



**I T H E A**



**International Journal**  
**INFORMATION** **MODELS**  
&  
**ANALYSES**



**2013** Volume 2 Number 3

**International Journal  
INFORMATION MODELS & ANALYSES  
Volume 2 / 2013, Number 3**

Editor in chief: **Krassimir Markov** (Bulgaria)

<b>Adil Timofeev</b>	(Russia)	<b>Levon Aslanyan</b>	(Armenia)
<b>Albert Voronin</b>	(Ukraine)	<b>Luis Fernando de Mingo</b>	(Spain)
<b>Aleksey Voloshin</b>	(Ukraine)	<b>Liudmila Cheremisinova</b>	(Belarus)
<b>Alexander Palagin</b>	(Ukraine)	<b>Lyudmila Lyadova</b>	(Russia)
<b>Alexey Petrovskiy</b>	(Russia)	<b>Martin P. Mintchev</b>	(Canada)
<b>Alfredo Milani</b>	(Italy)	<b>Nataliia Kussul</b>	(Ukraine)
<b>Anatoliy Krissilov</b>	(Ukraine)	<b>Natalia Ivanova</b>	(Russia)
<b>Avram Eskenazi</b>	(Bulgaria)	<b>Natalia Pankratova</b>	(Ukraine)
<b>Boris Tsankov</b>	(Bulgaria)	<b>Nelly Maneva</b>	(Bulgaria)
<b>Boris Sokolov</b>	(Russia)	<b>Olga Nevzorova</b>	(Russia)
<b>Diana Bogdanova</b>	(Russia)	<b>Orly Yadid-Pecht</b>	(Israel)
<b>Ekaterina Detcheva</b>	(Bulgaria)	<b>Pedro Marijuan</b>	(Spain)
<b>Ekaterina Solovyova</b>	(Ukraine)	<b>Rafael Yusupov</b>	(Russia)
<b>Evgeniy Bodyansky</b>	(Ukraine)	<b>Sergey Krivii</b>	(Ukraine)
<b>Galyna Gayvoronska</b>	(Ukraine)	<b>Stoyan Poryazov</b>	(Bulgaria)
<b>Galina Setlac</b>	(Poland)	<b>Tatyana Gavrilova</b>	(Russia)
<b>George Totkov</b>	(Bulgaria)	<b>Valeria Gribova</b>	(Russia)
<b>Gurgen Khachatryan</b>	(Armenia)	<b>Vasil Sgurev</b>	(Bulgaria)
<b>Hasmik Sahakyan</b>	(Armenia)	<b>Vitalii Velychko</b>	(Ukraine)
<b>Iliia Mitov</b>	(Bulgaria)	<b>Vladimir Donchenko</b>	(Ukraine)
<b>Juan Castellanos</b>	(Spain)	<b>Vladimir Ryazanov</b>	(Russia)
<b>Koen Vanhoof</b>	(Belgium)	<b>Yordan Tabov</b>	(Bulgaria)
<b>Krassimira B. Ivanova</b>	(Bulgaria)	<b>Yuriy Zaichenko</b>	(Ukraine)

IJ IMA is official publisher of the scientific papers of the members of  
the **ITHEA® International Scientific Society**

IJ IMA rules for preparing the manuscripts are compulsory.

The **rules for the papers** for ITHEA International Journals are given on [www.ithea.org](http://www.ithea.org).

The camera-ready copy of the paper should be received by ITHEA® Submission system <http://ij.ithea.org>.

Responsibility for papers published in IJ IMA belongs to authors.

General Sponsor of IJ IMA is the **Consortium FOI Bulgaria** ([www.foibg.com](http://www.foibg.com)).

**International Journal "INFORMATION MODELS AND ANALYSES" Vol.2, Number 3, 2013**

Edited by the **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria, in collaboration with

Institute of Mathematics and Informatics, BAS, Bulgaria,

V.M.Glushkov Institute of Cybernetics of NAS, Ukraine,

Universidad Politecnica de Madrid, Spain,

Hasselt University, Belgium

Institute of Informatics Problems of the RAS, Russia,

St. Petersburg Institute of Informatics, RAS, Russia

Institute for Informatics and Automation Problems, NAS of the Republic of Armenia,

and Federation of the Scientific - Engineering Unions /FNTE/ (Bulgaria).

Publisher: **ITHEA®**

Sofia, 1000, P.O.B. 775, Bulgaria. [www.ithea.org](http://www.ithea.org), e-mail: [info@foibg.com](mailto:info@foibg.com)

Technical editor: **Ina Markova**

**Printed in Bulgaria**

**Copyright © 2012-2013 All rights reserved for the publisher and all authors.**

© 2012-2013 "Information Models and Analyses" is a trademark of Krassimir Markov

© ITHEA is a registered trade mark of FOI-Commerce Co.

**ISSN 1314-6416 (printed)**

**ISSN 1314-6424 (CD)**

**ISSN 1314-6432 (Online)**

---

## CLASSIFICATION OF INCOMPLETE DATA

Vasily V.Ryazanov, Vladimir V.Ryazanov

**Abstract:** *The problem of reconstructing the feature values in samples of objects given in terms of numerical features is considered. A numerical study of different approaches of solving this problem on one model and three practical problems at different levels of data incompleteness is held. A modification of the model calculation of estimates, not requiring metrics for signs, is suggested. The advantage of a local and recognition approaches over filling gaps with sample averages is shown.*

**Keywords:** *data mining, missing data, classification, pattern recognition.*

**ACM Classification Keywords:** *I.5 Computing Methodologies – Pattern recognition*

---

### Introduction

---

In many problems of data analysis, supervised or unsupervised classification, regression problems as the study data samples  $X = \{\mathbf{x}_i, i = 1, 2, \dots, m\}$  are used, where  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{in})$  - is the feature description of some object,  $x_{ij} \in R$ .

In this paper, we consider the case when part of feature values  $x_{ij}$  of objects is unknown (unknown values are denoted  $\Delta$ ). Various approaches to restore the feature values are possible. Some of them take into account the kind of task (clustering, classification or regression), probabilistic characteristics of features, additional prior knowledge and hypotheses.

Currently, there are different approaches to solve the problem of restoration missing data which are conventionally divided into «marginalization», «imputation» и «projection».

In case of «marginalization» or «skipping incomplete objects» incomplete descriptions are simply excluded from the sample. As a result a new sample of complete descriptions is formed [Little, Rubin, 1987]. In this case, obviously, a lot of information can be lost.

In case of «imputation» the most suitable value for a gap is evaluated from the entire sample. General methodology of «imputation» uses the sample means, median, random choice [Little, Rubin, 1987; Zloba, 2002]. In paper [Morin, Raeside, 1981] nearest neighbor algorithm was used. Recently Zhang [4] proposed a method of «partial imputation». Unknown data is estimated by full description of the object in a small neighborhood of incomplete descriptions. Delavallade and Ha [Delavallade, Dang 2007] proposed a new approach which uses entropy to estimate unknown values.

In «projection»-methods (or «imputation by regression») feature space is reduced by one for each unknown characteristic values. This requires a special construction of the classifier in the reduced space. Usually full description from the training set for construction of an optimal hyperplane are used. In [Honghaj etc., 2005], Honghaj etc. investigated «imputation» - approach with the use of support vector machines (SVM) and conducted its comparison with the replacement of the unknown values with mean, median, average of the next two values, and the values of the nearest neighbor method. The results show that the SVM method showed the highest accuracy in comparison to other methods. The algorithm of filling the gaps with maximum likelihood is well-known and reliable (EM algorithm) [Little, Rubin, 1987]. The disadvantage of the method is low rate of convergence, if a

lot of data is missed. In solving the problem locally optimal solutions are found. A probability model of forming data is assumed reasonable.

In paper [Ryazanov, 2011] three approaches to the restoration of feature values are considered. The first is based on the organization of the iterative procedure of the sequential refinement of missing feature values. Herewith, the analysis of local information for each object with gaps is performed. The second approach is based on solving the optimization problem. Previously unknown feature values are found according to the rule that there is maximal conformity of metric relations between objects in subspaces of known values and their found spaces of full descriptions. In the third approach, each missing feature value is found by solving a series of problems of recognition.

In this paper, a numerical study of the first and third approaches in various model and real practical problems is conducted. Same time a new model to detect unknown feature values is suggested. This model doesn't require presence of metrics for features. In the numerical experiments for comparison the «means» method is used (replacement of unknown feature values with sample averages).

During restoring unknown feature values the following ideas are implemented:

- all training data objects regardless of the number of gaps in them are used;
- no probabilistic assumptions about the data sample are used;
- the only initial information is sample data;
- features in general are not independent and metric.

We assume that  $x_{ij} \in \{M_j, \Delta\}$ ,  $M_j \subseteq R$ ,  $R$  - set of real numbers with the metric  $\rho(a, b) = |a - b|$ .  $M_j$  - finite set of possible values of  $j$  - feature, which consists of all of its values recorded at the training data. Let the set of pairs of indices  $J$  sets all unknown feature values of the objects of the training set  $J = \{\langle i, j \rangle, i = 1, 2, \dots, m, j = 1, 2, \dots, n : x_{ij} = \Delta\}$ . The task of restoration of unknown feature values is to find a sample  $X^* = \{\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_m^*\}$  of full descriptions  $\mathbf{x}_i^* = (x_{i1}^*, x_{i2}^*, \dots, x_{in}^*)$ ,

$x_{ij}^* = \begin{cases} x_{ij}, & x_{ij} \neq \Delta, \\ \in M_j, & x_{ij} = \Delta, \end{cases}$  which «maximum corresponds» to sample set of partial descriptions  $X$ .

## 1. Local method of restoration feature values

The idea is simple. Unknown values of feature should be similar to the known values of a small neighborhood of the object. Firstly, all unknown feature values are filled with random numbers from the feature tolerance range  $x_{ij} \in M_j, j = 1, 2, \dots, n$ , or are filled based on some a priori considerations. Then unknown values are sequentially modified by a combination of  $k$  - neighbors method and shift procedure.

Fix a metric in the space of feature descriptions.

Step 0. Initialize the initial  $x_{ij}^{(0)} \in M_j, \forall \langle i, j \rangle \in J$ . Obtain a table full descriptions.

Step  $t=1, 2, \dots$ . We have  $\mathbf{x}_i^{(t-1)} = (x_{i1}^{(t-1)}, \dots, x_{in}^{(t-1)})$ . For each pair  $\langle i, j \rangle \in J$  calculate  $x_{ij}^{(t-1)*}$ , as the average value of the feature with index  $j$  of  $k$  neighbors of  $\mathbf{x}_i^{(t-1)}$  object. Set  $x_{ij}^{(t)} = x_{ij}^{(t-1)} + \theta(x_{ij}^{(t-1)*} - x_{ij}^{(t-1)}), \forall \langle i, j \rangle \in J$ ,  $x_{ij}^{(t)} = x_{ij}^{(t-1)}, \forall \langle i, j \rangle \notin J$ . Here  $k$  is integer,  $1 \leq k \leq m - 1, 0 < \theta \leq 1$  - algorithm parameters. The step is repeated, until the stopping criterion is executed.

As a stopping criterion are used the followings: the maximum number of iterations  $N$ , condition

$$\sum_{\langle i, j \rangle \in J} |x_{ij}^{(t)} - x_{ij}^{(t-1)}|^2 \leq \varepsilon, \text{ etc.}$$

Choice of metric, parameters  $\theta$ ,  $N$ ,  $\varepsilon$ ,  $k$ , initial values of the unknown features is made by the user.

## 2. Optimization method of restoration feature values

The essence of this approach is that missing values should take such values, with which the metric relationships between objects in space of «full descriptions» maximally correspond to the metric relationship in subspaces of «known descriptions» or «more known descriptions».

Let  $\mathbf{x}_i, \mathbf{x}_j$  - any pair of training table rows. Introduce notation:

$$\Omega_i^0 = \{t : x_{it} \neq \Delta\}, \quad \Omega_i^1 = \{t : x_{it} = \Delta\}. \quad \text{Set } \Omega_{ij}^{00} = \Omega_i^0 \cap \Omega_j^0, \quad \Omega_{ij}^{01} = \Omega_i^0 \cap \Omega_j^1, \quad \Omega_{ij}^{10} = \Omega_i^1 \cap \Omega_j^0, \quad \Omega_{ij}^{11} = \Omega_i^1 \cap \Omega_j^1.$$

The Euclidean metric is considered:

$$\rho(\mathbf{x}_i, \mathbf{x}_j) = \left( \sum_{t \in \Omega_{ij}^{00}} (x_{it} - x_{jt})^2 + \sum_{t \in \Omega_{ij}^{01}} (x_{it} - y_{jt})^2 + \sum_{t \in \Omega_{ij}^{10}} (y_{it} - x_{jt})^2 + \sum_{t \in \Omega_{ij}^{11}} (y_{it} - y_{jt})^2 \right)^{\frac{1}{2}}, \text{ where, for}$$

convenience, the unknown values of  $x_{it}$  features are replaced with the variables  $y_{it}$ .

Then  $\rho^+(\mathbf{x}_i, \mathbf{x}_j) = \left( \sum_{t \in \Omega_{ij}^{00}} (x_{it} - x_{jt})^2 \right)^{\frac{1}{2}}$  will be the distance between  $\mathbf{x}_i$  and  $\mathbf{x}_j$  in subspace, in which the values of the features of both objects are known, and

$$\rho^{++}(\mathbf{x}_i, \mathbf{x}_j) = \left( \sum_{t \in \Omega_{ij}^{00}} (x_{it} - x_{jt})^2 + \sum_{t \in \Omega_{ij}^{01}} (x_{it} - y_{jt})^2 + \sum_{t \in \Omega_{ij}^{10}} (y_{it} - x_{jt})^2 \right)^{\frac{1}{2}} - \text{the distance in subspace, in which}$$

the values of each feature for at least one of the objects are known.

In [Ryazanov, 2011] the problem of minimizing the following two criteria of filling the gaps was considered:

$$\Phi(\langle y_{ij} \rangle) = \sum_{\substack{i, j=1 \\ i > j}}^m (\rho(\mathbf{x}_i, \mathbf{x}_j) - N_{ij}^+ \rho^+(\mathbf{x}_i, \mathbf{x}_j))^2,$$

$$F(\langle y_{ij} \rangle) = \sum_{\substack{i, j=1 \\ i > j}}^m (\rho(\mathbf{x}_i, \mathbf{x}_j) - N_{ij}^{++} \rho^{++}(\mathbf{x}_i, \mathbf{x}_j))^2, \text{ where } N_{ij}^+, N_{ij}^{++} - \text{some constants.}$$

This approach is visual and natural, numerical experiments have confirmed its prospects. However, there are significant difficulties in its implementation, and adaptation to the data parameters. These optimization tasks have multiple extremums, so the choice of the initial approximation and analysis of local optima is the subject of a separate study. In addition, with a large number of gaps optimization task has a greater dimension and it should be taken into account in the optimization method. It seems that a practical method for optimizing should significantly depend on the degree of data incompleteness.

## 3. Restoring the features as a solution the problem of recognition

The task of restoring signs is solved sequentially for each pair  $\langle i, j \rangle \in J$  as a special recognition task. Let for object  $\mathbf{x}_i$  from sample  $X$  value  $x_{ij}$  is unknown. Then we denote,

$\Omega_i = \{j_1, j_2, \dots, j_\tau\}$ ,  $\Theta_i = \{k_1, k_2, \dots, k_\sigma\} = \{1, 2, \dots, n\} \setminus \Omega_i$ , where  $\Omega_i$  - set of numbers from  $\{1, 2, \dots, n\}$ , for which  $x_{ij} = \Delta$ . One considers that  $M_j = \{a_1, a_2, \dots, a_N\}$ ,  $a = a_1 < a_2 < \dots < a_N = b$ . The main task is to recognize  $a_t$ , which will be the true value for  $x_{ij}$ .

General algorithm consists of solving  $[\log_2 N] + 1$  dichotomous recognition problems. Assume we have some standard recognition algorithm  $A$ , which according to training set  $X$  of features descriptions solves the problem of classification with two classes  $K_1, K_2$ :  $A: \mathbf{x} \rightarrow \mathbf{K}(\mathbf{x})$ , where  $\mathbf{K}(\mathbf{x}) \in \{K_1, K_2\}$ ,  $\mathbf{x}$  - any object to be recognized. Let's describe the general scheme of the reconstruction algorithm of feature values, based on the solution of dichotomous recognition problems.

### 3.1. General algorithm

1. We have a set of numbers  $a = a_1 < a_2 < \dots < a_N = b$ . Consider two classes:

$$K_1 = \{\mathbf{x} \mid a \leq x_j \leq a_{\lfloor \frac{N}{2} \rfloor}\},$$

$$\tilde{K}_1 = K_1 \cap X, K_2 = \{\mathbf{x} \mid a_{\lfloor \frac{N}{2} \rfloor} < x_j \leq b\}, \tilde{K}_2 = K_2 \cap X.$$

2. Based on the training set  $\{\tilde{K}_1, \tilde{K}_2\}$  recognition algorithm  $A_1: \mathbf{x} \rightarrow \mathbf{K}(\mathbf{x})$ ,  $\mathbf{K}(\mathbf{x}) \in \{K_1, K_2\}$  is constructed.

3. The task of recognition  $\mathbf{x}$  is solved with classes  $K_1, K_2$ .

4. If the training set of class  $\mathbf{K}(\mathbf{x})$  consists only of objects  $\mathbf{x}_\alpha$ , for which  $x_{\alpha t} = a_t$ , we set  $x_{ij} = a_t$  and the feature value is considered to be calculated. Otherwise  $\mathbf{K}(\mathbf{x})$  (and corresponding training objects) analogously 1) is divided into two sets of values  $a_1, a_2, \dots, a_N$ , go to step 2) and the process is repeated for new classes. It is clear that no more than after  $[\log_2 N] + 1$  steps, we obtain the solution.

Let each feature has  $N_j, j = 1, 2, \dots, n$  different values. Then one needs to solve no more than

$$\sum_{j=1}^n ([\log_2 N_j] + 1) \text{ training tasks. In practical calculations was used a model of calculating estimates, which}$$

does not require prior solutions of laborious training tasks.

### 3.2. Algorithms for calculating estimates with non-metric proximity function

Considered further modification of the model calculation of estimates [Zhuravlev, Nikiforov, 1971] can be used for object recognition with ordinal features by changing the proximity function. The set  $R$  is the set of real numbers.

Consider the proximity function of recognized  $\mathbf{x}$  comparatively to the pair of objects  $\mathbf{x}_t$  и  $\mathbf{x}_\tau$ :

$$B_\Omega(\mathbf{x}_t, \mathbf{x}, \mathbf{x}_\tau) = \begin{cases} 1, & (x_{t\beta} \leq x_\beta \leq x_{\tau\beta}) \vee (x_{\tau\beta} \leq x_\beta \leq x_{t\beta}), \forall \beta \in \Omega, \\ 0, & \text{otherwise.} \end{cases}$$

To calculate the estimates the expression (1) is considered:

$$\tilde{\Gamma}_j(\mathbf{x}) = \frac{2}{|\tilde{K}_j|(|\tilde{K}_j| - 1)} \sum_{\mathbf{x}_t, \mathbf{x}_\tau \in \tilde{K}_j, t > \tau} \sum_{\Omega \in \Omega_A} B_\Omega(\mathbf{x}_t, \mathbf{x}, \mathbf{x}_\tau) \quad (1)$$

It can be shown (similar to standard proximity functions [Zhuravlev, Nikiforov, 1971]), that there is a formula of effective computing estimates (1):

$$\tilde{\Gamma}_j(\mathbf{x}) = \frac{2}{|\tilde{K}_j|(|\tilde{K}_j| - 1)} \sum_{\mathbf{x}_t, \mathbf{x}_\tau \in \tilde{K}_j, t > \tau} C_{d(\mathbf{x}_t, \mathbf{x}, \mathbf{x}_\tau)}^k \quad (2)$$

where  $d(\mathbf{x}_t, \mathbf{x}, \mathbf{x}_\tau) = \left| \{ \beta: (x_{t\beta} \leq x_\beta \leq x_{\tau\beta}) \vee (x_{\tau\beta} \leq x_\beta \leq x_{t\beta}) \}, \beta=1, 2, \dots, n \right|$ .

Natural limitation of the model is  $|\tilde{K}_j| \geq 2$ .

Note that the model of computing estimates in task of recognition unknown feature value  $x_{ij}$  can be used «directly», without assigning any values to other unknown feature values in the learning table. For this, during calculation of the quantities  $d(\mathbf{x}_\alpha, \mathbf{x}_t)$  (thereafter,  $d(\mathbf{x}_t, \mathbf{x}_\alpha, \mathbf{x}_\tau)$ ) it is sufficient to consider the cases  $\beta \in \Theta_\alpha \cap \Theta_t$  (thereafter,  $\beta \in \Theta_\alpha \cap \Theta_t \cap \Theta_\tau$ ). During solving the problems of recognition training data the calculation of any parameters is not required.

#### 4. Numerical results

This section presents the results of numerical experiments on simulated data and two real practical problems. Because in fact all methods of reconstructing feature values require initial values, as initial values sample means for known feature values were used («mean» or «average» method [Little, Rubin, 1987]). This choice was driven by good results of given initial approximation and ability to compare the recognition results with the original averages. In addition, with a large number of gaps, random selection of these gaps results in a large scatter of reconstructed values and, therefore, leads to new task of choosing found values.

The experiments for each task were performed in the following way. Based on the training sample, sample with gaps was built. Each line of a new sample has  $\alpha\%$  unknown features. This choice was performed randomly according to uniform distribution law. Other feature values in new sample were the same as in initial sample. Then unknown feature values were restored as the sample means.

Then they were «refined» with local method (method №1) or by solving the tasks of recognition (method №3). Optimization method was not considered, some preliminary results of its application are considered in [Ryazanov, 2011]. As a criterion for the solution of the restoration task were used two approaches: calculation of the quantity

$$\Psi_1(\alpha) = \sum_{(i,j) \in J} |x_{ij} - x_{ij}^*| / |J| \text{ or } \Psi_2(\alpha) = \sqrt{\sum_{(i,j) \in J} (x_{ij} - x_{ij}^*)^2} / |J|, \text{ where } x_{ij}^* - \text{restored feature values. The}$$

second approach is to measure the accuracy of the solution of the supervised classification problem on the recovered data.

##### 4.1. Restoring gaps for model problem

The model problem has been formed as a mixture of random samples, obtained according to a normal distribution with 10 independent features. The advantage of the model problem is that we know the configuration of classes. Relatively «complicated» model example was considered. The first class is the union of four samples with 25 objects in each, features of which were created by normal distributions with means, respectively, 1, 5, 9, 13, and a variance 0.5. The second class was formed similarly, except that the features had mathematical expectations 3, 7, 11, 15. Visualization of the original sample of 200 objects is shown in Fig. 1. (rendering algorithm is described in [O Duda, Hart, 1973]). Figure 2 shows a visualization of a sample in which 50% of the random features in each object are replaced with sample mean for the sample. Figure 3 shows a visualization of the sample with missing data after unknown features values restoration. We see that the reconstructed data has a smaller spread in class than after replacing gaps with averages. In Tab. 1 the results of recognition accuracy under different levels of incompleteness  $\alpha$  are shown. Each time a new sample was used. Implementation of well-

known methods was used in system RECOGNITION [Zhuravlev etc., 2006]: « k-nearest neighbors » (k-neighbors), « linear machine » (LM), « binary decision trees» (BT), «support vector machines » (SVM), «logical regularities» (LR) [Ryazanov, 2007]. During the experiments, the task to select control parameters of programs in order to build the best algorithms and further comparing algorithms hasn't been set. Values of their control parameters were usually chosen "by default". The aim was to compare different methods of reconstruction feature values and to study the behavior of the separate method depending on the level  $\alpha$ . The recognition accuracy (the percentage of correctly recognized objects in the sliding mode control) of methods SVM and LM on the original samples has amounted, respectively, 75.0% and 54.5%. Therefore, on this problem were considered only « k- neighbors » methods (100%), «BT» (99.5%), «LR» (96.0%). Note that the number of nearest neighbors of the "default" is 3, which explains the high accuracy of the method in this example.

