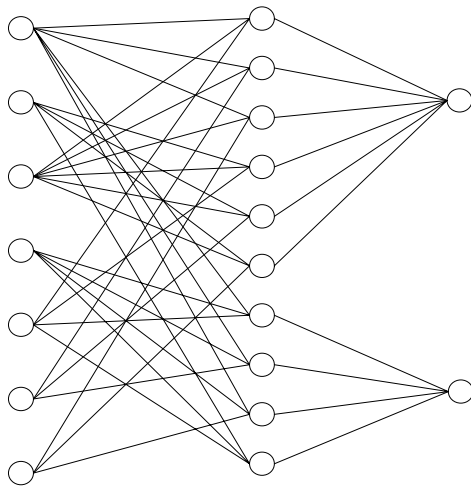


Fig. 8. Example of the factual data on graduates: x – graph node; X – set of nodes

The application of the “traversal of one node vicinities” algorithm:

– select all nodes from set X_2 which have ribs with node $x_{1,3}$ from set U_2 and place them into set:

$$X'_2 = \{x_{2,1}, x_{2,2}, x_{2,3}, x_{2,4}, x_{2,5}, x_{2,6}\};$$

– check each node $x'_{2,i} \in X'_2$ for possessing a full set of ribs with nodes of set U_2 . If the condition is not fulfilled, the node $x'_{2,i}$ is removed from set X'_2 . As a result, in set X'_2 there are the following elements: $\{x_{2,1}, x_{2,3}\}$.

Element values of set X'_2 will be taken from table. To calculate the complex indicator – “quantity of graduates in 2007” we will sum the elements of the obtained subset and receive the required value – “15”.

Applying the algorithm: “Mark of nodes”:

– $i = 1$. Choose node $u_{2,i}$ from set U_2 : ($u_{2,1} = x_{1,3}$).

We will introduce a set of labels M the dimensions of which are equal to $|X_2| = 10$; and we will appropriate “0” to each element of this set;

– for node $u_{2,i}$ we find set $\Gamma(u_{2,i}) \subset X_2$ for $\forall i$. For each $x_{2,j} \in X_2$ for which $x_{2,j} \in \Gamma(u_{2,i})$ the condition is satisfactory, we will increase corresponding label $m_j \in M$ value by “1”.

$$i = 1: u_{2,1} = x_{1,3}, \Gamma(u_{2,1}) = \{x_{2,1}, x_{2,2}, x_{2,3}, x_{2,4}, x_{2,5}, x_{2,6}\}, \\ M = \{1, 1, 1, 1, 1, 0, 0, 0, 0\}$$

$$i = 2: u_{2,2} = x_{1,5}, \Gamma(u_{2,2}) = \{x_{2,1}, x_{2,3}, x_{2,10}\}, \\ M = \{2, 1, 2, 1, 1, 0, 0, 0, 1\}$$

$$i = 3: u_{2,3} = x_{3,1}, \Gamma(u_{2,3}) = \{x_{2,1}, x_{2,2}, x_{2,3}, x_{2,4}, x_{2,5}, x_{2,6}\}, \\ M = \{3, 2, 3, 2, 2, 2, 0, 0, 1\};$$

– find a node in set U_2 with the least quantity of ribs, in case $x_{1,5}$ and move the nodes connected to it from X_2 to set $X'_2 = \{x_{2,1}, x_{2,3}, x_{2,10}\}$;

– in the resultant set X'_{res} from set X'_2 we select the nodes having a full set of ribs with nodes from set U_2 : $X'_{res} = \{\forall x_{2,i} \in X'_2 \mid m_i = |U_2|\} = \{x_{2,1}, x_{2,3}\}$.

The complex indicator is calculated similarly to the previous example and equals 15.

The presented approach for representing dynamic data structures allows designing information systems with changeable information structures of the subject domain objects, and its processing on region level. The algorithms

$x_{1,1}$

$x_{2,1}$

$x_{2,2}$

$x_{2,3}$

described in article, can be used to project automated systems with dynamic structures, which will be built on the basis of the proposed approach; for calculating complex indicators when processing statistical information.