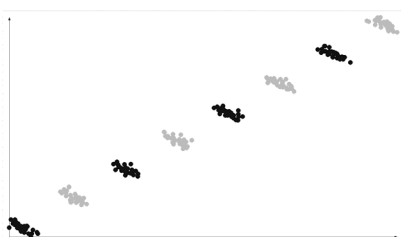


Figure1. Initial sample

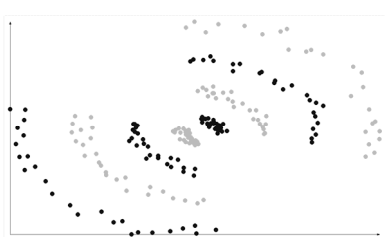


Figure 2. Replacement gaps with feature averages

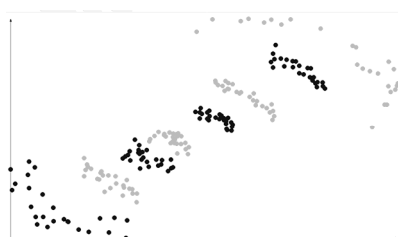


Figure 3. Restored data

method/ $\alpha$	10			20			30			40		
	avr.	loc.	recog.	avr.	loc.	recog.	avr.	loc.	recog.	avr.	loc.	recog.
$\kappa$ -neighbors	90.0	100	100	94.0	100	100	82.0	97.0	100	70.5	98.0	99.5
BT	94.0	99.0	92.5	93.0	98.5	88.0	91.5	97.5	94.0	84.5	97.0	90.0
LR	72.5	98.0	86.5	50.5	97.0	81.0	52.5	95.0	81.0	51.5	94.5	85.5

method/ $\alpha$	50			60			70			80		
	avr.	loc.	recog.	avr.	loc.	recog.	avr.	loc.	recog.	avr.	loc.	recog.
$\kappa$ -neighbors	67.5	97.0	97.0	64.5	90.5	84.5	66.5	87.0	66.5	63.5	78.0	72.8
BT	90.0	94.0	89.5	84.5	88.5	88.0	79.0	78.0	83.0	80.0	71.0	73.5
LR	58.5	86.5	85.0	47.0	80.0	74.0	44.0	64.0	64.0	55.5	66.5	70.0

Table 1. Recognition accuracy after the restoration of feature values

The use of recovery methods 1) and 3) significantly improves the accuracy comparatively to filling with averages. In gray are marked only those rare results when accuracy decreases. We see that for small and medium  $\alpha$  the best accuracy is obtained by the "k-neighbors", which is not surprising due to compactness of the subclasses. The method of "LR" shows significantly higher results after application of 1) and 2) because the compactness of classes is considerably improved.

Mean values for the three methods of recognition accuracy are shown in Figure 4.

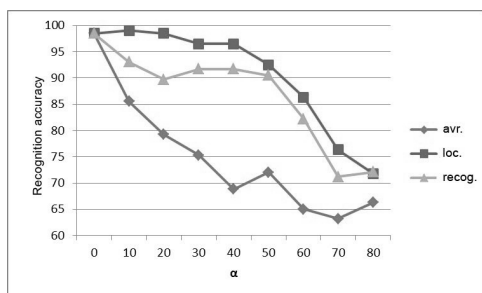


Figure 4. The average recognition accuracy on methods

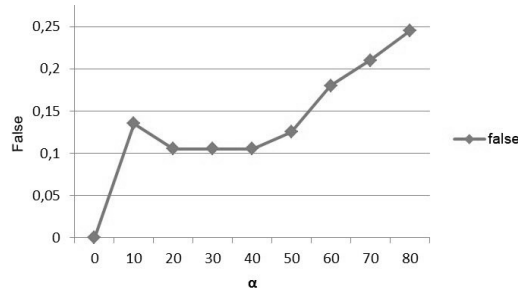


Figure5. Dependence the criterion  $\Psi_2(\alpha)$  at model data



During replacing the gaps with mean values occurs that  $\Psi_2(\alpha) \approx 0.26$ . Dependence of the average values for the algorithm  $\Psi_2(\alpha)$  is shown in Fig. 5. On other tasks, this relationship had a similar appearance.

#### 4.2. Restoration of gaps on the example of "wine" task

Original data of this and future applications are taken from [Frank, Asuncion, 2010]. This problem was set by Forina, M. et al. (PARVUS - An Extendible Package for Data Exploration, Classification and Correlation. Institute of Pharmaceutical and Food Analysis and Technologies, Via Brigata Salerno, 16147 Genoa, Italy. We study a sample of 178 objects, which are the results of chemical analysis, expressed in 13 features, such as content of alcohol, malic acid, magnesium, hue and others. The sample is divided into three classes (59, 71 and 48 objects), corresponding to 3 types of wine made from grapes grown in the same region of Italy.

Visualization sample is shown in Fig. 6-8 is similar to visualization of model problem.

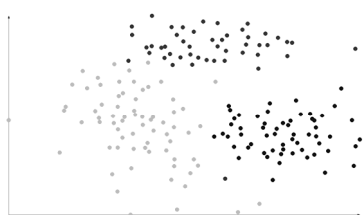


Figure 6. Initial sample

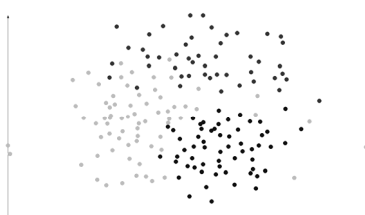


Figure 7. Replacement gaps with feature averages

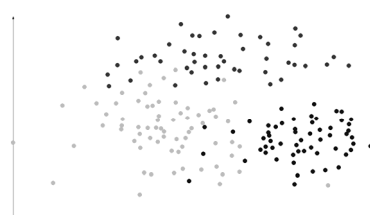


Figure 8. Restored data

In Tab. 2 the recognition accuracy under different levels of incompleteness  $\alpha$  is shown.

метод\ $\alpha$	0	25			50			75		
		avr.	loc.	recog.	avr.	loc.	recog.	avr.	loc.	recog.
k-neighb.	97.2	84.8	93.8	96.1	83.1	83.1	89.3	76.4	75.8	79.2
LR	97.8	95.5	96.6	94.4	82.0	86.0	86.0	83.1	79.2	78.7
SVM	97.8	94.9	97.8	96.6	89.3	91.6	92.1	87.1	79.8	76.4
BT	94.9	86.0	92.7	93.8	62.9	77.0	86.0	76.4	70.2	68.5
LR	94.9	86.5	96.6	92.7	69.1	86.0	88.2	0.00	79.8	66.3

Table 2.

It can be seen that for small and medium  $\alpha$  precision of recovery with methods 1) and 3) for subsequent recognition tasks is better than with the averaging method. However, with high level of gaps (here is the case of 75%) averages for the features are, apparently, more natural. The average value of accuracy here lowers the presence of LR method. During recovering feature values by the method of «mean» with a large number of gaps logical regularities in general are not found. Meanwhile, the average recognition accuracy without LR method is 80.5%.

#### 4.3. Restoration of gaps on the example of "home" task

This problem of recognition the value of housing in the suburbs of Boston on set of 13 numerical signs (the crime rate in the city, the concentration of nitrogen oxides, the average number of rooms in the home, the age of the home, the weighted distance to five employment centers, etc.) is actually the problem of restoration of regression [Harrison, Rubinfeld, 1978]. The problem of classification with 5 classes was considered by dividing the cost of the house at intervals. Each class included, respectively, 16, 91, 93, 27 and 15 objects. The specifics of this problem was in "ordering of classes." The accuracy of recognition by the five methods is presented in Tab. 3:

method\α	0	15			30			40			50		
		avr.	loc.	recog.	avr.	loc.	recog.	avr.	loc.	recog.	avr.	loc.	recog.
κ-nb.	66.9	62.8	66.5	57.4	57.0	55.8	56.2	49.0	58.3	41.7	51.2	42.6	39.7
LR	73.6	68.6	66.9	67.4	65.7	65.7	67.4	62.8	58.7	57.9	59.9	48.3	55.0
SVM	74.0	67.8	70.7	70.7	66.5	67.8	64.9	54.1	61.2	55.4	54.9	57.9	61.6
BT	65.3	65.7	64.9	59.5	53.7	61.6	64.5	55.8	53.3	61.2	50.0	54.1	47.1
LR	64.5	62.0	63.6	62.8	61.2	54.5	58.3	49.2	55.4	52.1	50.4	54.5	53.3

Table 3.

Here, the quality of all methods is about the same, due, apparently, to simplicity of the task.

#### 4.4. Restoration of gaps on the example of "image" task

The problem of classification 7 types of images ("heaven," grass ", " cement ", etc.) was considered. Data source: Vision Group, University of Massachusetts, Carla Brodley. Image Segmentation data. Each class was represented by 30 precedents, each of which was described by a set of 19 numerical features. The accuracy of recognition with five methods is shown in Tab 4, and the average accuracy graph (of recognition methods) is shown in Fig. 9:

method\α	0	10			25			50		
		avr.	loc.	recog.	avr.	loc.	recog.	avr.	loc.	recog.
κ-nb.	87.6	81.4	83.3	77.6	73.8	79.5	82.4	63.3	68.6	66.8
LR	88.1	81.9	83.3	82.4	76.2	82.4	75.4	60.3	68.9	62.7
SVM	90.0	83.8	86.2	83.8	76.2	83.8	83.8	63.5	70.6	69.0
BT	84.8	85.2	84.8	82.9	70.5	82.4	76.2	59.4	69.0	59.7
LR	87.6	82.9	89.0	85.2	75.2	85.2	80.0	29.5	68.6	63.3

Table 4.

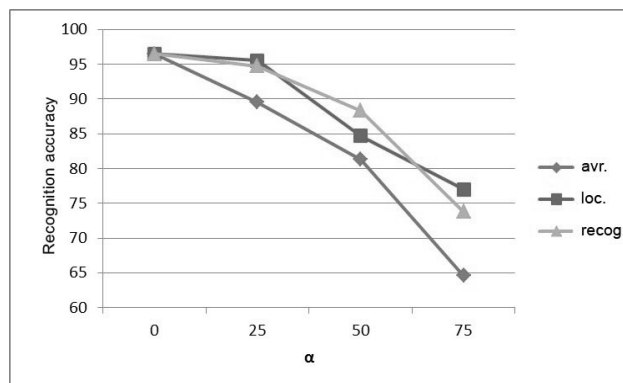


Figure 9

## Conclusion

Experiments conducted confirm a priori expectations. The accuracy of reconstruction of feature values with methods 1, 3, is generally higher than simple averaging features. This is confirmed by both criteria: higher values of the classification accuracy, and lower values of the mean square deviation of feature values. The ratio of number of tasks where the classification accuracy of the solution after application of 1<sup>st</sup> and 3<sup>rd</sup> method was higher than accuracy of solving similar problems after averaging feature values is 107:37. Note that the thesis "the more gaps the worse recognition" is sometimes not fulfilled, due to small sample sizes and conducting of each experiment on a new sample. Creation of new algorithms to recover unknown feature values is important.

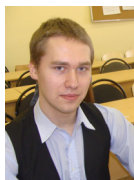
Having a set of  $N$  different recovery algorithms  $A^i, i = 1, 2, \dots, N$ , we can build sets  $\{\mathbf{x}^i, i = 1, 2, \dots, N\}$  of complete descriptions, corresponding to the original incomplete description  $\mathbf{x}$ . Then, solving the problem of classification for each of the objects  $\{\mathbf{x}^i, i = 1, 2, \dots, N\}$  and using the procedure of constructing collective decisions, we can construct the final classification for  $\mathbf{x}$ . It appears that the collective classification accuracy of  $\mathbf{x}$  is higher than the accuracy of the classification of individual  $\{\mathbf{x}^i, i = 1, 2, \dots, N\}$ .

This work was supported by the Program of the Presidium of RAS № 15 and № 5 "Basic Sciences - Medicine", program №2 of Department of Mathematical Sciences RAS, RFBR projects 12-01-00912, 12-01-90012 Bel\_a, 11-01-00585.

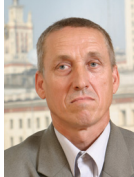
## Bibliography

- [Delavallade, Dang, 2007] T. Delavallade and T.H. Dang. Using Entropy to Impute Missing Data in a Classification Task. In: IEEE International Conference on Fuzzy Systems, London, 1-6, 2007.
- [Frank, Asuncion, 2010] A. Frank and A. Asuncion. UCI Repository of machine learning databases. In: <http://www.ics.uci.edu/~mlern/MLRepository.html>. Irvine, CA: University of California, Department of Information and Computer Science, 2010.
- [Harrison, Rubinfeld, 1978] Harrison, D. and Rubinfeld, D.L. Hedonic prices and the demand for clean air, J. Environ. In: Economics & Management, vol.5, 81-102, 1978.
- [Honghaj etc., 2005] F. Honghaj, C. Guoshun, Y. Cheng, Y. Bingru and C. Yumei. A SVM Regression Based Approach to Filling in Missing Values. In: LNCS - Knowledge-Based Intelligent Information and Engineering Systems, Springer Berlin - Heidelberg, vol. 3683, 581-587, 2005.
- [Little, Rubin, 1987] R.J.A. Little, D.B. Rubin. Statistical Analysis with Missing Data. In: Wiley, New York, 1987.
- [Morin, Raeside, 1981] R. L. Morin and D. E. Raeside. A reappraisal of distance-weighted k-nearest neighbor classification for pattern recognition with missing data. In: IEEE Transactions on Systems, Man and Cybernetics 11(3): 241–243, 1981
- [O Duda, Hart, 1973] Richard O.Duda, Peter E.Hart. Pattern Classification and Scene Analysis. In: A Wiley-Interscience Publication John Wiley&Sons, 1973.
- [Ryazanov, 2007] V.V. Ryazanov. Logical Regularities in Pattern Recognition (Parametric Approach). In: Computational Mathematics and Mathematical Physics, Vol. 47, No. 10, pp. 1720–1735. © Pleiades Publishing, Ltd., 2007. Original Russian Text © V.V. Ryazanov, published in Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, 2007, Vol. 47, No. 10, pp. 1793–1808, 2007
- [Ryazanov, 2011] V.V. Ryazanov. Some Imputation Algorithms for Restoration of Missing Data. In: "Lecture Notes in Computer Science" (LNCS), vol. 7042, pp. 372-379, 2011.
- [Zhang, 2008] S. Zhang. Parimputation: From imputation and null-imputation to partially imputation. In: IEEE Intelligent Informatics Bulletin, vol. 9(1), 32-38, 2008.
- [Zhuravlev etc., 2006] Yu. I. Zhuravlev, V.V. Ryazanov, O.V. Senko. Recognition. Mathematical methods. The software system. Applications. In: Moscow: Fazis, 2006.
- [Zhuravlev, Nikiforov, 1971] Yu. I. Zhuravlev and V. V. Nikiforov. "Recognition Algorithms based on Estimate Evaluation". In: Kibernetika, No. 3, pp. 1–11, 1971.
- [Zloba, 2002] E. Zloba. Statistical methods of reproducing of missing data. In: Computer Modelling & New Technologies. — Vol.6, No.1 — P. 51-61, 2002

## Authors' Information



**Vasily V. Ryazanov** – Student; Moscow Institute of Physics and Technologies, Russia, 141700, Moscow reg., Dolgoprudny, Institutsky per. 9; e-mail: vasyarv@mail.ru  
Major Fields of Scientific Research: Data Mining, Missing Data, Mathematical models in Economics.



**Vladimir V. Ryazanov** – Head of Department; Institution of Russian Academy of Sciences Dorodnicyn Computing Centre of RAS, Russia, 119991, Moscow, Vavilov's street, 40; e-mail: rvv@ccas.ru  
Major Fields of Scientific Research: Pattern recognition, Data mining, Artificial Intelligence

## ON A LOGICAL REGULARITIES BASED METHOD OF DEFINITE QUALITY OBJECT SYNTHESIS

**Alexander Dokukin**

**Abstract:** *In this paper a method of definite quality object synthesis is proposed. The quality is described only by precedent information, i. e. examples of its presence or absence in similar objects. The method is based on using of special logical regularities which come out of solving a standart recognition task. It is applicable in a certain case of data and cannot be considered a general one.*

**Keywords:** *object synthesis, algebraic approach, pattern recognition.*

**ACM Classification Keywords:** *1.5 Pattern Recognition — 1.5.0 General.*

---

### Introduction

The task which is considered in the present article arises in connection to a general theme of definite structure protein synthesis. Yet we will try to abstract away from the application domain and to refrain from stating any biophysical conclusions and even affixing of the proposed methods to a specific data. The only way we intend to use the data is illustrating recognition quality at some point. Nevertheless, the end goal of this and following articles, that is developing of a method for construction of polipeptide chain with the definite profile of structural information and some restrictions to the amino-acid composition, demands such an interpretation. That is why the proposed methods are designed with additional requirement of interpretability.

---

### Problem stating

In general the synthesis task can be described as following. Let we have a sample of objects which do or don't have some quality. Let we also have some set of restrictions. The problem is finding an object that satisfies the restrictions and has the quality. It is assumed that the initial sample can be strongly incomplete and inconsistent that eliminates simple enumeration. In addition it is considered desirable to find an object that has the target quality "in uppermost degree" for reliability of the solution.

Let's write down the above-said more formally.

**Definition 1.** *A set  $P = \{P_1, \dots, P_k\}$  will be called the alphabet and its elements will be referred to as letters.*

**Definition 2.** *By a word of length  $t$  a vector  $\beta = (\beta_1, \dots, \beta_t)$  will be considered, such that  $\beta_i \in P$ .*

**Definition 3.** *A set  $B_1 \times \dots \times B_t$ , where  $B_i \subset P$  and either  $|B_i| = 1$  or  $B_i = P$  will be called a template of length  $t$ .*

The meaning of a template is quite simple. It is a set of words restricted to a single specific letter in some positions and unrestricted in others.

Let us have a training sample  $\beta_1, \dots, \beta_m$  of words of length  $n$ . Let this sample be split into two classes  $K_1$  and  $K_0$ , that is words having and not having some  $\mu$ -quality correspondingly. Finally, let us have a template  $T$  of length  $n$ .

**Definition 4.** *By the elementary synthesis task we will call a search for the word  $\beta' \in T$ , such that*

$$\beta' = \arg \max_{\beta \in T} d(\beta),$$

where  $d(\beta)$  is an estimate of presence of  $\mu$ -quality in a word  $\beta$ .

**Remark 1.** The task is called elementary since the length of the template is equal to the length of sample words. The synthesis of a word by a longer template some parts of which are marked as having or not having the quality is of interest also. Such a general task is not considered in the present paper.

The key quality of the task is that the specific of the application domain doesn't allow defining the functional  $d$ . That is why it is proposed using the training sample for its calculation. Hereat it seems natural using recognition algorithms for the purpose.

The standard recognition task [Zhuravlev, 1977] is stated as follows. Let us have a training sample  $S_1, \dots, S_m$ , described by vectors of some nature  $S_i = (a_{i1}, \dots, a_{in})$ . The sample is split into  $l$  classes  $K_1, \dots, K_l$  that can overlap in general case. The training sample classification that is vectors  $\alpha_i = (\alpha_{i1}, \dots, \alpha_{il})$  is known. Here  $\alpha_{ij}$  is the value of " $S_i \in K_j$ " predicate. It is required to construct an algorithm  $A$  that can calculate classification of a new object  $S$ .

There is a theorem [Zhuravlev, 1977] among the basic propositions of the algebraic recognition theory stating that any recognition algorithm  $A$  can be represented as a composition  $CB$ . Here  $B$  is a recognition operator calculating some realvalue estimates of class membership

$$\{S_1, \dots, S_m, \alpha_1, \dots, \alpha_m, S\} \xrightarrow{B} (\Gamma_1, \dots, \Gamma_l) \in \mathbb{R}^l,$$

and  $C$  is a decision rule

$$\mathbb{R}^l \xrightarrow{C} (0, 1, \Delta)^l,$$

where  $\Delta$  means denial.

Let's construct a recognition operator  $B$  using synthesis training sample.

**Definition 5.** The functional  $d$  is defined as  $d(\beta) = \Gamma_1(\beta) - \Gamma_0(\beta)$ , where  $\Gamma_1(\beta)$  and  $\Gamma_0(\beta)$ , are estimates of the word  $\beta$  for classes  $K_1$  and  $K_0$  calculated by the operator  $B$  correspondingly.

---

## Recognition algorithm

---

Now it is important to mention that all the experimental research of ideas and algorithms described here is made with quite a specific data. The alphabet consists of 20 letters corresponding to amino-acid residues. The words of length 5 describe information units [Nekrasov, 2004] that are basic segments of polipeptide chain that can have structure forming quality. The sample is divided into two classes by presence of the quality.

During initial experiments a recognition algorithm has been proposed that provides 88,8% recognition quality [Senko et al., 2011] for the task. There is a couple of drawbacks preventing the solution from being used in synthesis task.

First of all, each letter was encoded by ten continuous features. That alone complicates interpretation of any results in terms of application domain. Ideally it is required to exercise synthesis based on a small number of simple and interpretable rules. Consequently a set of simple rules must underlie the recognition algorithm.

Secondly, initial test were carried out with a one thousand objects sample that has been split into two parts for the case. Now we deal with a sample containing at least five hundred thousand information units equally divided into two classes. This quantities themselves prevent use of conventional recognition methods. Besides, there is a five percent overlap of classes that was not the case in initial study.

Thus, we require the recognition algorithm to have the following features:

1. recognition results must be interpretable in terms of application domain,
2. recognition quality must not be much worse than the previously achieved one,

3. the algorithm has to deal with enormous data sets.

We propose using special kind of logical regularities [Ryazanov, 2007] for the task.

The regularities will be sought in form of templates (see Definition 3). Its significance will be determined by Fisher's test [Fisher, 1922] using the approximate formula [Vorontsov, 2007].

**Definition 6.** Let's consider a template  $T$ . We will call it a logical regularity of class  $K_1$  with significance  $P(T)$  if

$$P(T) = -\ln \left( \frac{C_{X+Y}^{x+y}}{C_X^x C_Y^y} \right) > 0 ,$$

where  $X = |\{\beta_1, \dots, \beta_m\} \cap K_1|$ ,  $Y = |\{\beta_1, \dots, \beta_m\} \cap K_0|$ ,  $x = |T \cap K_1|$ ,  $y = |T \cap K_0|$ .  
Class  $K_0$  regularities are defined the same way.

Let  $R_1$  and  $R_0$  be sets of logical regularities of classes  $K_1$  and  $K_0$  correspondingly.

**Definition 7.** We will call an estimate  $\Gamma_s(\beta)$  of the word  $\beta$  for  $K_s$  class the value

$$\Gamma_s(\beta) = \sum_{U \in \{T_k \in R_s \mid \beta \in T; T_i \not\subset T_j, i \neq j\}} P(U) .$$

**Remark 2.** As an estimate for the class a word gets sum of significances of all satisfied regularities of that class minus sum of significances of satisfied regularities of the other one. Hereat if the word satisfies two regularities one of which contains the other the latter's significance is not taken into account.

The described method satisfies all the requirements we demand from it.

First, the templates of the proposed structure allow determining dependencies between letters in different positions in a most explicit way. Nested templates then can be used to refine those dependencies if needed.

Second, the structure of logical regularities allows performing complete enumeration of templates for the required amounts of data.

**Statement 1.** Let  $n$ ,  $m$  and  $p$  be the sizes of template, training sample and alphabet correspondingly. There is an algorithm that calculates values  $x$  and  $y$  of Definition 7 for all possible regularities in  $C$  operations, where

$$C = n2^{n-1}m .$$

Hereat  $2((p+1)^n - 1)$  integer values is enough to store all the required counters.

Finally, we have achieved 78,4% recognition quality after testing the method on an independent subset that was equal to training sample in size (the whole set of data had been divided into two equal subsets for training and testing). The quality is achieved with significance threshold 20. On the one hand, the quality is ten percent worse than the initial one. On the other, we already mentioned that the new training sample has classes overlapped by 5,2% objects. That is why we consider the result acceptable.

---

## Synthesis algorithm

---

Template sizes in elementary synthesis task also allow searching the solution by complete enumeration. But having in mind possible generalizations we must take into account its extreme growth. That is why we consider important achieving an algorithm less dependent of the template size even for elementary synthesis.

Let's consider a template  $T = B_1 \times \dots \times B_n$ . Let  $B_i = P$ , i. e. all letters are allowed in  $i$ -th position.

**Definition 8.** Let's denote by  $T[i]\gamma$  the refining of template  $T$  in the position  $i$ ,

$$T[i]\gamma = \{\beta \in T \mid \beta_i = \gamma\},$$

where  $\gamma$  is some letter from the alphabet,  $\gamma \in P$ .

We propose the following synthesis algorithm. For each position  $i$  of the template in which the letter is not constricted we determine it by the formula

$$\gamma'_i = \arg \max_{\gamma \in P} \left( \sum_{U \in \{B_k \in R_1 \mid T[i]\gamma \subset B_k\}} P(U) - \sum_{V \in \{B_k \in R_0 \mid T[i]\gamma \subset B_k\}} P(V) \right).$$

It means that each letter will be independantly refined by the value that achieves the highest estimate as refinement of the initial template. The estimate of the refined template is calculated as difference of significances of templates containing it.

**Theorem 1.** Let the functional  $d$  be determined by Definitions 5 and 7. If the proposed algorithm finds a unique solution for each unrestricted letter of template  $T$  then it calculates the optimal solution of the elementary synthesis task (see Definition 4).

**Proof.** Let's prove the theorem ad absurdum. The following is the definition of the optimal solution of the synthesis task  $\beta'$ :

$$\beta' = \arg \max_{\beta \in T} \left( \sum_{U \in \{B_k \in R_1 \mid \beta \in B_k; B_i \not\subset B_j, i \neq j\}} P(U) - \sum_{V \in \{B_k \in R_0 \mid \beta \in B_k; B_i \not\subset B_j, i \neq j\}} P(V) \right).$$

Let there be a position  $h$  in which  $\gamma'_h \neq \beta'_h$ , where  $\gamma'_h$  is achieved by the synthesis algorithm.

Let's consider a set of regularities generating estimates of the optimal solution for the class  $K_s$ .

$$R'_s = \{B_k \in R_1 \mid \beta' \in B_k; B_i \not\subset B_j, i \neq j\}.$$

This set is divided into two parts

$$R'_s = \{U \mid U_h = \beta'_h\} \cup \{U \mid U_h = P\}.$$

Hereat the second item takes part in estimation regardless the letter in position  $h$ . The first one consists of the unique element  $U$  such that  $T[h]\beta'_h \subset U$ . Indeed one such regularity must exist by the theorem condition because the algorithm comes to a solution. Refinements of the regularity in turn do not take part in estimating.

Since the regularity  $U$  must be used in search of  $\gamma'_h$  by the synthesis algorithm definition, we come to a contradiction that proves the theorem. ■

Since we don't have exact criteria for estimating synthesis algorithm quality, that for example can be used in comparison of two methods, we will show some indirect indicators achieved experimentally. For the experiment the same testing sample was used that consisted of the half of initial data. For each word a template was generated by removing two letters in random positions. These templates were used for synthesis without duplicate filtering. Synthesis algorithm was made searching structure forming units if the initial word corresponded to a structure forming one and vice versa.

It is important that in all cases the proposed algorithm has come to a solutions, i. e. the theorem condition was satisfied. Hereat in most cases both letters were changed comparing to the initial ones that seems natural.

Only in 39, 7% of cases the synthesized word was found in the initial sample (among all the five hundred thousand objects) that justifies its name in some kind. Among them in 91, 1% cases the target class matched the class of the found precedent.

---

## Conclusion

In the present paper a synthesis task has been described, i. e. the task of searching for an object of definite quality that is described by a set of precedents. An approach for solving the task has been proposed. An algorithm has been constructed for an important task family.

At the same time the question of formalizing the method quality estimate remains. Its solution as well as different generalization of the task, such as synthesis by the longer template or number of classes growth, are the most obvious ways for development of these results.

---

## Acknowledgements

The paper is supported by RFBR, grant No. 12-01-31166.

The author also thanks A. N. Nekrasov, senior researcher in IBC RAS for the granted data.

---

## Bibliography

Fisher R. A. On the Interpretation of  $X^2$  from Contingency Tables, and the Calculation of P // Journal of the Royal Statistical Society. — 1922. — Vol. 85, No. 1. — Pp. 87–94.

Zhuravlev Yu. I. Correct algebras over sets of incorrect (heuristic) algorithms I (in Russian) // Cybernetics. — 1977. — No. 4. — Pp. 14–21.

Nekrasov A. N. Analysis of the Information Structure of Protein Sequences: A New Method for Analyzing the Domain Organization of Proteins // Journal of Biomolecular Structure and Dynamics. — 2004. — Vol. 21, Iss. 5. — Pp. 615–623.

Ryazanov V. V. Logical regularities in pattern recognition (parametric approach) // Computational Mathematics and Mathematical Physics. — 2007. — Vol. 47, No. 10. — Pp. 1720–1735.

Vorontsov K. V. Lectures on classification algorithms (in Russian). — 2007. — URL:  
<http://www.ccas.ru/voron/download/LogicAlgs.pdf>.

Senko O. V., Nekrasov A. N., Ryazanov V. V., Dokukin A. A. Prediction of Structure forming Properties of Protein with Help of Pattern Recognition Methods // Proceedings of 8-th Open German-Russian Workshop "Pattern Recognition and Image Understanding" OGRW-8-2011. — Pp. 38–40.

---

## Authors' Information



**Alexander Dokukin** — *Computing Centre of Russian Academy of Sciences, researcher, 40 Vavilova St., Moscow, Russia, 119333; e-mail: [dalex@ccas.ru](mailto:dalex@ccas.ru). Major Fields of Scientific Research: Algebraic Approach to Pattern Recognition.*



---

## MODELING OF AN INTELLECTUAL PROBLEM SOLVER BY TRANSFORMATION OF SEMANTIC MODELS

Elena Shalfeeva

**Abstract:** *The method of construction of a problem solver of intelligent software by transformation of its models is presented. The representations of all these models are convenient both for the person and for program processing. Based on this method supporting tools for intelligent software engineering will be constructed.*

**Keywords:** *intelligent software, problem solver, maintenance, hierarchical semantic network.*

**ACM Classification Keywords:** *D.2.11 Software Architectures (Domain-specific architectures)*

---

### Introduction

An automation of intelligent professional activity is widespread sphere of application of efforts of IT specialists. Modern software systems optimize circulation of documents, keep account expenses, form reports and personnel work records. However such program complexes don't include any tools of support of experts' decision-making [Белоцерковцева, 2010].

Successes in automation of support of decision-making of intellectual tasks are related to knowledge-based systems. But in spite of considerable successes in area of the theory and technology of their creation, they didn't take noticeable practical use. As a rule, such systems are created within research, innovative program projects and don't develop further in view of viability problems [Клещев, 2011].

Nowdays the creation of the integrated complex of problem solvers, knowledge bases and tools of their maintenance is expected from automation of an intelligent professional activity [Клещев, 2011; Румянцева, 2007]. Some subsystem which solves some intellectual problem or gives support in decision-making, using the knowledge base, is called here as a problem solver. The integrated complex has to be scalable and controlled [Клещев, 2010]. It demands the integration of the ideas from artificial intelligence with the approaches of modern programming [Шалфеева, 2013; Грибова, 2012]. The new approach to development of the controlled long-living intelligent software with conceptual knowledge base is presented in the [Клещев, 2010]. It is based on the hypothesis of separation of declarative and procedural components in specialized solvers.

The purpose of this research - to develop a method of transformation of domain and functional models of a specialized intellectual problem solver to its declarative design representations. Such method is necessary for creation of tools for development of maintainable intelligent software (ISW).

---

### The Creation of functionality models of a solver

The result of system analysis (for automation of intelligent professional activity) are the set of models of subject domain, system and its subsystems [Клещев, 2012], among them – the domain ontology, the mathematical statement of each intellectual problem, algorithm of execution of each such problem. Domain ontology consists of

two parts: "reality ontology" (the structure of the input and output data) and "knowledge ontology" (the structure of knowledge used by experts).

A developer of mathematical statement of problem, as a rule, defines a method of the problem solution. The example of method of diagnostics in medicine – "search of all possible values of the diseases; for each disease the following is carried out: forming of all possible variants of each cause-and-effect relation development and searching among all of them the variants corresponded to values of patient signs" [Москаленко, 2006]. The analyst or knowledge engineer builds the algorithm of the execution in according to the method of problem solution and the knowledge ontology. The most traditional representation for algorithm is the flowchart. This algorithm can be presented and as step-by-step transition from a compose task to elementary subtasks.

An example. The fragment of the algorithm for diagnostics [Москаленко, 2006] can be presented so:

*"...to check the hypothesis "the patient is healthy";*

*(to organize the Cycle on observed signs;*

*((to check the hypothesis "all observed values of the sign can take place at any healthy patient;*

*to record the value of a sign, if it disproves the hypothesis))*

*to sum up the hypothesis analysis);..."*

Before designing of each subsystem the detailed analysis of its functions is carried out. The functional model of any problem solver can depend on the user requirements also.

Universal approach to software functionality analysis is partitioning [Pressman, 2001]. The subtask of some level (consecutive or having some options of possible actions) breaks into some less complex subtasks. For some of them it can be necessary "cyclic address". Thus, the hierarchical model of subtasks is the reflex of algorithm contents with addition of input and output subfunctions, and probably subfunctions providing some user requirements.

The variant of representation of this model is following. It is advisability to keep hierarchical model of solver's subtasks (HMST) in the form of a hierarchical semantic network (HSN).

During partitioning it is important to distinguish, whether element of hierarchy of some level consists in "simple" submission or in "cyclic" submission, and whether it will be "optional" [Mori, 2002].

Meta-information for HMST (i.e. the structure of HSN containing HMST) can be described as follows (alternative elements of a network, nonempty and any sets of elements are marked here with qualifiers "alt", "set" and "setmm"):

*task {*

*name: string;*

*structuredness: ~alt (the consecutive; with - options)*

*~set subtask*

*{ subordination: ~alt (in "simple" submission;*

*in "cyclic" submission;*

*"chosen on a condition")*

*название: строка;*

*-> structuredness;*

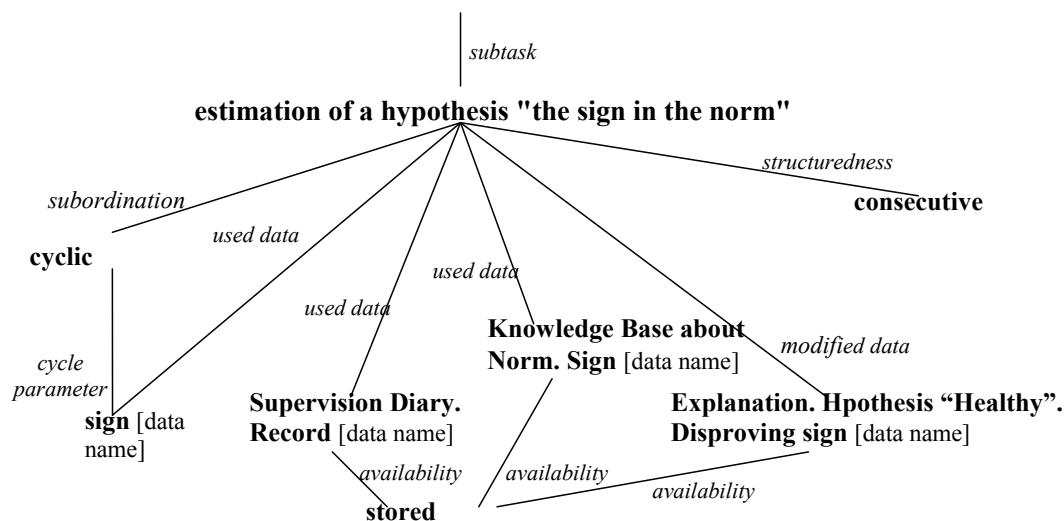
*[~setmm {-> subtask }]}.*

Traditionally the software requirements analysis presents an associations of functions with received and modified elements of information. For the accounting of all processed data of a subsystem every used (or modified) data can be classified to one of three types: stored, received interactively, formed during decision and not-stored.

The expanded model of subtasks (EMsT) enriches HMsT with associations of data. The variant of representation of this model (meta-information for EMsT) is following.

```
task {
  name: string;
  structuredness: ~alt (consecutive; with options)
  ~set subtask
    { subordination: ~alt (in "simple" submission;
      in "cyclic" submission;
      "be chosen on a condition")
      name: string;
      -> structuredness;
      ~set used data { data name: string;
        availability: ~alt { stored, interactive, in process};
      ~set modified data { -> data name; -> availability };
      ~setmm {-> subtask }}}.
```

An example. The HSN fragment for EMsT of a solver of medical diagnostics is presented in the next figure.



**Figure 1.** The HSN fragment for EMsT for a subtask of estimation of a hypothesis "the sign in the norm".

## The Creation of architectural representations of a solver

Software system Designing includes the construction of architectural representations of each subsystem.

In the course of architectural design of rather independent subsystem (such as the solver) designers focus attention, first, on modular decomposition and functionality distribution between these modules (units), and secondly, on control modeling – definition of interaction between units [Pressman, 2001]. Quality of the solver realization and possibility of its maintenance depends on these architectural models [Pressman, 2001, Басс

2006]. Procedures, functions, objects\classes, agents or others can become the units of a specialized solver, depending on approach to its realization.

The construction of a specialized solver by means of call-return architecture can be carried out by known design methods [Pressman, 2001]. In the other approach to architectural designing of the specialized solver processed knowledge and data in the form of ISS, there are two types of units. The first type – the active, rather independent units communicated each other by means of messages (let's call them agents [Грибова, 2012]), the second type – the operations over classes of semantic networks. For such architecture at first the model of necessary is defined agents (according to models of subtasks), then the interactions between units are modelled, further necessary operations are designed.

**Model of all needed units.** The functionality (presented by HMST) has to be distributed by the corresponding quantity of units. It is specified how agents involve each other for cooperation. The agent can be made of two or more structural units (blocks) corresponding to strongly coupled subfunctions.

The variant of representation of this model is following (for above-designated "agents-approach").

```
network of agents {
  root agent :string;
  ~ set cooperation {
    agent-collaborator name :string;{
    agent type ~alt (coordinating, processing, grouping)
    ~ setmm ->cooperation } } }
```

The method of models transformation: root agent is confronted with the main task; agents with two blocks are confronted with "composite" tasks, processing agents - with "leaf" subtasks. For optimization of units number some "leaf" subtasks can be realized as complementary blocks of certain agents.

An example. The HSN fragment for "network of agents" of medical diagnostics solver is presented in the next figure 2.

Here unit for Hypothesis "Healthy" for own performance has to check "normality" of all relevant signs, therefore it (for each of signs) addresses to unit "estimation of a hypothesis "a sign in the norm". After all patient's signs will be estimated, the general conclusion is formulated, whether he is healthy.

Model of control flow between units. The model of agents communication by messages (called by control graph) is intended for its use at a runtime: the processor gets to know about the recipient of the next message from this model. By changing of elements of this model, it is possible to influence a solver's configuration.

The model of control flow between units is being constructed on base of model of subtasks and a network of agents. The method of control graph construction can be follows. For every communication specification one of reusable message templates is selected (or the new template is specified and designed). The agents, confronted to subordinated subtasks, receive the message by "task" template. The agent for a subtask in "simple" subordination, as a rule, forms the response message on the "feedback" template, and in "cyclic" subordination – "for cycle communication" template.

The most convenient representation of control flow between units for person's perception is in the form of graph. However HSN- representation is needed to support uniform storage of all models of a solver. For example, the representation of control graph nodes may be following.

```
~ set agent's block node{
  block name :string;
```

*name of the agent :string;*  
*launching message template : \$ template}.*

An example. Fig. 3 shows some fragment of the control graph defining a scheme for transmitting messages between agents for medical diagnostics problem solver. The agents with several blocks are depicted with multiple entry points.

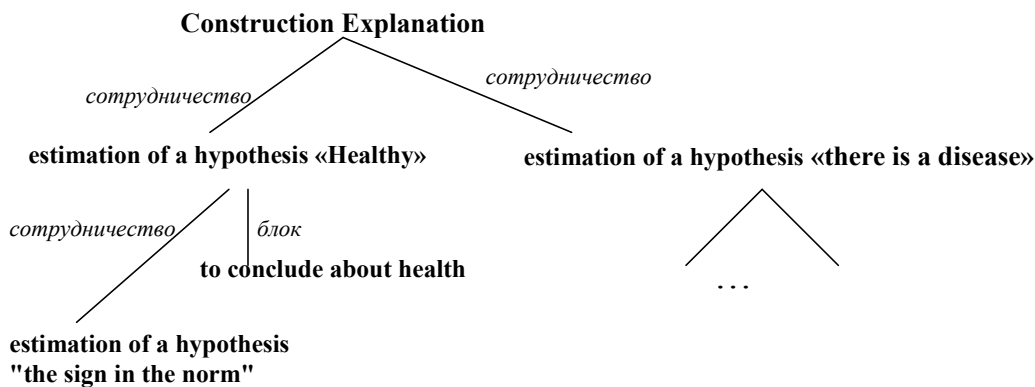


Figure 2. The fragment of the model of all needed units for medical diagnostics solver

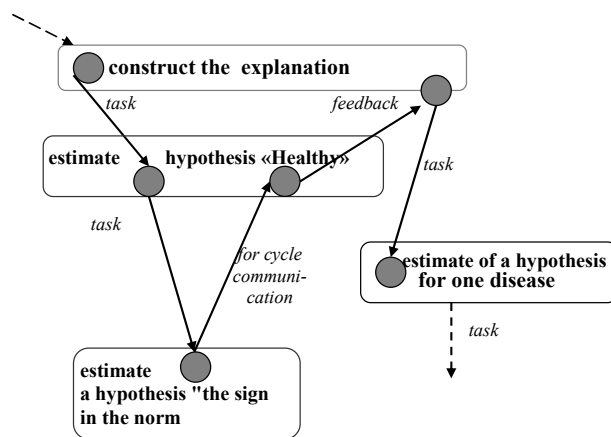


Figure3. The fragment of the control graph for medical diagnostics problem solver

Further designing is aimed to addition in architectural model of information communications of units, associations with information resources (IR). It is aimed to designing of separate units (adapted for a reuse). For ensuring quality of each solver of ISW it is required to construct such models as the design of information communications of agents, the specifications and interfaces design of each agent and designs of each IR with sets of operations of access to these stored IR.

---

## Conclusion

---

For increase of the maintainability of systems, automated the intelligent activity, the new approach to the development of each of its subsystems (intellectual problem solver) has proposed.

This approach is relevant for systems in which knowledge base, the first, are controlled by the experts, secondly, are stored in the conceptual resources (semantic networks), and, the third, are available by software.

---

This approach consists of the representation of results of problem solver development in the form of the declarative models clear to the developer. Architectural models of solvers under construction are realized within the declarative agent programming [Грибова, 2012]. Decision-making agents, which process data or use stored knowledge, have means to access to the semantic networks storing data and knowledge.

The retention of declarative models in the form of semantic networks provides convenient program access to them. Existence of such access to all solver models and existence of a method of transformation of one (semantic) models for construction of the following models are base for creation of those intelligent systems development tools components, which support designing of solvers.

---

### Acknowledgements

---

*This work is performed with financial support of grants of the RFBR № 12-07-00179-a and FEB RUS № 12-I-ОНИТ-04.*

---

### Bibliography

---

- [Mori, 2002] Mori G., Paternò F., Santoro C.: CTTE: support for developing and analyzing task models for interactive system design. IEEE Trans. Softw. Eng. 2002, № 28(8). С. 797–813.
- [Pressman, 2001] Pressman R.S. Software Engineering: Practitioner's Approach. Fifth edition. McGraw-Hill Inc., 2001. 860 p.
- [Басс 2006] Басс Л., Клементс П., Кацман Р. Архитектура программного обеспечения на практике. 2-е изд. – Л.:Питер, 2006. 576 с.
- [Белоцерковцева, 2010] Белоцерковцева Л.Д. Опыт автоматизации деятельности медперсонала перинатального центра города Сургута // Федеральный справочник "Здравоохранение России". Т. 9. 2010. С. 469-474.
- [Грибова, 2012] Грибова В.В., Клещев А.С., Крылов Д.А., Москаленко Ф.М., Тимченко В.А., Шалфеева Е.А. Агентный подход к разработке интеллектуальных интернет-сервисов // Труды Конгресса по интеллектуальным системам и информационным технологиям «IS&IT'12». М.: Физматлит, 2012. Т.1. с. 218-223.
- [Клещев, 2010] Клещев А.С., Грибова В.В. Управление интеллектуальными системами. Известия РАН. Теории и системы управления. 2010. № 6. С. 122-137.
- [Клещев, 2011] Клещев А.С., Грибова В.В. Проблемы создания жизнеспособных интеллектуальных систем и методы их решения // International Journal "Information Technologies & Knowledge" Vol.5, Number 3, 2011. С. 250-258.
- [Клещев, 2012] Клещев А.С., Шалфеева Е.А. Системный анализ при автоматизации интеллектуальной профессиональной деятельности // Труды XIII Национальной конференции по искусственному интеллекту с международным участием «КИИ-2012», Белгород: Изд-во БГТУ, 2012. т.2. С.128–135.
- [Москаленко, 2006] Москаленко Ф.М. Алгоритм диагностики, основанный на реальной онтологии медицины, для многопроцессорной ЭВМ / Доклад Расо2006, 2006. Издатель: Институт проблем управления им. В.А. Трапезникова РАН.
- [Румянцева, 2007] Румянцева Е.Л., Слюсарь В.В. Информационные технологии. Общая характеристика систем поддержки принятия решений // М.: Форум, Инфра-М, 2007.
- [Шалфеева, 2013] Шалфеева Е.А. Семантические модели представления решателя Интеллектуальной системы // Материалы III Междунар. научн.-техн. конф. "Открытые семантические технологии проектирования интеллектуальных систем" (OSTIS-2013) / отв. ред.: В. В. Голенков. Минск: БГУИР, 2013. С. 257-264.