The development of the approach is planned in the following ways:

– the revealing of the structure representing the indicator graph, features of its construction and the traversal algorithms;

– researching invariable database control system storage methods for graph indicator, qualifiers, and factual data; working out techniques to work with account elements in these systems;

– researching possible automatic processing ways for factual data: in order to reveal doubtful data, new complex indicators, and new data classes for further analysis.

References

1. Post-relational SUBD Cache 5. Objective-orientated development of applications / V. Kirsten, M. Iringer, M. Kjun, B. Rerig. 2 ed. Moscow : LLC «Binom-Press», 2005.
2. Kite T. Oracle for professionals. Saint-Petersburg : DiaSoftUP, 2005.
3. Faulmer M., Scot K. UML. Basics. Saint-Petersburg : Symbol-Plus, 2002.
4. Volkov V. N., Denisov A. A. System theory : textbook. Moscow : Higher School, 2006.
5. Shovkun A. V. Constructing a corporative informative-analytical system in conditions of constantly changing business // Scientific-technical information. Moscow : VINITI, 2004. № 9. P. 1–6. Series 1.
6. Harari F. Graphs theory. Moscow : Editorial URS, 2003.
7. Anany V., Levitin A. Algorithms: Introduction to the design and analysis of algorithms. Moscow : Williams, 2006. P. 189–195.

© Yakunin Yu. Yu., Gorodilov A. A., 2010

O. V. Zaitsev

JSC Russian Space Systems, Russia, Moscow

THE APPLIED METHOD FOR CARRIER FREQUENCY RESTORATION OF THE TELEMETRY SYSTEMS SIGNAL BY DIGITAL PROCESSING

The paper considers the issue of restoring the level of the telemetry signal carrier frequency at digital processing in the automatic carrier control tract and the calculation of the threshold for taking the decision about the validity of received information symbol from the spacecraft and carrier rocket. We have described the applied method and algorithm for calculating the level of the carrier frequency and the level of threshold for making a solution based on histogram processing of the signal from the output of the frequency detector.

Keywords: control system, signal processing, telemetry.

The control of flight task execution by class SC/CR onboard systems is performed based on the processing of telemetric data (TMD) on the status of most units and devices of the object [1]. For a satisfactory receipt and processing of TMD achieving proper ground means that the following parameters are required for positive detection of received char (“0” or “1”): the restoration of radio signal carrier frequency and the calculation of decisions making the threshold.

Currently, the fast Fourier transformation (FFT) or data parameter linear filtering procedures are usually applied to repair carrier frequency of carrier-shift radio signal [2]. The use of FFT requires a large number of processing operations to get a necessary result, and consequently, a significant time to analyze the signal. The linear filtering (LF) usage leads to complicity for unbiased signal estimate under processing. Besides, the usage of conventional approaches (FFT and LF) for signal processing on a significant noise background does not allow positive detecting of carrier frequency radio-signals [3]; this considerably decreases the sensitivity of radio receiver digital channels.

Alternatives to FFT and LF correction carrier frequency procedures and algorithms are still insufficiently presented in Russian and foreign studies [4–6]. Therefore a development of procedures and algorithms (P&A) for carrier frequency correction (CFR) based on the histogram procedure is considered to be rather urgent, since this method would require considerably less computing sources than FFT.

The CFR procedure presented in this paper is developed based of the histogram procedure. The point of the procedure is the following: the range of possible signal levels at the detector’s output frequency is divided into an optimal number of control levels or intervals. The signal beyond the frequency detector output is transferred to histogram level construction units (HLCU). The amplitude of the signal is compared to the value of the control levels in HLCU; the number of values for the signal amplitude within the interval $\Delta_k = \frac{A_k - A_{k-1}}{2}$ is registered. A_k and A_{k-1} – values of neighbor control levels. A one-dimensional vector (vector of the amount values within interval $\Delta(k)$ is generated –