---

### Authors' Information

---



**Elena Shalfeeva** – senior researcher of Intelligent system laboratory, IACP, Vladivostok, Russia;  
e-mail: shalf@iacp.dvo.ru

Major Fields of Scientific Research: artificial Intelligence and decision making, program models and systems

---

## MODELING OF THE ORGANIZATIONAL AND EXPERT ACTIVITIES OF THE SCIENTISTS AT THE INSTITUTE OF MATHEMATICS AND INFORMATICS, BULGARIAN ACADEMY OF SCIENCES, IN 2005-2010

J. Tabov, S. Hristova

**Abstract.** In September 2010 attestation of the scientists at the Institute of Mathematics and Informatics of the Bulgarian Academy of Sciences (IMI-BAS) was held. The data in the attestation forms give us the opportunity of doing evaluations and analyses of different aspects of the scientific activities in IMI and on the basis of these to do forecasts and recommendations for the future development of the Institute. It was already made for the research activity – the analysis of the results and forecasts about it is the subject of the paper [Tabov, Sotirova and Hristova, 2012]. The subjects of the present paper are the organizational and expert activities in IMI. Our research rebuts a logical expectation – that the amount (more exactly, the average amount) of the expert activity of the scientists in IMI increases in the course of time and with the accumulation of experience; at the same time our research generally confirms an analogical expectation about the organizational activity.

**Keywords:** *research policy, science, attestation, BAN, IMI.*

---

### 1 Evaluation, attestation and administration of the scientific activity

---

If such attestation is held periodically, it could be used for researching of the alterations in the creative activity and other activities of the scientists during different periods. Similarly, their scientific-organizational, learning-educational and other activities could be monitored. The conclusions from these researches could be helpful and should be taken into account when someone makes administrative decisions.

The attestation in IMI in 2010 was part of a general to BAS procedure undertaken within the framework of the efforts to reform the Bulgarian Academy of Sciences and its units. The indicators for evaluation in the attestation forms and given points were and still are subjects of discussions and criticism, often reasonable. But, on one hand, the evaluation and attestation are significant for the administrative solutions and could serve as a reference point and self-control of the scientists, and this makes them necessary elements of the modern organization of the scientific units and institutions. On the other hand, development of appropriate criteria and scales for evaluation is not trivial and should be based on some experience; that is why the attestation in IMI in 2010 is the first step in the right direction.

Many colleagues think that not only the BAS, but the whole field of science and higher education in Bulgaria should be fully reformed. Independently of those if these reforms will be undertaken in each institution separately or they will be general, we believe that they should be preceded by careful evaluations of the scientific manifestation of the researchers and units. What is more, procedures for attestation and evaluation should be carried out periodically.

Holding of a large-scale attestation of the scientists at BAS is a positive thing. Unfortunately, the universities in Bulgaria, which for the present neglect the control of the quality of the education and science, did not follow the example of BAS.

That is why the scientific community should put a lot of effort for the gradual establishment of adequate evaluations of the professional abilities and manifestation of the scientists, which to stimulate their creative

attitude and, at the same time, to help the governing bodies of the scientific units and institutions with taking the right administrative decisions.

The analysis of the results of attestation, which we will show below, is also connected with an extremely important for IMI problem: "aging" of the scientists in the institute. The average age of the researchers is constantly growing and the young scientists are rare.

## 2 Presentation of the results of attestation

The staff with scientific functions filled in special attestation forms with a description of their activities in the period 2005-2010 and data about the results of them. The number of scientists filled in attestation forms is 133.

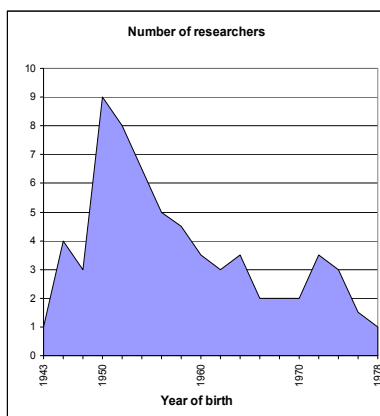
The indicators for evaluation were grouped into several sections:

- I. Research activity.
- II. Applied research activity.
- III. Learning-educational activity.
- IV. Scientific-organizational activity.
- V. Expert activity.

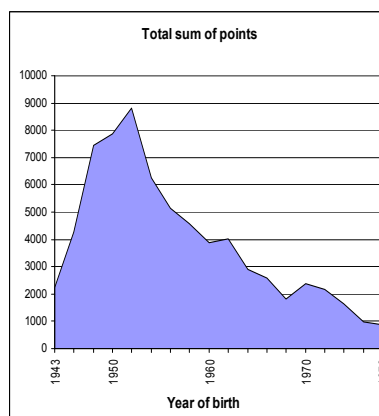
The analysis of the results from research activity, the construction of a special model and relevant forecasts and recommendations are objects of a specific paper [Tabov, Sotirova and Hristova, 2012]. Here we use the same methodology to analyze the results of scientific-organizational and expert activities in IMI for the mentioned period.

Fig.1. shows the distribution of researchers filled in the attestation forms, by age. Fig.2. shows the distribution of total sum of points given to the scientists, by age.

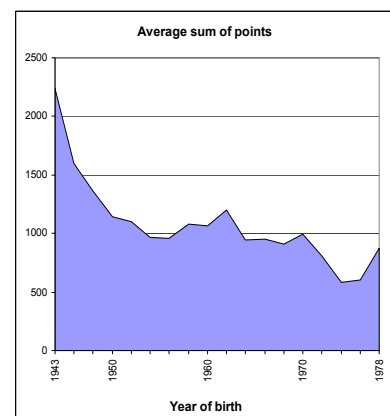
The next graph – in Fig.3 – shows the average sum of points got by a scientist born in the respective year.



**Figure 1.** The distribution of scientists by age. The abscissae are the years of birth, starting from 1943



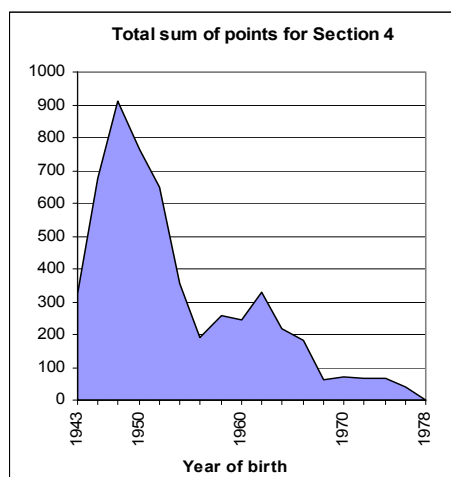
**Figure 2.** The distribution of points by age of the scientists



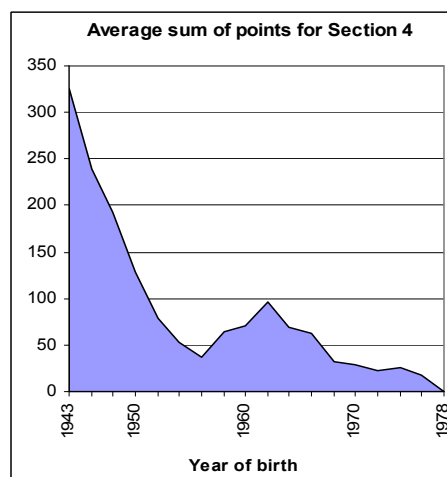
**Figure 3.** The average sum of points, by age of the scientists

The evaluations of the activities in IV and V sections of the attestation ("IV. Scientific-organizational activity" and "V. Expert activity") are of special interest to us. The next two graphs (in Fig.4 and Fig.5) are analogous respectively to the graphs in Fig.2 and Fig.3 but they are based on the evaluations for IV section only. Fig.6 and Fig.7 show the evaluations of the activities in V section.

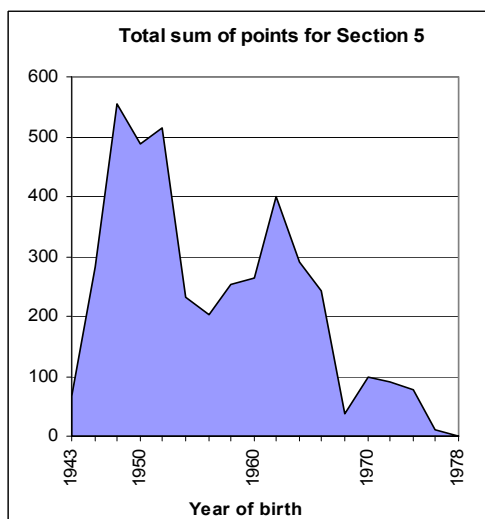




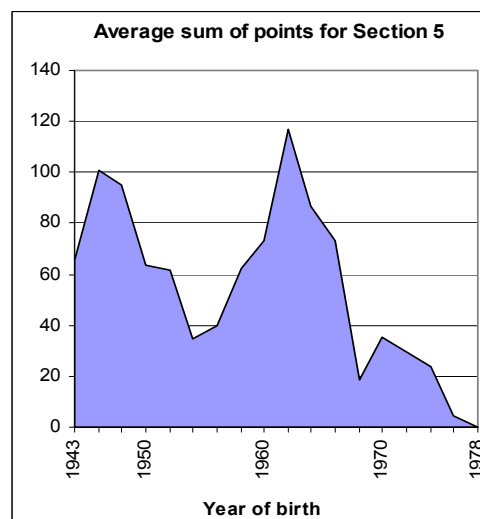
**Figure 4.** The distribution of points for Section IV by age of the scientists.



**Figure 5.** The average sum of points for Section IV, by age of the scientists.



**Figure 6.** The distribution of points for Section V by age of the scientists.



**Figure 7.** The average sum of points for Section V, by age of the scientists.

### 3 Analysis of the results

We can analyze the above shown results if we know the purpose and some details of attestation.

In fact, it gives a numerical evaluation of the activities of scientists, scientific bodies and organizations during the period 2005-2010. But subject of this paper are mainly some specific aspects and results from these activities.

To illustrate the process of attestation we will mention some of the rules for evaluation of the activities in IV and V sections of the attestation form – Scientific-organizational and scientific-administrative activities, and Expert activity. In each section are separated several subsections:

#### IV. Scientific-organizational and scientific-administrative activities

##### 4.1. Participation in administrative bodies of BAS

##### 4.2. Participation in administrative bodies of Specialized Scientific Units in BAS

##### 4.3. Participation in organizational and program committees of international and national scientific forums

4.4. Participation in scientific, expert councils, commissions, etc. in the field of science and higher education in Bulgaria and abroad

4.5. Participation in administrative bodies of scientific institutions, organizations and universities in Bulgaria and abroad

4.6. Participation in editorial boards and councils of national, foreign and international scientific journals

4.7. Prizes

V. Expert activities carried out as a scientist at the Bulgarian Academy of Sciences

5.1. Participation in state and governmental bodies

5.2. Participation in national and international councils, commissions and other public bodies and organizations

5.3. Expert, consulting and other activities to help institutions and administrative bodies

5.4. Public reviewer's activity

5.5. Anonymous reviewer's activity

In the subsections are counted the activities for which points are awarded. For example, we have the following activities in subsection 4.2.:

4.2.1. Director, deputy director, scientific secretary, chairman of the General Assembly, chairman of the Scientific Council

4.2.2. Deputy chair, secretary of the Scientific Council

4.2.3. Member of the Scientific Council

4.2.4. Member of a commission or other structure with expert functions in Specialized Scientific Units (certifying commission, etc.)

4.2.5. Head of Section (laboratory, sector)

4.2.6. Scientific Secretary of Section at the Institute of Mathematics and Informatics, BAS

The points, fixed for these activities, are:

4.2.1. 25 points per year

4.2.2. 10 points per year

4.2.3. 8 points per year

4.2.4. 3 points per year

4.2.5. 6 points per year

4.2.6. 5 points per year

The activities in the other subsections are evaluated in a similar way.

Let us look at the graphs shown in figures above and see what follows from them.

Fig.1 shows that the group of scientists which are 56-62 years old (born in 1948-1954) is numerous (51 people); one can say that it is the core of the Institute. Together with the elder colleagues (born before 1948 – 17 in number) these scientists (over 55 years old in 2010) are 67 in number, i.e. they represent more than half of the scientific potential of the Institute.

Or, in other words, at this rate of development of IMI, in 2014 the half of its scientists will be over 60 years old.

In our opinion, the average age of the scientists in a research institute as IMI should be around 48-50 years and the number of scientists that are 25-40 years old should be not less than that of scientists over 55 years old. From this perspective, the age distribution of the scientists in IMI is unfavourable and this fact marks a tendency to its deterioration in the future ([Tabov, Sotirova and Hristova, 2012]).

The negative predictions for the future are enhanced by Fig.2 and Fig.3. They show that the scientists over 55 years old give a large contribution (82526 points) to the total volume of scientific events in IMI (143655 points).

---

If we imaginatively move the left end of the graph in Fig.2 to the right till 1955 we will see that a large amount of scientific activities will be left out of IMI results ([Tabov, Sotirova and Hristova, 2012]).

Fig.4 and Fig.5 show that the group of scientists over 60 years old is the most engaged of all groups in scientific-organizational and administrative activities: their average sum of points is greater than 100. To a certain extent unexpected is the sudden decrease in the same activities of the scientists which are 52-58 years old (i.e. born between 1952 and 1958).

Fig.6 and Fig.7 show that the group of scientists between 44 and 52 years old and the group of scientists over 60 years old are the most engaged of all groups in expert activity (i.e. the average sum of points is bigger than 60). Here we see the same sudden decrease (as in Fig.4 and Fig.5) in the expert activity of the scientists between 52 and 58 years old (i.e. born between 1952 and 1958). This poses an interesting problem. Of course, we can suppose that this decrease is determined from the fact that the scientists direct their efforts to doctor degrees and dissertations but we need additional researches to understand the reasons for this.

---

## Bibliography

---

[Tabov, Sotirova and Hristova, 2012] Tabov, J., Sotirova, K. and Hristova, S. Researchers in IMI-BAS 2010 – Attestation Analysis – Results and Recommendations. BJSEP 2012, Number 2, Volume 6.

---

## Authors' Information

---



**Jordan Tabov** – professor; Institute of Mathematics and Informatics of the Bulgarian Academy of Sciences, block 8, Acad. G. Bonchev Street, 1113, Sofia, Bulgaria; e-mail: [tabov@math.bas.bg](mailto:tabov@math.bas.bg)

*Major Fields of Scientific Research: Applications of Mathematics and Informatics in the Humanities, Didactics of Mathematics and Informatics*



**Svilena Hristova** – research assistant; Institute of Mathematics and Informatics of the Bulgarian Academy of Sciences, block 8, Acad. G. Bonchev Street, 1113, Sofia, Bulgaria; e-mail: [svilenajh@abv.bg](mailto:svilenajh@abv.bg)

*Major Fields of Scientific Research: Applications of Mathematics and Informatics in the Humanities*

## METHODS AND MODELS FOR SELECTION OF RATIONAL SOLUTIONS IN DECISION-MAKING SYSTEMS

Oleksandr Kuzomin, Mikhail Shtukin

**Abstract:** *The problems of analysis and synthesis of the intelligent systems for decision-making under conditions of uncertainty and multi-objectiveness are considered. The models and methods for the multi-objective presentation of emergencies are developed taking into account the micro situation concepts and knowledge quanta under conditions of uncertainty.*

**Keywords:** *decision-making, extreme situations, solution models, micro situations, knowledge engineering.*

---

### Introduction

Analysis of development of natural catastrophic phenomena on the Earth shows that despite scientific and technological progress protection of people and techno sphere against naturally occurring dangers doesn't rise. Annually the number of victims of destructive natural phenomena in the world increases by 4,3% and preys – by 8,6%. Economic losses increase on the average about 6% per year. At present there is an understanding in the world that natural catastrophes represent the global problem being the source of the most profound humanitarian shocks and it remains one of the most important factors defining the stable development of economy. The chief causes for retaining and aggravating of natural dangers are as follows: the increase in anthropogenic action on the environment; irrational placement of the economy objects; settling of people in the zones of potential natural danger; insufficient efficiency and lack of development of the environment monitoring system: weakening of the state system of observation over the natural processes and phenomena; absence or a poor state of hydro technical, anti mudflow and other protective engineer constructions as well as protective afforestation; insufficient volumes and low rates of seismological construction, strengthening of buildings and structures in seismically dangerous areas; lack or insufficiency of cadasters of potentially dangerous areas (regularly flooded, particularly seismically dangerous, sill dangerous, avalanche dangerous, tsunami dangerous etc.).

An extremely complicated decision-making (DM) problem before and after extreme situations (ES) emergence has to be solved under uncertainty and multi-objectiveness conditions. Artificial intelligence (AI) methods are used in solution models and methods [1, 2, 3].

But the hope to receive efficient solutions requires additional investigations in many applications.

It is pertinent to note that the information collected from different sources and at different time is aging [4]. To increase the reliability the rate of inquiry into the date transducers and sources is associated with some delay. Moreover, change in the situation during the period of the ES prediction often represents a random process and collection of the precedent and ontological information is always associated with a time factor and variations in the information quality and quantity. This is, by no means, incomplete enumeration of the factors acting on the information aging [4].

The information aging, development of rational knowledge base for making decisions with minimal expenditures of time and material resources when preventing ES and eliminating their results should be taken into account

when developing the quantum approach to the knowledge engineering [3] the concept of algorithmic structures of  $k$  – knowledge [1,2] and concepts of micro situations  $\{\text{Sit}_i\}$ ,  $i = \overline{1, n}$  [3]. In this case using data and knowledge systems analysis one should estimate the micro situation taking into account the information aging, its reliability and value of knowledge quanta, revisions or additions of quantitative and qualitative data during estimation and overcoming of the ES results, which act on changing the ES development scenario, machine algebras needed for quantitative and qualitative estimation of the ES. The number and numerical sequence described by the knowledge quanta of the 0<sup>th</sup> level; the vector or function described by the quantum of the 1<sup>st</sup> level; the matrix or composition of functions described by the quantum of the 2<sup>nd</sup> level etc. [1] are needed to reveal micro situations [2, 3] according to the theorems of Godel and Gabor [5, 6]. In this case it is necessary to realize the principles of self-organization of the model, of the outer addition, Godel approach at self-organization of the models, to take into account the outer criteria of models selection, to carry out partitioning of the data table into parts, to use the hypothesis of selection and principle of conservation of freedom in choosing, to apply heuristic methods, to realize simultaneous simulation at different levels of generality of the language of the mathematical description of objects [6]. Optimal complexity of the ES prediction models is defined by the uncertainty and multi criteria conditions.

---

### Setting of the problem

---

**The problem** – it is necessary to develop models for prediction of natural emergency situations which will give the possibility to raise reliability and safety of the quantitative data of the situations analysis, knowledge base and provide efficiency of prevention and elimination of the natural emergency situations results.

**Solution of the problem.** The problem of control and prediction of the natural environment  $Sit$ . To solve the problem stated above it is offered to introduce the optimal procedures of prediction  $Pr_p$  and search for  $Pr_r$  solution.

The procedures  $Pr_p$  are carried out using the models –  $\{Mod\}$ , which represent the properties of the natural environment in the ES zone of control and give the possibility to predict the ES coming with a sufficient degree of precision as soon as possible. In so doing the  $Pr_p$  procedures should be realized as best as possible the requirements and characteristics of the statistical model  $\{Mod\}$  for defining the set of micro situations for precedent data

$$\{\text{Sit}_i\}, i = \overline{1, I}$$

Notation of the micro situations has the following form presented in fig.1

In this case it is possible to separate a set of the knowledge quanta  $\{kz_j\}$ ,  $j = \overline{1, J}$  of the best ontological solutions. Besides the geo information context  $G_n$ ,  $n = \overline{1, N}$ , the already obtained ontological experience for the knowledge quanta is singled out in the form of the best rational solutions from the set of resources  $\{Res\}$ , which can be or must be at the disposal of the decision-making person (DMP).

In the general form the ES class pattern for modeling of the object-oriented mode can be defined as follows:

Emergency situation (ES):=

Group microsituations |for several points of control

declaration-concepts of a group –microsituations|  
 declaration-concepts of a property-properties -notions  
 group-of microsituations

**group-of microsituations :=**

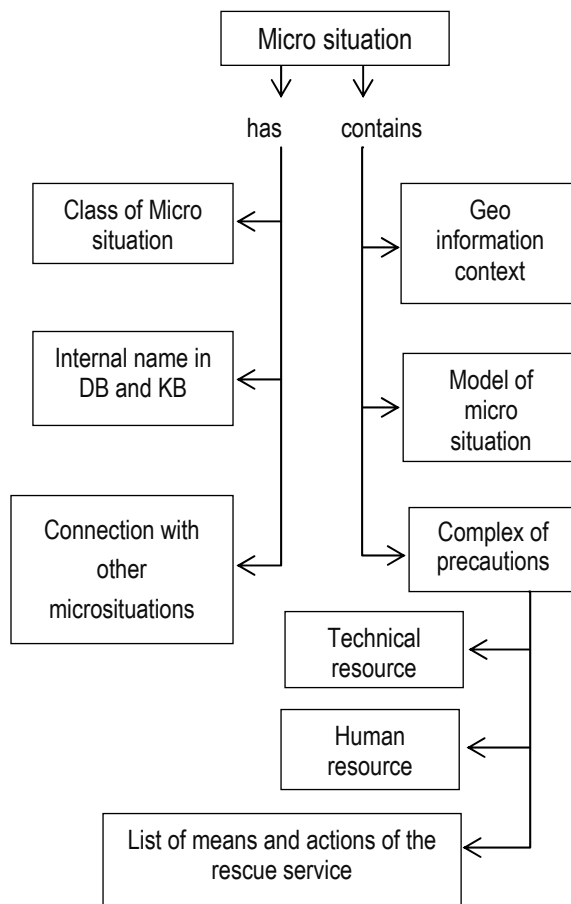
Precedent microsituations: ;

**microsituations :=**

microsituation 1;  
 microsituation 2  
 ...  
 microsituation N.

**microsituation i:=**

central-concept. |  
 central-concept, geo information context.  
 secondary-concept, making decisions.



**Figure 1.** Micro situation notation

Precedent knowledge  $kz_j$  quanta are defined for the corresponding geoinformation context in the ontological areas of ES. Procedures  $Pr_r$ , should ensure making of rational decisions on the basis of the systems analysis of

the precedent data and knowledge base with the best ontological decisions. The  $Pr_r$  procedures in the rational decision search methods use switching of the controlled and precedent microsituations models built on the data and knowledge bases.

As a whole the problem of simulation under uncertainty conditions can be specified formally by a set of the following form

$$SS = \langle \text{Mod}, R(\text{Mod}, F(\text{Mod}), F(SS)) \rangle,$$

where  $\text{Mod} = \{\text{Mod}_1, \dots, \text{Mod}_m\}$  – is a set of formal or logical-linguistic models based on the precedent data and knowledge of the situation which realize definite intelligent situation functions;

$R(\text{Mod})$  – is a set of rules for selection of the microsituations representing the data acting the most strongly on the origin of ES, on the needed model or a totality of models in the current situation, i.e. the rules for realization of the mapping  $R(\text{Mod}) : \text{Sit} \rightarrow \text{Mod}$ , where  $\text{Sit}$  – is a set of the possible situations (states) which can be either an open one, or  $\text{Sit}' \rightarrow M$ , where  $\text{Sit}'$  – is some set of the generalized microsituations (states), for example, normal (standard), anomalous or emergency, the model change takes place when falling within them;

$F(\text{Mod}) = \{F(\text{Mod}_1), \dots, F(\text{Mod}_m)\}$  – is a set of the modification rules of the models  $\text{Mod}_m$ ,  $m = 1, \dots, M$ .

Each rule  $F(\text{Mod}_m)$  realizes the mapping

$$F(\text{Mod}_m) : \text{Sit}'' \times \text{Mod}_m \rightarrow \text{Mod}' ,$$

where  $\text{Sit}'' \subseteq \text{Sit}$ ,  $\text{Mod}_m'$  – is some modification of the model  $\text{Mod}_m$ ;

$\text{Mod}_m$  – is the modification rule of the proper model system  $SS$  – of its base constructions  $\text{Mod}, R(\text{Mod}), F(\text{Mod})$  and, probably, of the rule  $F(SS)$  itself, i.e.  $F(SS)$  realizes a diversity of mappings (or a complex mapping)

$$F(SS) : \text{Sit}''' \times \text{Mod} \rightarrow \text{Mod}' ,$$

$$\text{Sit}''' \times R(\text{Mod}) \rightarrow R'(\text{Mod}) ,$$

$$\text{Sit}''' \times F(\text{Mod}) \rightarrow F'(\text{Mod}) ,$$

$$\text{Sit}''' \times F(SS) \rightarrow F'(SS) ,$$

where  $\text{Sit}''' \subseteq \text{Sit}''$ ,  $\text{Sit}''' \cap \text{Sit}' = \emptyset$ ,  $\text{Sit}''' \cap \text{Sit}'' = \emptyset$  i. e. the given type modification rules are used in the situations when the available sets of models, selection rules and modification rules are insufficient for searching for a decision (decisions) in the problematic situation at hand. In this case both the internal means for models and rules (hypotheses) generation, and the external meta knowledge representing the pragmatic aspect of the problematic situation can be used for modification of  $F(SS)$ .

The particularity of the ES appearance is associated with variations in the environment, the decision-making is accompanied by limitations or insufficient resources which are available to a DMP. The decision which has been made is generally called a rational one. Such a decision is made in the course of a very limited moment

$t_m \in [t_n, t_k]$ , where  $t_n$  – is the interval low boundary,  $t_k$  – is the interval upper boundary, but sometimes even this time is insufficient. Thus it is offered to use the data and knowledge precedents ontological experience in the given work to solve this complicated problem.

Ontological quantitative and qualitative precedents distinguished for successful and qualitative solution should be used in the constructive solution to prevention or eliminate the ES results, as it was mentioned above.

In this case it is necessary to strive for the optimal decision using the allowable resource set  $\{Res_c\}$ ,  $c = \overline{1, C}$

$$Res^0 = \arg \operatorname{extr}_{Res_c} Q(Res_c), c = \overline{1, C},$$

where the symbol  $\arg \operatorname{extr} Q(Res_c)$  makes it possible to define  $Res_c$  – quantitative and qualitative means of evacuation, announcement, building-assembly means for elimination of the ES results or its prevention, communication means etc.

So, to eliminate results or prevent the ES it is important that the decisions made by the DMP should be timely in the specified time interval and solve the problem in a sufficiently complete and complex form. That is why it is necessary to solve the multicriteria problem of optimization taking into account the local criteria presented by the criteria finite sequence

$$Q^\Sigma = \arg \operatorname{extr}_{Res_i \in \{Res\}} \langle Q(Res) \rangle, \forall i = \overline{1, K},$$

where  $K$  – is the number of local criteria;

$Q^\Sigma$  – is the local criteria finite sequence presentation, among them there can be: minimization of a risk, material resources for prevention or elimination of the ES results, minimization of time expenditures for searching rational solutions, time for announcement of population about the ES threat, maximization of protective constructions reliability etc.

In this case the valid decision of the set problem can be presented in the following form:

$$\{X_k, X_q, Mod, Sit, kz, Q\} \xrightarrow{\{Pr_p, Pr_r\}} Res^0.$$

Mod – is the model of the ENS presentation should reflect the connection between natural situations with the search of the managerial, organizational solutions and will require the necessary resources –  $Res$ , directed to prevention or elimination of the emergencies consequences which ensure the minimum risk for a human being vital activity  $\mathfrak{R}$ .

In other words, to solve the problem set above for the controlled situation of the environment in the specified controlled region it is necessary to develop the model of the ENS in the following form:

$$Mod_m = \operatorname{extr}_{r_i^m \in \mathfrak{R}} Mod \{X_k, X_q, \mathfrak{R}, Res\}.$$

Integration of the indicated elements in one model can exist if the following regularities are taken into account: A – associations;  $\Pi$  – sequences; K – classifications; KL – clusterizations; Pr – predictions.



Association occurs in the case if the current problematic natural situation  $Sit_t$ , where  $t$  – is current time and preceding or the precedent situation occurred earlier  $Sit_i$  were related. There is a communication  $A : Sit_t \rightarrow Sit_i$ .

For  $\Pi$  it is true  $\Pi : Sit_i^{t-1} \rightarrow Sit_i^t$  as there is a sequence of situations connected in time.

Using the classifications  $Sit_i = Kl\{(Sit_i \in Kl_{kr}) \vee (Sit_i \in Kl_{\neg kr})\}$ , where  $Kl_{kr}$  – is a class of emergencies, and  $Kl_{\neg kr}$  – is a class of not emergency situations, the features are revealed through the introduced microsituations [1, 2, 4], which characterize the group of situations, one or other situation belongs to this group [2, 3]. This is performed through the analysis of the already classified situations and formulation of some set of rules [3, 4].

Clusterization of differs from the classification in that the groups themselves are not set in advance  $Sit_i = Kl\{(Sit_i \in Kl_{kr}) \vee (Sit_i \in Kl_{\neg kr})\}$ . The Data Mining means separate independently different similar data groups using the clusterization.

Under conditions of uncertainty and multicriteriality the informative core of the offered knowledge-oriented microsituation approach to prediction of the ES and decision-making before and after the ES rise is solution of the following problems decision:

1. Formalization of different level  $k$  – knowledge, estimation of sets of microsituations for these levels, taking into account the information aging and self-organization of the ES mathematical models corresponding to the principles of the non-final decisions of D. Gabor, external addition of Godel and mass selection of A.G. Ivakhnenko.
2. Identification (recognition), proximity to the standard situation.
3. Prediction of ES.

The standard microsituation is defined on the basis of systems analysis of a set of microsituations of knowledge quanta of  $\alpha$  – ,  $\beta$  – ,  $\lambda$  – uncertainty for presentation of cluster of each level using models and methods for the models self-organization. The proximity of the standard microsituation makes it possible to group together the knowledge quanta for every level by the most influencing factors for the specified geoinformational context. The generalized standard microsituation is defined for all precedents and ontologies. The current controlled situation is compared to the generalized standard microsituation. The comparison results are based the on the self-organizing models of knowledge quanta of  $\alpha$  – ,  $\beta$  – ,  $\lambda$  – uncertainty.

1. Estimation of the standard microsituation of the set of microsituations of knowledge quanta of  $\alpha$  – ,  $\beta$  – ,  $\lambda$  – uncertainty for the geoinformation context situation being analyzed.
  - 1.1. Estimation of the knowledge quanta microsituations.

Statement 1. The principle of the non-ultimate solutions offered by D. Gabor, 1971: "...each uniserial procedure can be changed for a multiserial one on retention of the sufficient freedom in choosing several best solutions in every step of self-organization" is valid at systems analysis of the precedent data on ES.

Statement 2. Selection of solutions is impossible on the basis of the precedent data coincidence with current indices of the situation in the region being controlled and should correspond to the principle of the external addition which is associated with indispensable fulfillment of Godel theorem (Naguel, Newman, 1970): "...only the

external criteria based on a new information make it possible to synthesize the object true model hidden in the noised experimental data".

None axioms system can be logically closed: one can always find such a theorem for the proof of which the external addition will be required – extension of the initial axioms system.

### 1.2. Base of data and precedent knowledge accumulation

Statement 3. Data and knowledge are aging. Digital representation of the continuous signal (temperature, pressure, humidity of environment etc.) in the form of the readings totality hampers further reconstruction of the continuous form of the signal, moreover, it is necessary to take into account the information "aging" factor (Efimov A.N.) in the system of control with a closed circuit stipulated by its delay in the feedback circuit. The action of the information delay and the signal continuous form restoration method on the system error is growing proportional to the frequency of the signals being converted and processed. Moreover, the concepts, statements and knowledge also are "aging" and require a regular renewal.

### 2. Models self-organization.

Statement 4. Long-term weather forecast is possible on the basis of use of the heuristic principle of mass selection at the cost of detection of factors acting the most on ES emergence with the help of microsituations and introduction of the models quality criteria. Such a principle of models self-organization was proposed by A.G. Ivakhnenko (1982): "...the model of the optimal complexity results in the course of the most expedient gradual complication of the self-organizing model at reaching the minimum of the external criterion of its quality".

Each of the system models is oriented to processing of some type of uncertainty. In particular, models and methods, using the approximate set apparatus for operating with incomplete and contradictory information, are effective in the situations of maximal uncertainty when there is no additional information on the ES identification problem.

As an additional information becomes available, for example, in the form of probabilities for the rules, the integrated approach combining methods of the approximate sets and probability theory can be applied, this increases the degree of probability of recommendations produced by the system for a decision-making person (DMP). With the complete information about a problematic situation at hand the solutions can be reliable and received as a result of the multicriteria optimization problem solution.

It should be noted that ES characterize the data and knowledge of the environment state  $\{X_k, X_q\}$ ,  $k = \overline{1, K}$ ,  $q = \overline{1, Q}$ , and the decision making is associated with the resource parameters  $\{Res_f\}$ ,  $f = \overline{1, F}$ , DMP can use them or influence on them. As a whole the state in the ES region can be presented as  $\{Sit_n\}$ ,  $n = \overline{1, N}$  in simulation.

For ES simulation it is possible to use the method of Kolmogorov-Gabor, where the ideas of self-organization and mechanisms of the living nature crossing (hybridization) and selection (choosing) are applied. It is possible to describe the model  $F(x)$  by the results of environment observations for the zone of the ES control. In this case the structure of the  $F(x)$  model is unknown.

Traditionally there is a sampling of  $N$  observations in the database to estimate the precedent information about ES:

$$\begin{array}{ll} \{X(1) & Sit(1)\} \\ \{X(2) & Sit(2)\} \\ & \dots \\ \{X(N) & Sit(N)\} \end{array}$$

The most complete dependence between the inputs  $X(i)$  and outputs  $\{Sit_n\}$ ,  $n = \overline{1, N}$  can be presented with the generalized polynomial of Kolmogorov-Gabor.

For  $X = \{x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_N\}$  the polynomial has the form:

$$Sit = a_0 + \sum_{i=1}^N a_i x_i + \sum_{j=1}^N \sum_{i \leq j} a_{ij} x_i x_j + \sum_{i=1}^N \sum_{j \leq i} \sum_{k \leq j} a_{ijk} x_i x_j x_k + \dots$$

where all coefficients  $a$  are unknown.

When constructing the model (when defining the coefficients values) the criterion of regularity (precision) can be used as the criterion:

$$\overline{\varepsilon^2} = \frac{1}{N} \sum_{i=1}^N (Sit_i - f(x_i))^2.$$

It is necessary that

$$\overline{\varepsilon^2} \rightarrow \min.$$

To ensure the raise the zero error can be ensured in the given sampling if to rise the degree of the model polynomial. If as a result of this there are  $N$  nodes of interpolation then it is possible to built a whole family models, each of them will give a zero error when passing through the experimental points

$$\overline{\varepsilon^2} = 0.$$

Generally nonlinearity degree is taken no higher than  $n - 1$ , if  $n$  – is the number of the sampling points. Let us symbolize by  $Sl$  – the complexity of the model (it is defined by the number of Kolmogorov-Gabor polynomial members). The value of the error  $\overline{\varepsilon^2}$  depends on the model complexity.

At the first stage of simulation the complexity of the ES model will be raised at the expense of the generalized estimate of the microsituations  $\{Sit_i\}$ ,  $i = \overline{1, n}$  for the values  $X = \{x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_N\}$  at different time for the precedent values of such ES. For such an ES the standard value of the precedent microsituation is defined

$$\{Sit_h^{et}\}, \quad h = \overline{1, H}$$

where  $H$  – is the general number of ES which are used for accumulation of the precedent data from the ontological sources for ES close to the ES type being considered. An expert defines the proximity by the type.

The knowledge quantum is defined at the next simulation stage using Euclidian metric. Therewith as the complexity increases it will decline first and it will rise then. But we need to choose such an optimal complexity at which the error will be a minimal one. Furthermore, the following moments can be singled out if the errors action is taken into account:

With different level of errors the dependence  $\overline{\varepsilon^2}$  on the complexity  $Sl$  will be changed preserving the general direction (it should be kept in mind that initially it will show a decrease with the increase in complexity, but then it will grow).

With the increase in the errors level the value of  $\min_{Sl} \overline{\varepsilon^2}$  will grow.

With the increase in the level of errors  $Sl = \arg \min \overline{\varepsilon^2}$  will decrease. In this case  $\overline{\varepsilon^2}(Sl_0) > 0$ , if the noise level is not zero.

Godel incompleteness theorem:

Each formal logical system contains a number of statements and theorems which can be neither disprove, nor prove remaining in the frameworks of this system of axioms.

In this case this theorem means that the sampling is always incomplete.

One of the ways to overcome this incompleteness is the principle of the external addition. An additional sampling (control) is used as an external addition, the points of this sampling were not use when training the system (i.e. when searching for the evaluating values of Kolmogorov-Gabor polynomial coefficients).

The search for the best model is realized in the following manner:

The whole sampling falls into the training sampling and the control one:  $N_{vyb} = N_{obuch} + N_{pruf}$ .

The values  $\hat{a}_0, \hat{a}_i, \hat{a}_{ij}$  are defined from the training sampling  $N_{obuch}$ .

The best models are chosen from the control sampling  $N_{pruf}$ .

The input vector has the dimensionality  $N$  ( $X = \{x_1, x_2, \dots, x_N\}$ ).

**The principle of freedom in choosing (inconclusiveness of the intermediate solution):**

For each pair  $x_i$  and  $x_j$  the partial descriptions (of the whole  $C_N^2$ ) type are constructed:

$$\text{or } \hat{Sit}^{(Sl)} = \varphi(x_i, x_j) = a_0 + a_i x_i + a_j x_j, \quad Sl = 1 \dots C_N^2 \text{ (linear);}$$

or

$$\hat{Sit}^{(Sl)} = \varphi(x_i, x_j) = a_0 + x_i + a_j x_j + a_{ii} x_i^2 + a_{ij} x_i x_j + a_{jj} x_j^2, \quad Sl = 1 \dots C_N^2$$

(quadratic).

1. We define coefficients of these models using the method of least squares (MLS) making use of the

learning sampling. I.e. we find  $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_j, \dots, \hat{a}_N, \hat{a}_{11}, \dots, \hat{a}_{ij}, \dots, \hat{a}_{NN}$ .

2. Further on we find an estimate for every of these models using the control sampling

$$\overline{\varepsilon_{Sl}^2} = \frac{1}{N_{pruf}} \sum_{i=1}^{N_{pruf}} \left[ Sit(k) - \hat{Sit}_k^{(Sl)} \right]^2,$$

where  $Y(k)$  – is the real output value in the  $k$ -th control point of the control sampling;

$\hat{Y}_k^{(Sl)}$  – is the output value in the  $k$ -th point of the control sampling according to the  $S$ -th model) and then we define  $F$  of the best models.

$$z_I = \varphi^{(2)}(x_i, x_j) = a_0^{(2)} + a_1^{(2)}x_i + a_2^{(2)}x_j + a_3^{(2)}x_i^2 + a_4^{(2)}x_ix_j + a_5^{(2)}x_j^2.$$

The estimate here is the same as in the first series. Selection of the best ones is realized again in the same way, but  $F_2 < F_1$ . The process of the series construction is repeated so long as the error mean square will fall. When

the increase in the error  $\overline{\varepsilon^2}$  is obtained, then the process is stopped.

If the partial descriptions are quadratic ones and the number of the polynomial series is  $Sl$ , then we receive the polynomial degree  $k = 2^{Sl}$ .

As opposed to the conventional methods of statistical analysis it is possible to receive sufficiently complicated dependence with such approach even having such a short sampling.

There is a problem: some variables  $x_i$  and  $x_j$ , not acting on the output data, can be thrown away.

In this connection the following modification is offered: to feed  $y_i$  and  $x_j$  at the second layer, i.e.:

$$z_I = a_0^{(2)} + a_1^{(2)}y_i + a_2^{(2)}x_j + a_3^{(2)}y_i^2 + a_4^{(2)}y_ix_j + a_5^{(2)}x_j^2.$$

This is important in the presence of a high level of noise to ensure nonbiasedness.

Two methods emerge for choosing the best candidates of partial description fed at a definite layer.

1. Criterion of regularity (precision)  $\overline{\varepsilon_{pr}^2}$ .

$$\overline{\varepsilon^2} = \frac{1}{N_{pruf}} \sum_{i=1}^{N_{pruf}} (Sit_i - Sit_i^*)^2,$$

$$\overline{\Delta_{pruf}^2} = \frac{\sum_{i=1}^{N_{pruf}} (Sit_i - Sit_i^*(x))}{\sum_{i=1}^{N_{pruf}} (Sit_i - \overline{Sit})^2}.$$

Criterion of nonbiasedness. Let us take the whole sampling and divide it into two parts  $R = R_1 + R_2$ . The first experiment:  $R_1$  – is the learning sampling,  $R_2$  – is the control sampling; define the model outputs  $y_i^*, i = 1 \dots R$ . The second experiment:  $R_2$  – is the learning sampling,  $R_1$  – is the control sampling; define the model outputs  $y_i^{**}, i = 1 \dots R$ , and compare. Criterion of nonbiasedness:

$$n_{sm} = \frac{1}{N} \sum_{i=1}^R (Sit_i^* - Sit_i^{**})^2$$

The less is  $n_{sm}$ , the greater is nonbiased the model.

Such a criterion is defined for each partial description of the first level and then  $n_{sm}$  is found for the level as a whole

$$n_{sm} = \frac{1}{F} \sum_{i=1}^F n_{sm,i}^{(1)}$$

for  $F$  of the best models. In a number of versions  $F = 1$ . The same is observed at the second layer  $n_{sm}^{(2)}$ .

The selection process lasts till this criterion stops decreasing, i.e. till meeting the condition

$$n_{sm}^{(2)} \rightarrow \min .$$

---

## Conclusion

The intelligent model of the multi-objective presentation of an emergency taking into account separation of knowledge quanta and precedent micro situations based on data of successful or rational decisions has been developed. The model of the situation, which features a priori indefinability taking into account the information aging, has been developed; random phenomena are joined by the cause and effect relations in this model.

---

## Acknowledgements

*The paper is published with financial support by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA ([www.ithea.org](http://www.ithea.org)) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine ([www.aduis.com.ua](http://www.aduis.com.ua)).*

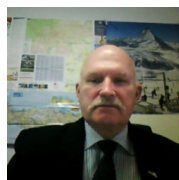
---

## Bibliography

- Efimov A.N. Information: value, aging, distance M., "Znanie", 1978. P.- 64. (Novelty in life, science and engineering, Series "Mathematics for Cybernetic Engineers" published monthly since 1967) (Rus)
- Kuzemin A, Fastova D, Yanchevsky I. Methods of adaptive extraction and analysis of knowledge for knowledge-base construction and fast decision making // International Journal on Information Theories&Applicatio.– Bulgaria. – 2005. – Vol. 12, №1.– P. 93 – 99
- Kuzemin A., Sorochan M., Y.anchevskiy I., Torojev A. The use of situation representation when searching for solutions in computer aided design systems International Journal // Information Theories & Applications. – 2005. – Vol.11, №1 .– P. 101 – 107.
- Martin Davis, "The Incompleteness Theorem", in Notices of the AMS vol. 53 no. 4 (April 2006), p. 414.
- Olivier Bousquet, Stéphane Boucheron, Gábor Lugosi «Introduction to Statistical Learning Theory» Advanced Lectures on Machine Learning Lecture Notes in Computer Science Volume 3176, 2004, pp 169-207.
- Sirodzha IB Quantum models and artificial intelligence for decision-making and management. - Kiev, Naukova Dumka, 2002, P.- 490.(Rus)

---

## Authors' Information



**Prof. Dr. Oleksandr Kuzomin** – Informatics chair Innovation-marketing department of Kharkiv National University of Radioelectronics; Kharkiv, Ukraine;

e-mail: [kuzy@daad-alumni.de](mailto:kuzy@daad-alumni.de); tel.: +38(057)7021515

Major Fields of Scientific Research: General theoretical information research, Decision Making, Emergency Prevention, Data Mining, Business Informatics.



**Mikhail Shtukin** – Aspirant of Kharkiv National University of Radioelectronics; Kharkiv, Ukraine;

e-mail: [shtumi@tut.by](mailto:shtumi@tut.by); tel.: +38(057)7021515

Major Fields of Scientific Research: General theoretical information research, Knowledge Discovery and Engineering, Business Informatics.

---

## STORING DICTIONARIES AND THESAURUSES USING NL-ADDRESSING

Krassimira B. Ivanova, Koen Vanhoof, Krassimir Markov, Vitalii Velychko

**Abstract:** *Our main goal in this paper is to propose a new approach for storing dictionaries and thesauruses using only the names but not pointers and this way to simplify and to speed up the corresponded computer programs. It is called "Natural Language Addressing" (NLA). This approach is a possibility to access information using natural language words or phrases as direct addresses of the information in the computer memory. For this purpose the internal encoding of the letters is used to generate corresponded address co-ordinates. This paper outlines the main idea of NLA approach. An extended example based on the WordNet thesaurus is given.*

**Keywords:** *addressing; natural language addressing.*

**ACM Classification Keywords:** *D.4.3 File Systems Management, Access methods.*

---

### Introduction

An Introduction to Natural Language Addressing (NLA) was given in [Ivanova et al, 2012; Ivanova et al, 2013]. Below we will remember the main idea of NLA.

We will use concept "address" in sense accepted in the Computer Science: *the code that identifies where a piece of information is stored* [WordNet, 2012]; a name or number used in information storage or retrieval that is assigned to a specific memory location; the *memory location* identified by this name or number; a name or a sequence of characters that designates an e-mail account or a specific site on the Internet or other network [AHD, 2009].

We usually make difference between human aspect of the concept "address" and its computer "understanding". The NLA approach is based on using human representation of the address (by natural language words) as computer memory address.

#### ➤ Computer indexes

The textual information of dictionaries and thesauruses may be stored in (internal or external) computer memory. Locating concepts and connected to them definitions may be done by:

- Direct scanning the files;
- Indexing and based on it search of the pointer to address of text element.

Scanning files is convenient only for small volumes of concepts. Some rationalization is possible using some algorithms like binary search.

*Indexing* is creating tables (indexes) that point to the location of folders, files and records. Depending on the purpose, indexing identifies location of resources based on file names, key data fields in a database record, and text within a file [PC mag, 2013].

The main idea of indexing is to facilitate the search by search in (multi-level) index and after that to ensure the direct access to address given by pointer. In other words, the goal of data indexing is to ease the search of (and

access to) data at all times. This is done by creating a data structure called index and providing faster access to the data. Accessing data is determined by the physical storage device being used. Indexing could potentially provide large increases in performance for large-scale analysis of unstructured data. In addition, the implementation of the chosen index must be suitable in terms of index construction time and storage utilization [Faye et al, 2012].

Indexing needs resources: memory for storing additional information and time for processing, which may be quite a long, especially for updating of the indexes when new elements are added or some old ones are removed.

### ➤ Names and addresses

Basic element of an index is couple: (name, address). In different sources the "name" is called "key", "concept", etc. The address usually is given by any "number", "pointer", etc.

There are two interpretations of the couple:

1) Address is a connection of the concept with its definition, i.e. practically we have triple:

(concept, address, definition).

2) Concept is a name of the address and may be used for user friendly style of programming and the third part of the triple (definition) may be variable.

NLA approach is based on using the computer encoding of name (concept) letters as address of connected to it information. This way no indexes are needed and high speed direct access to text elements is available. It is similar to natural order addressing in a dictionary book where no explicit index is used but the concept by itself locates definition. For instance, let have the next concept and corresponded definition:

*"London: The capital city of England and the United Kingdom, and the largest city, urban zone and metropolitan area in the United Kingdom, and the European Union by most measures."*

In computer memory, for example, the definition may be stored at address "FF084920". The index couple is:

("London", "FF084920"),

i.e. at memory address "FF084920" the main text, *"The capital ... measures."* will be stored.

To read/write the main text, firstly we need to find name "London" in the index and after that to access memory address "FF084920" to read/write the definition.

If we assume that name "London" in the computer memory is encoded by six numbers (letter codes), for instance by using ASCII encoding system London is encoded as (76, 111, 110, 100, 111, 110), than we may use these codes as direct address to memory, i.e.

("London", "76, 111, 110, 100, 111, 110")

One may remark that above we have written two times the same name and this is truth. Because of this we may omit this couple and index, and read/write directly to the address (76, 111, 110, 100, 111, 110).

For human this address will be shown as "London", but for the computer it will be "76, 111, 110, 100, 111, 110".

From other point of view, the array (76, 111, 110, 100, 111, 110) may be assumed as co-ordinates of point in multidimensional (in this case – six dimensional) information space and the definition can be stored in this point.

The natural language does not contain words only of six letters long. The length of the words is variable and in addition there exists names as phrases. This means that we really have multidimensional address space defined by set of all natural words and phrases.

What we need is a program function which converts such multidimensional addresses in concrete linear machine locations (on the hard disk, for example) and corresponded access method.



---

### ➤ Storing Dictionaries and Thesauruses

**Dictionary** is a reference resource, in printed or electronic form, which consists of an alphabetical list of words with their meanings and parts of speech, and often a guide to accepted pronunciation and syllabification, irregular inflections of words, derived words of different parts of speech, and etymologies [Collins, 2003].

Dictionary may be a book, optical disc, mobile device, or online lexical resource containing a selection of the words of a language, giving information about their meanings, pronunciations, etymologies, inflected forms, derived forms, etc., expressed in either the same or another language; lexicon; glossary. Print dictionaries of various sizes, ranging from small pocket dictionaries to multivolume books, usually sort entries alphabetically, as typical CD or DVD dictionary applications do, allowing one to browse through the terms in sequence. All electronic dictionaries, whether online or installed on a device, can provide immediate, direct access to a search term, its meanings, and ancillary information [Dict, 2013].

**Thesaurus** is a book or catalog of words with the same or nearly the same meanings, or synonyms, and their opposites, or antonyms. An online thesaurus provides immediate electronic access to lists of alternate terms for the queried word, covering its various shades of meaning [Collins, 2003; AHD, 2009; Dict, 2013]

What is important is that the main dictionary construction is the couple (concept, definition). In the electronic variant it becomes triple (concept, address, definition).

The computer storing model for dictionaries is simple because the one-one correspondence "word-definition". For computer dictionary it is enough to have a sorted list of words and a tool for binary search and pointers to the corresponded definitions.

The computer storing models for thesauruses are more complicated due to existing more than one corresponded definitions for a given word (synonyms and etc.). Because of this, below we will outline and analyze one such model – the storing model of WordNet thesaurus [WordNet, 2012].

---

### WordNet thesaurus

---

WordNet® is a large lexical database of English. It was created and is being maintained at the Cognitive Science Laboratory of Princeton University under the direction of psychology professor George A. Miller. Development began in 1985. WordNet 3.0 database contains 155 287 words organized in 117659 synsets for a total of 206941 word-sense pairs; in compressed form it is about 12 megabytes in size [WordNet, 2012].

WordNet superficially resembles a thesaurus, in that it groups words together based on their meanings. However, there are some important distinctions.

- First, WordNet interlinks not just word forms — strings of letters — but specific senses of words. As a result, words that are found in close proximity to one another in the network are semantically disambiguated.
- Second, WordNet labels the semantic relations among words, whereas the groupings of words in a thesaurus do not follow any explicit pattern other than meaning similarity.

In developing WordNet lexical database, it has been convenient to divide the work into two interdependent tasks which bear a vague similarity to the traditional tasks of writing and printing a dictionary [Fellbaum, 1998]:

- One task was to write source files that contain the basic lexical data — the contents of those files are lexical substance of WordNet.
- The second task was to create a set of computer programs that would accept source files and do all the work leading ultimately to the generation of a display for the user.

The WordNet system falls naturally into four parts:

- The WordNet lexicographers' source files;
- The software to convert these files into the WordNet lexical database;
- The WordNet lexical database;
- The suite of software tools used to access the database.

WordNet's source files are written by lexicographers. They are the product of a detailed relational analysis of lexical semantics: a variety of lexical and semantic relations are used to represent the organization of lexical knowledge.

WordNet organizes nouns, verbs, adjectives and adverbs into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. Further they are arranged into a set of lexicographers' source files by syntactic category and other organizational criteria. Adverbs are maintained in one file, while nouns and verbs are grouped according to semantic fields. Adjectives are divided between two files: one for descriptive adjectives and one for relational adjectives.

The "*Grinder*" utility compiles the lexicographers' files. It verifies the syntax of the files, resolves the relational pointers, then generates the WordNet database that is used with the retrieval software and other research tools. To build a complete WordNet database, all of the lexicographers' files must be processed at the same time.

The Grinder is a multi-pass compiler that is coded in C++. The first pass uses a parser to verify that the syntax of the input files conforms to the specification of the input grammar and lexical items, and builds an internal representation of the parsed synsets.

In its second pass, the Grinder resolves all of the semantic and lexical pointers. To do this, the pointers that were specified in each synset are examined in turn, and the target of each pointer (either a synset or a word form in a synset) is found. The source pointer is then resolved by adding an entry to the internal data structure which notes the "location" of the target. In the case of reflexive pointers, the target pointer's synset is then searched for a corresponding reflexive pointer:

- If a reflexive pointer is found, it is replaced by the pointer to the original source.
- If a reflexive pointer is not found, the Grinder automatically creates one with all the pertinent information.

A subsequent pass through the list of word forms assigns a polysemy index value, or sense count, to each word form found in the on-line dictionary. There is a separate sense count for each syntactic category that the word form is found in.

The Grinder's final pass generates the WordNet database.

The resulting network of meaningfully related words and concepts can be navigated with a specialized browser [Fellbaum et al, 1998; Miller, 1995].

---

### **WordNet database**

---

WordNet database is divided into five lexical and grammatical classes: noun, verb, adjective, adverb and functional words. The parts of speech are organized in a hierarchy of synsets (nodes).

The internal representation of the lexicographic data is a network of interrelated linked lists. A *hash table* of word forms is created as the lexicographers' files are parsed. Lower-case strings are used as keys; the original orthographic word form, if not in lower-case, is retained as part of the data structure for inclusion in the database files.

As the parser processes an input file, it calls functions which create data structures for the word forms, pointers, and verb frames in a synset. Once an entire synset had been parsed, a data structure is created for it which includes pointers to the various structures representing the word forms, pointers, and verb frames.

All of the synsets from the input files are maintained as a single linked list. The Grinder's different passes access the structures either through the *linked list* of synsets or the *hash table* of word forms. A list of synsets that specify each word form is maintained for the purposes of resolving pointers and generating the database's index files.

For each syntactic category, two files represent the WordNet database — *index.pos* and *data.pos*, where "pos" is either "noun", "verb", "adj" or "adv" (the actual file names may be different on platforms other than Sun-4). The database is in an ASCII format that is human- and machine-readable, and is easily accessible to those who wish to use it with their own applications. Each index file is an alphabetized list of all of the word forms in WordNet for the corresponding syntactic category. Each data file contains all of the lexicographic data gathered from the lexicographers' files for the corresponding syntactic category, with *relational pointers resolved to addresses in data files*.

The index and data files are interrelated. Part of each entry in an index file is a list of one or more byte offsets, each indicating the starting address of a synset in a data file. The first step to the retrieval of synsets or other information is typically a search for a word form in one or more index files to obtain all data file addresses of the synsets containing the word form. Each address is the byte offset (in the data file corresponding to the syntactic category of the index file) at which the synset's information begins [Fellbaum, 1998].

The main relation among words in WordNet is synonymy, as between the words "shut" and "close" or "car" and "automobile". Synonyms (words that denote the same concept and are interchangeable in many contexts) are grouped into unordered sets (synsets). Each of WordNet's 117 000 synsets is linked to other synsets by means of a small number of "conceptual relations". Additionally, a synset contains a brief definition ("gloss") and, in most cases, one or more short sentences illustrating the use of the synset members. Word forms with several distinct meanings are represented in as many distinct synsets. Thus, each form-meaning pair in WordNet is unique [WordNet, 2012].

---

### WordNet storing model

---

Consider a representation of a synset of the word "accession" in the WordNet lexical database:

**00047131 04 n 02 accession 0 addition 0 001 @ 09536731 n 0000 |**

**something added to what you have already;**

**"the librarian shelved the new accessions";**

**"he was a new addition to the staff"**

The number 00047131 is a unique identifier of the synset of the noun {accession, addition}. The part of record between the symbols "@" and "|" indicates that this synset is subordinated to the synset with ID 09536731 which correspond to meaning "acquisition". The last part of the record (after the symbol "|") is interpretation of synset and some examples of using the words included in the synset.

From a software standpoint, this record requires a number of additional indexes for service the access, which of course needs additional resources.

As an example, consider the information about the word "accession". As an answer to the request for "accession", the WordNet system returns the following information (Figure 1):

The noun accession has 6 senses (no senses from tagged texts)

1. **{13251723}** <noun.process> accession#1 -- (a process of increasing by addition (as to a collection or group); "the art collection grew through accession")
2. **{13170404}** <noun.possession> accession1#2 -- ((civil law) the right to all of that which your property produces whether by growth or improvement)
3. **{13082910}** <noun.possession> accession#3, addition#4 -- (something added to what you already have; "the librarian shelved the new accessions"; "he was a new addition to the staff")
4. **{07078650}** <noun.communication> accession2#4, assenting#1 -- (agreeing with or consenting to (often unwillingly); "accession to such demands would set a dangerous precedent"; "assenting to the Congressional determination")
5. **{05115154}** <noun.attribute> entree#2, access#1, accession#5, admittance#1 -- (the right to enter)
6. **{00232781}** <noun.act> accession3#6, rise to power#1 -- (the act of attaining or gaining access to a new office or right or position (especially the throne); "Elizabeth's accession in 1558")

The verb accession has 1 sense (no senses from tagged texts)

1. **{00989696}** <verb.communication> accession#1 -- (make a record of additions to a collection, such as a library)

**Figure 1.** Answer by WordNet system to a query for the word "accession"

WordNet stores information about words in four main data files (for nouns, verbs, adjectives and adverbs). The data structure is the same in each of these files – one or more synsets are stored for every word and the access is performed by the address of the first bytes of the synsets, which is apparently given by an eight digit number beginning namely in this byte (Figure 2 and Figure 3). This value is the unique identifier of the synset. The synset data elements are separated by spaces. We should note that links to other synsets are given again by the absolute addresses.

```
13251723 22 n 01 accession 0 001 @ 13323403 n 0000 | a process of increasing by addition (as to a collection
or group); "the art collection grew through accession"
13170404 21 n 01 accession 1 002 @ 13070995 n 0000 ;c 08338303 n 0000 | (civil law) the right to all of that
which your property produces whether by growth or improvement
13082910 21 n 02 accession 0 addition 0 001 @ 13082742 n 0000 | something added to what you already have;
"the librarian shelved the new accessions"; "he was a new addition to the staff"
07078650 10 n 02 accession 2 assenting 0 002 @ 07076600 n 0000 + 00795631 v 0102 | agreeing with or
consenting to (often unwillingly); "accession to such demands would set a dangerous precedent";
"assenting to the Congressional determination"
05115154 07 n 04 entree 0 access 0 accession 0 admittance 0 003 @ 05113619 n 0000 + 02426186 v 0401 ~
05119817 n 0000 | the right to enter
00232781 04 n 02 accession 3 rise_to_power 0 003 @ 00060914 n 0000 + 01989112 v 0101 + 02358456 v
0101 | the act of attaining or gaining access to a new office or right or position (especially the throne);
"Elizabeth's accession in 1558"
```

**Figure 2.** Synsets of the word "accession" in WordNet data file for nouns

```
00989696 32 v 01 accession 0 002 @ 00990286 v 0000;c 00897092 n 0000 01 + 08 00 | make a record of
additions to a collection, such as a library
```

Figure 3. Synsets of the word "accession" in WordNet data file for verbs

What is important for us now, is the algorithm of reaching the synsets.

There are four index files of WordNet (for nouns, verbs, adjectives and adverbs). They are sorted in alphabetical order of words and for each word a special record is stored at separated line. Its structure is simple: at the first place the word is given and, after some coded information, absolute addresses of corresponded synsets are given (Figure 4 and Figure 5).

```
accession n 6 4 @ ~ + ; 6 0 13251723 13170404 13082910 07078650 05115154 00232781
```

Figure 4. Record for the word "accession" in the index of nouns

```
accession v 1 2 @ ; 1 0 00989696
```

Figure 5. Record for the word "accession" in the index of verbs

To reach all synsets of a word, firstly a binary search is made in all index files, the corresponded absolute addresses are collected and then system reads synsets directly from data files.

Algorithmic complexity in this case is  $O(\log(n_n)+\log(n_v)+\log(n_a)+\log(n_r))$ , where  $n_n$ ,  $n_v$ ,  $n_a$  и  $n_r$  are the quantities of nouns, verbs, adjectives and adverbs, respectively.

There is a second way to reach synsets. It is served by so called "sense index". This index is also sorted, but for every word there exist as much records as number of synsets exists for given word in all data files. For example, the word accession has seven records: six for its meanings as a noun and one for its meaning as a verb. Each record contains only one absolute address of a synset (Figure 6).

```
accession%1:04:03:: 00232781 6 0
accession%1:07:00:: 05115154 5 0
accession%1:10:02:: 07078650 4 0
accession%1:21:00:: 13082910 3 0
accession%1:21:01:: 13170404 2 0
accession%1:22:00:: 13251723 1 0
accession%2:32:00:: 00989696 1 0
```

Figure 6. Records for the word "accession" in the sense index

In this case, to reach all synsets of a word, firstly a binary search is made in the sense index and the corresponded absolute addresses are collected from all records for the word. Then, the system reads the synsets directly from the data files.

Algorithmic complexity in this case is greater than  $O(\log(n))$ , where  $n = n_n + n_v + n_a + n_r$ , i.e. total number of words in the database (nouns + verbs + adjectives + adverbs), because the words may be repeated many times, and further work is needed to retrieve all occurrences of the word.

---

## Disadvantages of WordNet storing model

---

WordNet storing model permits quick response of the system during its everyday using. The (binary) search in four types sorted index files and one general sense index, using corresponded hash tables, allows high speed of search and, based on it, extracting needed information via direct access based on absolute addresses in data files.

Many disadvantages of the WordNet organization are discussed in [Poprat et al, 2008]. An important shortcoming of the WordNet database's structure is that although all files are in ASCII, and are therefore editable, and in theory extensible, in practice *this is almost impossible*. One of Grinder's primary functions is the calculation of addresses for synsets in data files. Editing any of the database files would (most likely) create incorrect byte offsets, and would thus derail many searching strategies. At present time, building a WordNet database requires using of the Grinder and processing of all lexicographers' source files at the same time [Fellbaum, 1998]. In addition, Grinder utility is available only for Unix/Linux operating systems.

In the main, the WordNet database organization has following important disadvantages:

1. Absolute addressing is convenient for computer processing, but it is difficult to be used by customer;
2. Manual creating of numerical addresses is impossible, and their use can be done only by corresponded program;
3. End user has access only to the static ("compiled") version of the database, which couldn't be extended and further developed;
4. Building the WordNet database requires the use of WordNet program "Grinder" and processing of all lexicographers' source files at the same time;
5. Using the current format is not only cumbersome and error-prone, but also limits what can be expressed in a WordNet resource.

*Our goal is to show that WordNet lexical database may be realized without using pointers and this way to avoid pointed above limitations and recompilation of the database after every update.*

---

## NL-addressing

---

Let see a small example shown on Figure 7 - two variants of the synset of the word "accession" from different compilations of data file for nouns:

(a) Version of WordNet from August, 2012,

(b) An older version of WordNet from 2011 year, published in [Palagin et al, 2011].

**13082910** 21 n 02 **accession** 0 addition 0 001 @ **13082742** n 0000 | something added to what you already have;  
"the librarian shelved the new accessions"; "he was a new addition to the staff"

a) Version from the 2012 year

**00047131** 04 n 02 **accession** 0 addition 0 001 @ **09536731** n 0000 | something added to what you have already;  
"the librarian shelved the new accessions"; "he was a new addition to the staff"

b) Version from the 2011 year

**Figure 7.** Synset of the word "accession" from the data files for nouns

The difference between absolute addresses is visible.

Due to absolute addresses that are used as pointers, any change which cause alteration of the number of bytes in any data file makes it unusable and it must be recompiled as well as the corresponded index files.

Our main goal in this work is to propose a NLA as a new approach for storing WordNet database, using only the names but not pointers and this way to simplify and to speed up the corresponded computer programs.

Analyzing the database structure of synset we may see an important peculiarity.

In the computer memory, all characters are represented by numerical codes (one, two or four bytes depending on the encoding system such as ASCII or UNICODE). This way, in WordNet storing model, the unique digital code of word is combined with another unique numeric code of its place in the file (absolute address). This duplicating of the codes could be avoided if NLA model of information storing is used.

Let remember, ASCII internal representation of word "accession" is the following unique sequence of numbers: (97, 99, 99, 101, 115, 115, 105, 111, 110). It can be used as a spatial address in nine-dimensional space. At this address we can record the synset and access it again via this address.

In this case, for customer, the address will be presented by word "accession" and for computer by vector (97, 99, 99, 101, 115, 115, 105, 111, 110).

If we apply this opportunity to WordNet data, we will obtain a result that is understandable for users and, at the same time, fully recognizable for the computer programs.

As example two variants of the synset of word "accession" are shown in Figure 8: (a) WordNet version and (b) NLA-version.

An important feature of NLA-version is that it can be updated dynamically without recompilation of the database.

**13082910** 21 n 02 **accession** 0 addition 0 001 @ **13082742** n 0000 | something added to what you already have; "the librarian shelved the new accessions"; "he was a new addition to the staff"

a) WordNet version

**accession** 21 n 02 ; 0 addition 0 001 @ **acquisition** n 0000 | something added to what you already have; "the librarian shelved the new accessions"; "he was a new addition to the staff"

b) NLA-version

**Figure 8.** WordNet and NLA-versions of the synset of the word "accession"

## Experiments

### ➤ Experiments for storing dictionaries

Our first experiment was to realize a small multi-language dictionary based on NL-addressing. For this purpose, we have taken data from the popular in Bulgaria "SA Dictionary" [Angelov, 2012]. SA Dictionary is a computer dictionary, which translates words from Bulgarian language to English and vice versa.

For experiments we take a list of 23412 words in English and Bulgarian with their definitions in Bulgarian, stored in a sequential file with size of 2410 KB.

The experiments were provided at PC SONY Vaio, with Intel® Core™2 Duo CPU T9550 @ 2.66GHz 2.67GHZ, RAM 4.00 GB, 64-bit operating system Windows 7 Ultimate SP1.

The experimental program "WordArM" used for the experiments is specially designed for storing dictionaries and thesauruses based on NL-addressing.

For storing dictionaries we use simple model: definitions are stored at the NL-addresses formed from words corresponded to them.

The speed for storing, accessing, and size of the work and permanent files are given in Table 1.

**Table 1.** Experimental data for NL-storing of a dictionary

operation	number of instances	total time in milliseconds	average time for one instance	work file	permanent file
NL-writing	23 412	22 105	0.94 ms	80 898 KB	2 939 KB
NL-reading	23 412	20 826	0.89 ms		

The analysis of the results in Table 1 shows that the NL-addressing in this realization permits access practically equal for writing and reading for all data. The speed is more than a thousand instances per second.

Reading is possible immediately after writing and no search indexes are created.

The work memory taken during the work was 80898 KB. To analyze the system performance, work memory was chosen to be in a file but not in the main memory. After finishing the work, occupied permanent memory for compressed archive is 2939 KB. This means that the NL-indexing takes 2103 KB additional compressed memory (the sequential file with initial data is 2410 KB and compressed by WinZip it is 836 KB).

In further realizations of WordArM, the work memory may be realized as a part of main memory of computer as:

- Dynamically allocated memory;
- File mapped in memory.

In this case, the speed of storing and accessing will be accelerated and used hard disk space will be reduced.

### ➤ Experiments for storing thesauruses

Storing thesauruses is similar to storing dictionaries. The main difference is that the computer storing models for thesauruses are more complicated due to existing more than one corresponded definitions for a given word (synonyms) as well as other relations between words (hyperonymy, hyponymy, meronymy, etc.) which need to be encoded in proper way and the database has to support such relations. We had no goal to realize our own thesaurus. Because of this for experiments we have chosen the WordNet thesaurus.

Using NL-addressing, the WordNet lexical database may be realized without using pointers and this way to avoid the limitations and recompilation of the database after every update.

The main source information of WordNet is published as lexicographer files.

The total number of instances (file records) is 117 871 from which 206 instances contain service information but not concepts' definitions, so we have 117 665 instances for experiments, distributed in 45 thematically organized lexicographer files.

It is important to remark that there are equal synsets in several lexicographer files. This has matter when we integrate the 45 files in one source file for representing a thesaurus.



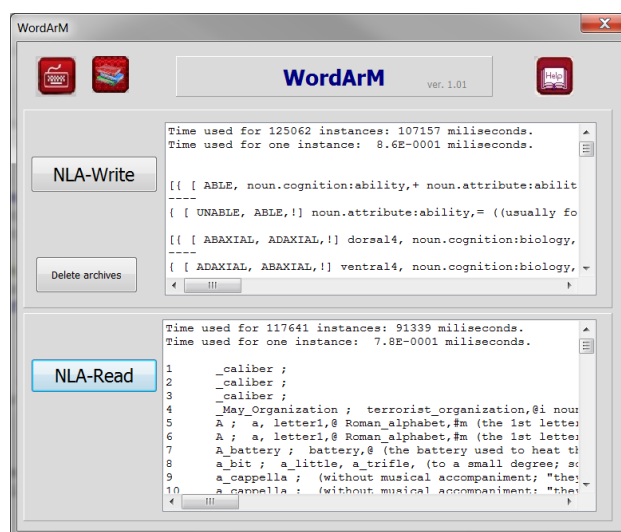
The program used for the experiments is the experimental program "WordArM". The results are given in Table 2. A screenshot from the WordArM for the case of WordNet as thesaurus is shown at Figure 9.

**Table 2.** Experimental data for storing WordNet as thesaurus

operation	number of instances	total time in milliseconds	average time for one instance
writing	125 062	107 157	0.86 ms
reading	117 641	91 339	0.78 ms
work memory: 385 538 KB; permanent memory: 15 603 KB; source text: 1 333 KB			

We receive practically the same results as for storing dictionaries. The analysis of results in Table 2 shows that the NL-addressing permits access practically equal for writing and reading for all data. The speed is more than a thousand instances per second. Reading is possible immediately after writing and no search indexes are created.

Work memory taken during the work was 385538 KB. To analyze the system performance, work memory was chosen to be in a file but not in main memory. In further realizations, to accelerate the speed and reduce of used disk space, work memory may be realized as part of main memory (as dynamically allocated memory or as a file mapped in memory).



**Figure 9.** WordArM results for the case of WordNet as thesaurus

Comparing results in Table 1 and Table 2 we may see that more great number of instances in the case of thesauruses is stored and accessed faster than less number of instances in the case of dictionaries. The main reason for this is that the system needs some time for preparing internal indexing structures. Main part of this time is spent during creating archive and initial empty structures.

After finishing the work, occupied permanent memory for compressed archive is 15603 KB, i.e. in this case NL-indexing takes 14270 KB additional compressed memory (the sequential file with initial data is 1333 KB).

To compare our results with "Grinder" (original WordNet program for generating the database) we provided experiments on separate computer under operating system "Linux open source 12.3". Computer configuration was: Processor: Celeron Dual Core T35.00 2.1 GH; Memory: 2GB; HDD: 230 GB. It is important that Grinder works in single command prompt mode, which permits high speed processing. Multitasking and Graphical User

Interface of MS Windows takes additional resources (processor time and main memory) for supporting its work and this way the work of the user program became slower. Pseudo parallelism of Windows 7 processes is additional factor for delay of applied programs work.

The steps of Grinder program were: Resolving pointers; Getting sense counts; Figuring out byte offsets; Dumping data files; Dumping index files; Dumping sense index. The time used for generating the data base was about 7 seconds. If we assume that access to hard disk blocks is at least 100 times slower than access to main memory blocks, we may accept that WordArM will process the same information for about one second if it is only in the main memory, i.e. working only in main memory, WordArM will store information at least seven times faster.

It was impossible to measure the WordNet access to information in compiled files.

The size of data and index files created by Grinder is 32.3 MB (33 921 109 bytes); compressed by WinZip they have size of 9584 KB, the size of the WordNet compressed archive is 15603 KB, i.e. our archive takes about one and half more disk space.

Analyzing other software systems for maintenance of dictionaries and thesauruses we found that for more of them there exist limitations for maximal length of the concepts which vary from 60 up to 255 bytes [Will & Will, 2004]. As example of the size of thesaurus database we may point the MultiTes system which thesaurus database needs 15 MB of disk space for every 50,000 terms [MultiTes, 2013]. WordArM needs the same disk space for two times more terms.

---

## Conclusion

---

Natural Language Addressing (NLA) is a possibility to access information using natural language words or phrases as direct addresses of the information in the computer memory. For this purpose the internal encoding of the letters is used to generate corresponded address co-ordinates. In this paper we outlined the main idea of NLA and illustrated its genesis.

NLA is applicable for storing dictionaries and thesauruses.

The conclusions about the gain and loss from using NL-Addressing for storing dictionaries and thesauruses are:

- The loss is additional memory for storing internal address structures. But the same if no great losses we will have if we will build balanced search trees or other kind in external indexing. It is difficult to compare with other systems because such information practically is not published.
- The benefit is in two main achievements:
  1. High speed for storing and accessing the information;
  2. The possibility to access the information immediately after storing without recompilation the database and rebuilding the indexes.

---

## Bibliography

---

- [AHD, 2009] The American Heritage® "Dictionary of the English Language" Fourth Edition copyright© 2000 by Houghton Mifflin Company, Updated in 2009; Published by Houghton Mifflin Company. All rights reserved.
- [Angelov, 2012] St. Angelov. SA Dictionary <http://www.thediction.com/> (accessed: 11.01.2013).
- [Collins, 2003] "Collins English Dictionary – Complete and Unabridged", HarperCollins Publishers, 1991, 1994, 1998, 2000, 2003
- [Dict, 2013] Dictionary.com LLC, <http://dictionary.reference.com/> (accessed: 14.04.2013)
- [Faye et al, 2012] David C. Faye, Olivier Cure, Guillaume Blin. A survey of RDF storage approaches. Received, December 12, 2011, Accepted, February 7, 2012, ARIMA Journal, vol. 15 (2012), pp. 11-35.
- [Fellbaum et al, 1998] Fellbaum, Christiane, ed. WordNet: An Electronic Lexical Database/MIT Press, Cambridge, MA, 1998. pp. 422

- [Fellbaum, 1998] Fellbaum Christiane (ed.) WordNet. "An Electronic Lexical Database", ISBN: 978026206197, MA: MIT Press; 1998, pp. 422
- [Ivanova et al, 2012] Krassimira Ivanova, Vitalii Velychko, Krassimir Markov. About NL-addressing (К вопросу о естественно-языковой адрессации). In: V. Velychko et al (ed.), Problems of Computer in Intellectualization. ITHEA@ 2012, Kiev, Ukraine - Sofia, Bulgaria, ISBN: 978-954-16-0061-0 (printed), ISBN: 978-954-16-0062-7 (online), pp. 77-83 (in Russian).
- [Ivanova et al, 2013] Krassimira B. Ivanova, Koen Vanhoof, Krassimir Markov, Vitalii Velychko, "Introduction on the Natural Language Addressing", International Journal of Information Technologies and Knowledge, Vol.7, Number 2, 2013, pp. 139–146. ISSN 1313-0455
- [Miller, 1995] Miller G. A. "WordNet: a lexical database for English"/G. A. Miller. – Communications of the ACM 38: 11, 1995. pp. 39–41.
- [MultiTes, 2013] MultiTes. <http://www.multites.com> (accessed: 18.04.2013)
- [Palagin et al, 2011] Palagin A.V., Krivii S.L., Petrenko N.G. Ontological methods and instruments for processing domain knowledge. (А. В. Палагин, С. Л. Крытый, Н. Г. Петренко. Онтологические методы и средства обработки предметных знаний: монография/Луганск: изд-во ВГУ им. В. Даля, 2011. – 300 с.), (in Russian).
- [PC mag, 2013] PC Magazine Enciclopedia [http://www.pcmag.com/encyclopedia\\_term/0,1237,t=indexing&i=44896,00.asp](http://www.pcmag.com/encyclopedia_term/0,1237,t=indexing&i=44896,00.asp) (accessed: 23.01.2013)
- [Poprat et al, 2008] Poprat Michael, Elena Beisswanger, Udo Hahn. Building a BioWordNet by Using WordNet's Data Formats and WordNet's Software Infrastructure — A Failure Story. Software Engineering, Testing, and Quality Assurance for Natural Language Processing, Columbus, Ohio, USA, June 2008. Association for Computational Linguistics, 2008, pp. 31–39.
- [Will & Will, 2004] L.D. Will, S.E. Will. Willpower Information. <http://www.willpowerinfo.co.uk/thestabl.htm> (accessed: 18.04.2013)
- [WordNet, 2012] Princeton University "About WordNet", WordNet, Princeton University, 2010 <http://WordNet.princeton.edu> (accessed: 23.07.2012)

---

## Authors' Information

---



**Ivanova Krassimira** – University of National and World Economy, Sofia, Bulgaria. Institute of Mathematica and Informatics, BAS, Sofia, Bulgaria.

e-mail: [krazy78@mail.bg](mailto:krazy78@mail.bg)

Major Fields of Scientific Research: Software Engineering, Business Informatics, Data Mining, Multidimensional multi-layer data structures in self-structured systems



**Vanhoof Koen** - Universiteit Hasselt, Belgium

e-mail: [koen.vanhoof@uhasselt.be](mailto:koen.vanhoof@uhasselt.be)

Main research areas: data mining, knowledge retrieval.



**Markov Krassimir** – Institute of Mathematics and Informatics, BAS, Bulgaria

e-mail: [markov@foibg.com](mailto:markov@foibg.com)

Major Fields of Scientific Research: General theoretical information research, Multi-dimensional information systems



**Velychko Vitalii** – Institute of Cybernetics, NASU, Kiev, Ukraine

e-mail: [Velychko@rambler.ru](mailto:Velychko@rambler.ru)

Major Fields of Scientific Research: Data Mining, Natural Language Processing

## ЗАДАЧА ВЫБОРА ТОПОЛОГИЧЕСКОЙ СТРУКТУРЫ СЕТИ ДОСТУПА

Галина Гайворонская, Антон Бондаренко

**Аннотация:** показаны причины, определяющие сложность и нетривиальность задачи синтеза сетей доступа и предложены некоторые подходы к решению одной из задач проектирования этих сетей, а именно к выбору топологической структуры сети доступа. При этом особое внимание уделено анализу возможности использования хордовых кольцевых структур.

**Ключевые слова:** телекоммуникационная сеть, сеть доступа, топология, узел доступа, сегмент локального доступа, хордовые кольцевые структуры.

**Ключевые слова классификации АСМ:** С.2. Computer-communication networks, Н. Information Systems - Н.1 Models and Principles, К. Computing Milieux - К.6 Management of computing and information system.

---

### Введение

Значительные усилия исследователей направлены на повышение эффективности использования сетевых ресурсов, чтобы найти приемлемый компромисс между необходимыми характеристиками сети и затратами на ее сооружение. Специфика теоретической модели, представляющей телекоммуникационную сеть (ТС) в задачах синтеза ее оптимальной структуры, определяет как постановку задач, так и методы их решения. Типичная задача синтеза оптимальной первичной сети предполагает следующую постановку [Давыдов, 1977]. Задана матрица каналов между пунктами ввода и вывода информации и указана дислокация этих пунктов, задан набор типовых систем передачи и линий связи. Требуется построить сеть линий между этими пунктами, обеспечивающую реализацию заданной матрицы каналов при ограничениях, налагаемых на структуру сети географическими особенностями, технологиями обслуживания и алгоритмами работы сети, спецификой используемого оборудования и ожидаемых внешних воздействий, а также требованиями к качеству работы сети и ее экономическим параметрам [Давыдов, 1985].

Синтезу топологии первичных сетей посвящено большое количество работ, в частности, Г. Б. Давыдова, В. Н. Рогинского, В. И. Неймана, А. Д. Харкевича, С. А. Аджемова, В. Г. Лазарева, Л. Клейнрока, Г. П. Захарова, Дж. Мартина, А. Я. Толчана, Х. Иноссе и многих других. Однако анализ опубликованных работ показывает, что синтез первичной сети в самой постановке задачи не предполагает выбор местоположения узлов коммутации (УК). Синтезу коммутируемых сетей, подразумевающему определение местоположения УК, их взаимосвязь и взаимоподчинение, уделено гораздо меньше внимания в отечественных и зарубежных публикациях. При создании сетей доступа необходимо синтезировать топологию именно коммутируемой сети, характеризуемой некоторыми особенностями, отличающими ее от топологических структур, используемых ранее.

Особенности СД связаны с тем, что в соответствии с рекомендацией G.902, сеть доступа состоит из трех основных элементов, размещенных между оборудованием пользователя и узлами, предоставляющими обслуживание (УПО) [Гайворонская, 2010]. Эти элементы включают сегмент локального доступа (СЛД),

---

узел доступа (УД) и сегмент транспортного доступа (СТД). Сегмент локального доступа СД характеризуется максимальной гарантированной пропускной способностью и минимально допустимой длиной физической линии доступа (ЛД). Узел доступа реализуют на оборудовании, концентрирующем информационные потоки от индивидуальных линий локального доступа к различным базовым сетям. Сегмент транспортного доступа – это групповые тракты передачи информации между УД и УПУ базовых сетей, и функционально является частью национальной транспортной сети. Узел предоставления услуг реализуют в виде универсального сетевого элемента, способного поддерживать все требования пользователей.

Тенденции эволюции трех основных параметров СД: пропускной способности, топологии и среды передачи проанализированы проф. Соколовым в работе [Соколов, 1999]. Перспективная СД должна обеспечивать полосу пропускания, необходимую пользователю, поэтому верхний предел скорости передачи будет измеряться не только сотнями Гбит/с, но и Тбит/с. Кроме этого отчетливо проявляется диверсификация сред распространения сигналов, которые могут применяться для создания и развития СД, но доминировать все же будет оптическое волокно. Реализация СД сопровождается появлением новых топологий, среди которых предпочтение отдается кольцевым и комбинированным структурам. Достаточно перспективным может оказаться применение хордовых, и решетчатых структур, что подразумевает создание линий межузловой связи. Такие линии могут понадобиться, если необходимо подключить пользователя, требующего набор услуг существенно отличающийся от услуг, заказанных остальными пользователями, обслуживаемыми тем же УД. Создать их можно путем прокладки новой кабельной канализации, что связано с большими затратами, и подвержено влиянию множества факторов, которые могут значительно затруднить этот процесс. Поэтому эффективность использования межузловых линий в СД требует дальнейших исследований, некоторые результаты которых отображены в этой работе.

---

### Выбор топологической структуры сети доступа

---

Задачи синтеза сетей весьма сложны, причем сложность резко увеличивается по мере увеличения масштаба сети. Поскольку рассматривается задача, в которой синтезируется новая СД, процесс ее проектирования отображает практическую реализацию решения задачи синтеза. Для которой выполнена декомпозиция и выделены отдельные этапы её решения. Первоначально этот подход предложен в работе [Гайворонская, 2012], его развитие представлено в [Гайворонская, 2013]. Предлагаемая авторами последовательность и взаимосвязь задач, решаемых на отдельных этапах проектирования сетей доступа, соответствующих концепции, сформулированной в рекомендации G.902 и последующих, развивающих ее рекомендациях МСЭ, показана на рис.1.

Анализ задач каждого этапа и методов их решения приведен в [Гайворонская, 2013]. При этом исходные параметры, используемые при проектировании СД, предлагается разделить на три группы, характеризующие территорию, пользователей и базовые сети, с которыми предполагается взаимодействие. Поскольку эти параметры отличаются большим разнообразием и плохо поддаются классификации в работе [Ганницкий, 2011] предложен подход повышающий эффективность предварительной обработки исходных данных, необходимых при проектировании СД. Выбор топологической структуры СД является, гораздо более сложной задачей, чем синтез структуры существующих ТС [Гайворонская, Бондаренко, 2013]. Поэтому решению этой задачи посвящен отдельный этап, при реализации которого выполняется фрагментирование территории, обслуживаемой СД [Гайворонская, 2013; Гайворонская, Бондаренко, 2013], выбирается способ подключения УД

(одноуровневый или двухуровневый) и местоположение этих узлов, оценивается целесообразность поперечных связей и определяется схема прокладки линий доступа.

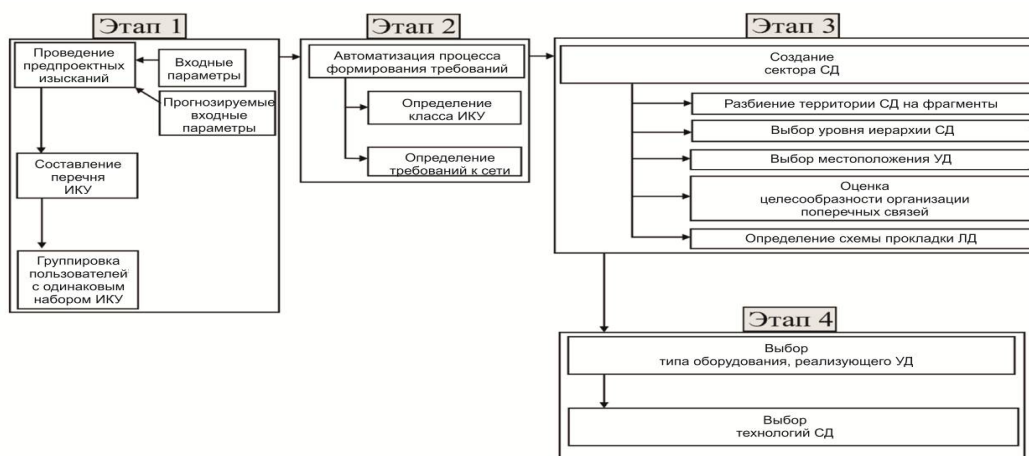


Рисунок 1. Последовательность и взаимосвязь задач задач, решаемых при

Возможным подходам решения совокупности задач третьего этапа посвящена данная работа, целью которой является повышение эффективности проектирования СД, объектом – топологическая структура СД, а предметом – методы оптимального синтеза ТС.

Поскольку основным требованием к топологической структуре СД является минимизация длины СЛД при гарантированной пропускной способности, основополагающим при выборе топологии СД является выбор местоположения УД. Некоторые методы решения задач определения количества и местоположения УД приведены в [Гайворонская, Котова, 2010; Гайворонская, Сахарова, 2011; Гайворонская, 2011; Гайворонская, Бондаренко, 2011]. Реализация основных требований, выдвигаемых МСЭ, ETSI и другими стандартизационными организациями к структуре СД приводит к увеличению территории, обслуживаемой этой сетью, и к достаточно сложным структурам сети, сочетающим различные топологии. Пример топологии СД показан на рис. 2.

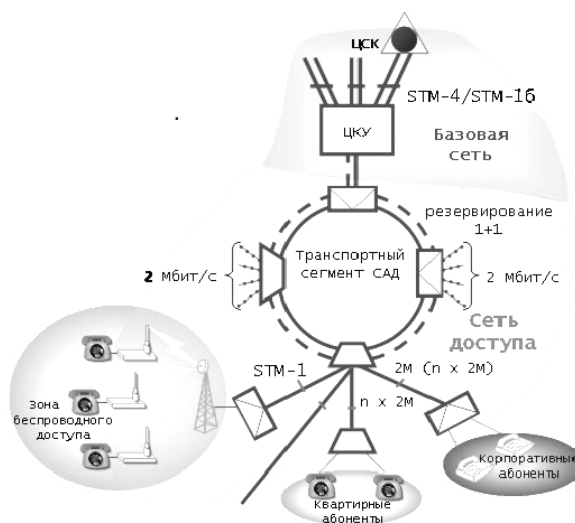


Рисунок 2. Пример структуры СД

Для выбора топологий, сочетание которых предполагается использовать на проектируемой СД, предложено использовать методы теории принятия решений, при этом решение задачи может быть автоматизировано с помощью программного продукта «Системы и методы принятия решения» (СМПР), реализованного под руководством д.т.н. проф. Волошина А.Ф. (Киевский национальный университет им. Т.Г. Шевченко) [Волошин, 2010].

Для выбора топологической структуры СД необходимо проанализировать целесообразность применения решетчатых, сетчатых и хордовых топологий. В этой работе остановимся только на хордовых кольцевых сетях, в которых наряду с каналом, соединяющим все узлы в кольцо (кольцевой канал), имеются прямые каналы между несмежными узлами кольца (каналы-хорды).

### Анализ возможности применения хордовых топологий в сетях доступа

В последние годы появился ряд работ, посвященных исследованию хордовых кольцевых сетей. Хордовые сети с однонаправленными каналами, в которых в каждом узле  $i$  имеются два исходящих и два входящих канала (рис. 3,а) исследовались в [Wong, 1974; Raghavendra, 1985; Hwang, 1987; Fiol, 1987].

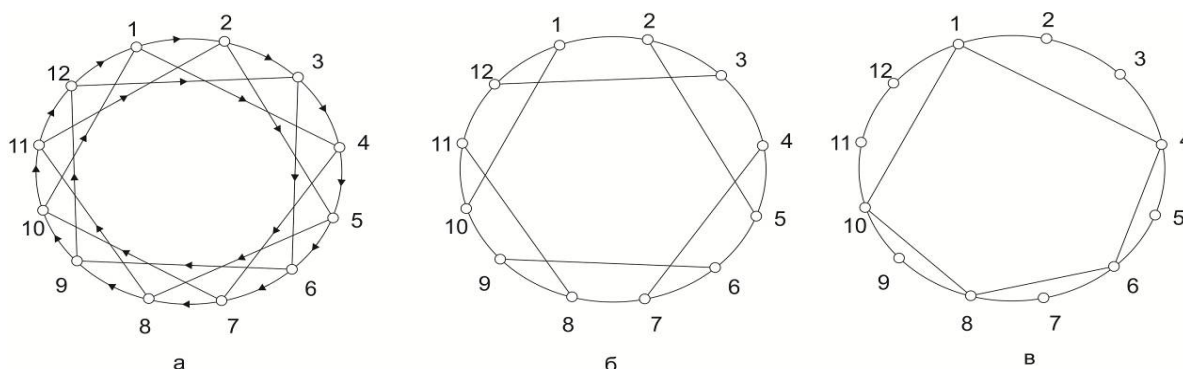


Рисунок 3. Кольцевые хордовые структуры

Исходящие каналы направлены к узлу  $(i + 1)$  и к узлу  $(i + b)$ , при этом  $b$  имеет одно и то же значение для всех  $i$ , где  $b$  – шаг хорды. Цель опубликованных исследований – предложить алгоритм выбора параметра  $b$ , минимизирующий максимальное и среднее расстояния между узлами хордовой сети. При этом под расстоянием  $d(i, j)$  между двумя узлами  $i$  и  $j$  понимается минимальное число шагов, за которое из узла  $i$  можно попасть в узел  $j$ . Диаметр сети считаем максимальное расстояние между ее узлами:  $D = \Delta_{ij} \max d(i, j)$ . Среднее расстояние между узлами определяется из выражения

$d = \sum_{i=1}^N \sum_{j=1}^N d(i, j) / N^2$ , где  $N$  – количество узлов. В [Wong, 1974] показано, что при  $b \approx \sqrt{N}$  диаметр

$$D \approx 2\sqrt{N}, \quad d \approx \sqrt{N}.$$

Примерно такие же результаты выводятся в [Raghavendra, 1985] для  $b \approx -\sqrt{N}$ , здесь показано, что это существенно лучше, чем при  $b = -1$  и при  $b = -2$ . В [Hwang, 1987] говорится об алгоритме выбора  $b$ , обеспечивающем  $D \approx \sqrt{3N} + 2\sqrt[4]{3N}$  для всех значений  $N$ , а в [Fiol, 1987] приведены полученные с помощью моделирования абсолютно лучшие значения  $D \approx \sqrt{3N}$  и  $d \approx 5\sqrt{3N} / 9$  с соответствующими

им значениями  $b$  для всех  $N \leq 256$ . В [Raghavendra, 1985; Fiol, 1987] приводится сравнение показателей  $D$  и  $d$  для колец, реализующих двойные кольцевые топологии. В [Arden, 1981] исследованы хордовые кольцевые сети с двунаправленными каналами и четным числом узлов (рис. 3,б), при этом сравниваются значения диаметра при различном  $b$  и показано, что при всех  $b$  выполняется соотношение  $D \geq \sqrt{N} - 1$ . Из любого узла  $i$  такой сети исходят три однонаправленных канала к узлам  $(i + 1)$  и  $(i - 1)$ , а также к узлу  $(i + b)$  либо  $(i - b)$ . При анализе двойных кольцевых топологий показано, что суммарная протяженность хордовых каналов существенно зависит от значения  $b$ . Если каждый канал математической модели сети является отдельным физическим каналом, стоимость которого пропорциональна его длине, такое соображение становится весьма существенным, если не решающим. В частности, если все каналы проложены в одном кабеле, то суммарная протяженность каналов двойной кольцевой топологии составляет в точности  $(|b| + 1)$  длин кольца.

В такой ситуации нас может интересовать, каким образом минимизировать диаметр и среднее расстояние между узлами хордового кольца (ХК). С точки зрения топологии СД, нас интересуют только двунаправленные ХК, в которых хорды следуют цепочкой друг за другом, образуя внутреннее хордовое кольцо (рис. 3в). В [Андреев, 1989] показано, что в двунаправленных ХК, имеющих  $n$  хорд с шагом  $s$  и содержащих  $N = ms$  узлов, диаметр вычисляется по формуле  $D = s + \lceil m/2 \rceil - 1$ , где  $\lceil m/2 \rceil$  - целая часть  $m/2$ .

Чтобы выбрать наиболее целесообразный диаметр кольца и расстояние между узлами, выполним аналитическое исследование модели хордовой сети, показанной на рис. 4.

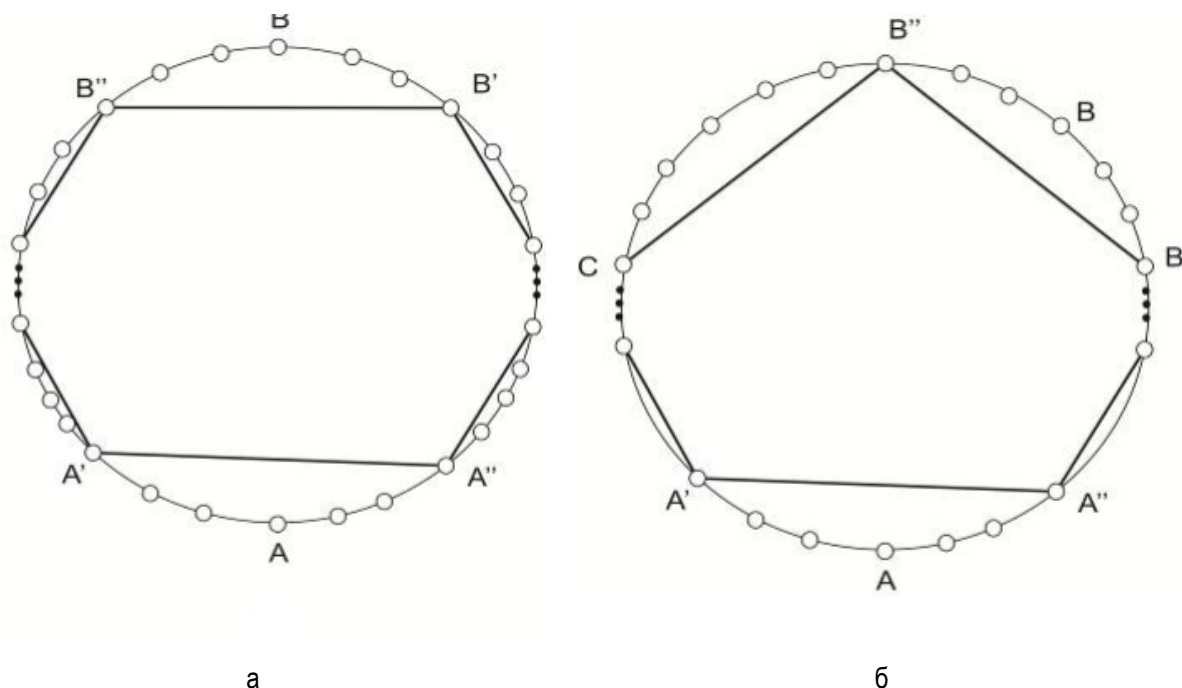


Рисунок 4. Хордовые кольца

Рассмотрим случай, когда  $n$  четно. При четном  $s$  вычислим расстояние между узлами  $A$  и  $B$ , находящимися точно посередине диаметрально противоположных дуг  $A'A''$  и  $B'B''$  (рис. 4а). Узлы, принадлежащие и к основному, и к хордовому кольцам (например,  $A'$ ,  $A''$ ,  $B'$ ,  $B''$ ), назовем узлами 2-го



порядка, а все прочие – узлами 1-го порядка. Поскольку и  $A$ , и  $B$  отстоят от ближайших узлов 2-го порядка на  $s/2$  шагов и расстояние между ними ( $A'B''$  или  $A''B'$ ) равно  $(n/2) - 1$ , в итоге имеем  $d(A, B) = s + \frac{m}{2} - 1$ . Легко убедиться, что расстояние между любыми другими двумя узлами сети не превышает  $d(A, B)$ . При  $s$  нечетном выберем  $A$  и  $B$  также из диаметрально противоположных хорд и на максимальном  $(s-1)/2$  отдалении от ближайших узлов 2-го порядка  $A'$  и  $B''$ , которые тоже диаметрально противоположны. Получаем  $d(A, B) = s - 1 + \frac{m}{2}$  поскольку маршруты  $AA'B'B$ ,  $AA''B''B$ ,  $AA''B'B$  все имеют такую длину. Результирующие выражения при четном и нечетном  $s$  совпадают, и нетрудно убедиться, что более удаленных пар узлов в сети нет.

Пусть теперь  $n$  нечетно. При четном  $s$  наибольшее расстояние в сети наблюдается между узлами  $A$  и  $B$ , находящимися точно посередине дуг  $A'A''$  и  $B'B''$ , максимально отстоящих друг от друга. По дугам  $AA''$  и  $BB''$  путь составляет дважды по  $s/2$  шагов плюс  $\left(\frac{m-1}{2} - 1\right)$  шагов по хордам от  $A''$  до  $B'$ , что в итоге дает  $d(A, B) = s + \frac{m-1}{2} - 1 = s + \lfloor m/2 \rfloor - 1$ , поскольку более короткого маршрута нет. А при нечетном  $s$  на максимальном удалении находятся, например, узлы  $A$  ( $\frac{s-1}{2}$  шагов до  $A'$ ) и  $B$  ( $\frac{s-1}{2}$  шагов до  $B''$ ) в наиболее отдаленных друг от друга дугах  $AA''$  и  $BB''$ , причем от  $A'$  до  $B''$  по хордам  $\frac{m-1}{2}$  шагов. Так что  $d(A, B) = s - 1 + \frac{m-1}{2} = s + \lfloor m/2 \rfloor - 1$ , поскольку другие пути от  $A$  до  $B$  не короче.

В двунаправленных ХК, имеющих нечетное число хорд с шагом  $s$  и числом узлов  $N = ms$ , среднее расстояние между узлами вычисляется по формуле

$$d = (((2s + m)m - s)s / 4N) - \begin{cases} (m + s/2 - 1) / N, \text{ при } s \text{ четном} \\ ((m + s) / 2 + (s + 1) / 4s) / N, \text{ при } s \text{ нечетном} \end{cases} \quad (1)$$

Обозначив через  $R(N)$  сумму расстояний от какого-либо узла кольцевой сети с одним каналом и  $N$  узлами до всех узлов сети, при нечетном  $N$  получим

$$R(N) = 2 \left( 1 + 2 + \dots + \frac{N-1}{2} \right) = (N-1)(N+1) / 4, \quad (2)$$

а при четном  $N$

$$R(N) = 2 \left( 1 + 2 + \dots + \left( \frac{N}{2} - 1 \right) \right) + \frac{N}{2} = N^2 / 4. \quad (3)$$

В рассматриваемых двунаправленных ХК под наиболее отдаленными парами дуг будем понимать такие пары дуг, между которыми пролегает  $\lfloor m/2 \rfloor - 1$  хорд. При нечетном числе хорд (рис. 4б) каждая дуга  $A'A''$  входит в две пары наиболее отдаленных дуг – ( $A'A''$ ,  $B'B''$ ) и ( $A'A''$ ,  $CB''$ ). Но и при четном, и при нечетном  $n$  таких пар будет ровно  $n$ . Вычислим сумму длин дуг, входящих в каждую такую пару, и максимальное значение этой величины обозначим через  $S$ . Для хордовых колец из предыдущих утверждений, понятно, что  $S = 2s$ . В двунаправленных ХК с  $n$  хордами с максимальным шагом  $s$  и минимальным шагом  $(s-1)$  диаметр вычисляется по формуле  $D = \lfloor (S + m) / 2 \rfloor - 1$ . При этом следует

учесть, что при произвольных длинах дуг наиболее удаленные узлы вовсе не обязательно лежат в наиболее отдаленных парах дуг. Среди двунаправленных ХК с произвольным фиксированным числом узлов  $N$  минимальный возможный диаметр  $D(N) = \lfloor \sqrt{2N + 1/4} - 1/2 \rfloor$  достигается среди хордовых и кольцевых сетей с примерно равными дугами.

В таблице 1 для каждого значения диаметра  $D$  вычислено максимально возможное число узлов  $N$  двунаправленного ХК, представлены параметры соответствующей сети с нечетным числом  $n$  дуг равной длины  $s$ , а также вычислено среднее расстояние между узлами  $d$  такой сети. Для сравнения показаны минимальные значения  $D$  и  $d$  для двойных кольцевых топологий с тем же числом узлов, а также на каких однонаправленных хордовых кольцах они достигаются. Анализируя результаты, приведенные в таблице, можно сделать вывод о том, что для произвольного числа узлов  $N$  среднее расстояние между этими узлами, при числе дуг равно  $m_s = N - (s(N) - 1)m(N)$ , составляет  $d \approx \sqrt{\frac{N}{2}}$ .

Таким образом, доказано, что при любом  $N$  имеются хордовые сети с диаметром  $D < \sqrt{2N}$  и расстоянием между узлами  $d \approx \sqrt{\frac{N}{2}}$ , что существенно лучше аналогичных показателей в двойных кольцевых топологиях, достигаемых к тому же при куда большей суммарной длине, а следовательно, и стоимости.

**Таблица 1.** Вычисления максимально возможного числа узлов  $N$  двунаправленного ХК

Двунаправленное хордовое кольцо					Двойная кольцевая топология		
$N$	$s$	$n$	$D$	$d$	$D$	$d$	$N$
3	1	3	1	0,66667	1	0,66667	2
6	2	3	2	1,08333	3	1,5	3
10	2	5	3	1,65	4	2,1	4
15	3	5	4	2,35556	5	2,8	5
21	3	7	5	2,92063	6	3,42857	10
28	4	7	6	3,32143	8	4,21428	6
36	4	9	7	3,86111	9	4,83333	12
45	5	9	8	4,46222	10	5,55556	21
55	5	11	9	4,99636	11	6,18182	22
66	6	11	10	5,41667	13	6,86364	16
78	6	13	11	5,94231	14	7,57692	19
91	7	13	12	6,50863	15	8,18681	18
105	7	15	13	7,03129	16	8,93333	30
120	8	15	14	7,46667	17	9,55	20
136	8	17	15	7,98529	19	10,32353	33
153	9	17	16	8,5345	20	10,94118	22
171	9	19	17	9,05138	21	11,68421	38
190	10	19	18	9,49737	23	12,3421	27
210	10	21	19	10,0119	24	13,07143	41
231	11	21	20	10,55096	25	13,80952	45
253	11	23	21	11,06432	26	14,43478	46

Простейшим вариантом использования таких структур в сетях доступа, является модель, представленная на рис. 5, где изображен пример реализации СД, обеспечивающей доступ к трем базовым сетям, представленными соответствующими УПО.

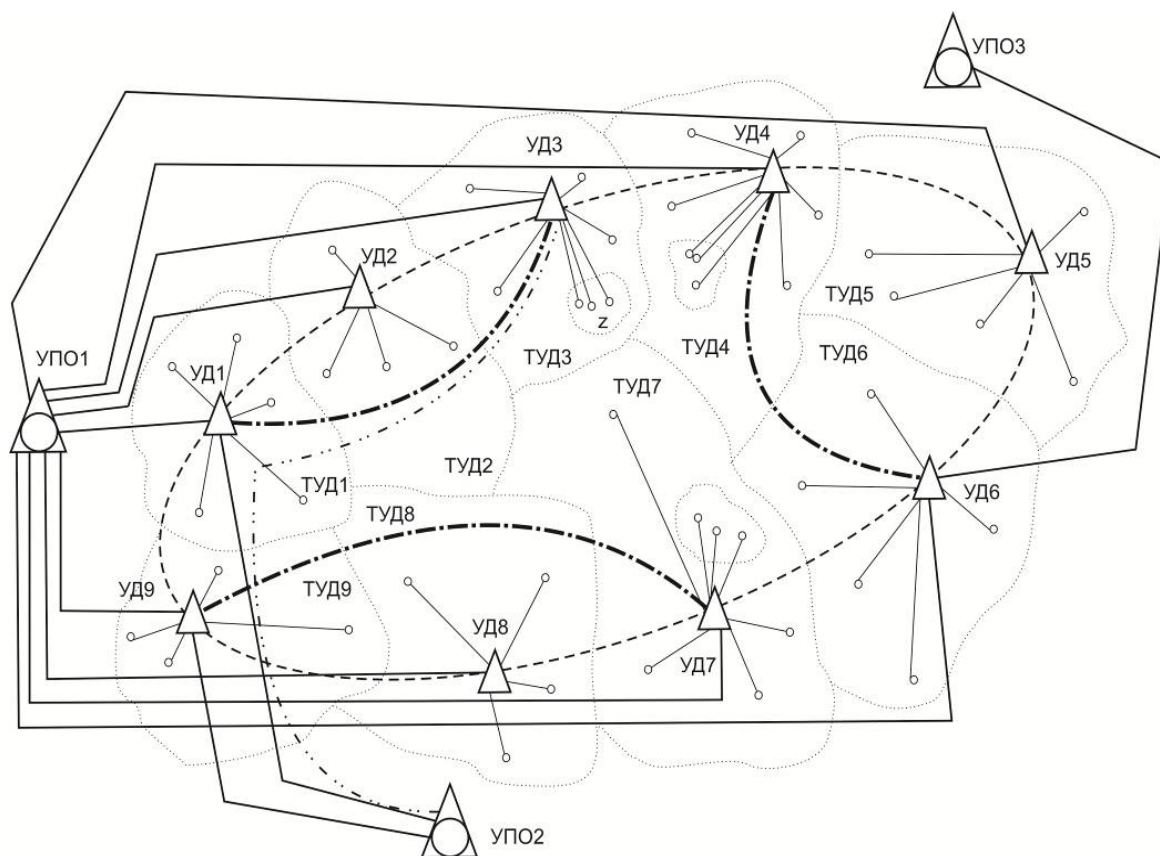


Рисунок 5. Пример реализации СД

На сети доступа организовано девять УД (условно соединенные по топологии кольцо, обозначенное штриховой линией), кроме этого реализованы три хорды: УД1-УД3, УД6-УД4, УД9-УД7 (обозначенные штрих-пунктирной линией), организация которых обусловлена тем, что на территории, обслуживаемой УД3 (ТУД3) расположены пользователи, набор услуг которых требует подключения к базовым сетям, связанным с УД1 (т.е. относящиеся к ТУД1). Аналогично на ТУД4 располагаются пользователи, с набором услуг требующим подключение к УД6, а на ТУД7 – к УД9. Путь предоставления услуг пользователям находящимся в секторе Z, размещенном на территории, обслуживаемой УД3 (ТУД3) с помощью УПО2 с которым УД3 не имеет связи, реализованный с помощью хорды, обозначен на рисунке линией штрих-точка-точка. Используя выражения, полученные выше нетрудно получить значения диаметра кольца и среднее расстояние между узлами  $D \leq 4$  и  $d = 2$ , минимизирующие общую протяженность линий доступа для рассматриваемого примера сети.

Определив местоположение УД и структуру связей между ними выбираем наиболее целесообразные пути прокладки ЛД от УПУ к УД и от УД к пользователям. Для этого в рамках сотрудничества с Киевским институтом кибернетики им. В.М. Глушкова разработан метод выбора маршрута прокладки ЛД, основанный на модели балансных сетей, предложенной д.т.н., проф. Гладуном В.П. [Гладун, 2000]. Реализация этого метода рассмотрена в [Величко, 2012].

---

## Заклучение

---

В работе предложено сочетание методов, последовательное применение которых позволит выбрать наиболее целесообразную топологическую структуру сети доступа, что даст возможность повысить эффективность и увеличить точность проектирования этих сетей. Дальнейшие исследования, проводимые коллективом кафедры ИКТ, посвящены решению задач четвертого этапа проектирования СД – выбору оборудования и транспортных технологий доступа. В результате планируется завершить разработку целостной методики проектирования СД, что позволит оптимизировать процесс создания сетей следующего поколения *Next Generation Networks (NGN)*.

---

## Литература

---

- [Arden, 1981] Arden B.W. Analysis of Chordal Ring Network / Arden B.W., Lee H. // IEEE Trans. Comput., 1981. – Vol.30, №4. – p.291-295
- [Fiol, 1987] Fiol M.A. A discrete optimization problem in local networks and data alignment / Fiol M.A., Yebra J.L.A., Alegre I., Valero M. // Ibid, 1987. – Vol.36, №6 – p.702-713
- [Hwang, 1987] Hwang F.K. Comments on « Reliable loop topologies for large local computer networks» // Ibid, 1987. – Vol.36, №3. – p.383-384
- [Raghavendra, 1985] Raghavendra C.S. Reliable loop topologies for large local computer networks / Raghavendra C.S., Gerla M., Avizienis A. // Ibid, 1985. – Vol.34, №1. – p.46-55
- [Wong, 1974] Wong C.K., Coppersmith D.A. Combinatorial problem related to multimodule memory organizations // JACM, 1974. – Vol.21, №9. – p.392-402
- [Андреев, 1989] Андреев Л.В. Хордовые кольцевые сети с двумя дуплексными каналами // Сборник научных трудов «Теория телетрафика в системах информатики». М: «Наука», 1989. – с.136-144
- [Величко, 2012] Величко В.Ю. Анализ возможности использования модели балансных сетей при проектировании сетей доступа / Величко В.Ю., Бондаренко А.А./ International Journal "Information Theories and Knowledge". – Sofia: ITNEA, 2012. – № 2 (Volume 6) – P. 126-130.
- [Волошин, 2010] Волошин О.Ф. Моделі та методи прийняття рішень: навчальний посібник для студ. вищ. навч. закл. / О.Ф.Волошин, С.О.Машенко // 2-ге вид. – К.: Видавничо-поліграфічний центр «Київський університет». – 2010.– 336 с.
- [Гайворонская, 2010] Гайворонская Г.С. Основные задачи модернизации сетей пользовательского доступа / Г.С. Гайворонская, А.И. Котова // Зв'язок. – 2010. – №2 (90). –С.32-36.
- [Гайворонская, 2011] Гайворонська Г.С. Особенности определения местоположения узлов доступа при использовании радиальной модели обслуживаемой территории. / Г.С. Гайворонская, С.В. Сахарова // Наукові праці ДонНТУ. – Донецьк: ДонНТУ, 2011.– №21 (183). – С. 82-86.
- [Гайворонская, 2012] Гайворонская Г.С. Основные принципы проектирования сетей доступа / Г.С. Гайворонская, С.В. Сахарова // Зв'язок. – 2012. – №4 (100). – С.22-27
- [Гайворонская, 2013] Гайворонская Г.С. Разработка методики проектирования сетей доступа / Г.С. Гайворонская, А.А. Бондаренко // Сучасний захист інформації. ДУІКТ. – 2013. – № 1. – С. 81-86
- [Гайворонская, Бондаренко, 2011] Гайворонская Г.С. Программная реализация проектируемой ортогональной сети доступа / Г.С. Гайворонская, С.В. Сахарова, А.А. Бондаренко // Комп'ютерні системи та мережі, № 11. – 2012.
- [Гайворонская, Бондаренко, 2013] Гайворонская Г.С. Метод фрагментации территории, обслуживаемый сетью доступа/ Г.С. Гайворонская, А.А. Бондаренко // Вісник ДУІКТ. – 2013. – №2. – С. 25-30

- [Гайворонская, Котова, 2010] Гайворонська Г.С. Метод визначення довжини ліній доступу для різноманітних конфігурацій території обслуговування / Гайворонська Г.С., Котова О.І., Сахарова С.В. // Вісник Нац. ун-ту «Львівська політехніка». – Львів, 2010. – № 688: Комп'ютерні системи та мережі. – С.65-69.
- [Гайворонская, Сахарова, 2011] Гайворонська Г.С. Метод определения местоположения узлов при использовании прямоугольной модели сети доступа / Г. С. Гайворонская, С.В. Сахарова // Холодильна техніка і технологія.– Одеса: ОДАХ, 2011.– №1 (129).–с.73-76.
- [Ганницкий, 2011] Ганницкий І.В. Підвищення швидкості обробки результатів вимірювань параметрів потоку викликів / Ганницкий І.В., Бондаренко А.А./ «Холодильна техніка і технологія». – Одеса, ОДАХ, 2011. – №1 (129), с. 69-72
- [Гладун, 2000] Гладун В.П., Партнерство с компьютером. Человеко-машинные целеустремленные системы, «Port Royal», Киев, 2000. – 128с
- [Давыдов, 1977] Давыдов Г. Б., Рогинский В. Н., Толчан А. Д. Сети электросвязи. – М.: Связь, 1977. – 380 с.
- [Давыдов, 1985] Давыдов Г.Б. Некоторые проблемы оптимизации развития сетей//Электросвязь.– 1985.– №2.– С. 1-5
- [Соколов, 1999] Соколов Н. А. Сети абонентского доступа. – Пермь: ИПК Звезда, 1999. – 154 с.

---

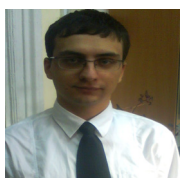
### Информация об авторах

---



**Галина Гайворонская** – Факультет Информационных технологий и кибербезопасности ОНАПТ, д.т.н., профессор, зав. кафедрой информационно-коммуникационных технологий, советник ректора по инфокоммуникациям; Украина, Одесса, 65026, ул. Дворянская, 1/3; тел. (048)-720-91-48; e-mail: [gsgayvoronska@gmail.com](mailto:gsgayvoronska@gmail.com)

*Области научных исследований: оптимизация переходных периодов при эволюции телекоммуникационных сетей. Потоки вызовов, нагрузка и межузловое тяготение в сетях. Проблемы создания сетей доступа. Проблема построения полностью оптических сетей и систем коммутации*



**Антон Бондаренко** – Факультет Информационных технологий и кибербезопасности ОНАПТ, аспирант кафедры информационно-коммуникационных технологий; Украина, Одесса, 65026, ул. Дворянская, 1/3; тел. (048)-720-91-48; e-mail: [divlaine@mail.ru](mailto:divlaine@mail.ru)

*Области научных исследований: Сети доступа: проблемы их создания и методы проектирования. Теория и методы принятия решений.*

## КАЧЕСТВО УПРАВЛЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫМИ УСЛУГАМИ В СЕТЯХ ПОСЛЕДУЮЩЕГО ПОКОЛЕНИЯ

Сергей Шестопапов

**Аннотация:** *Выполненные исследования относятся к области проектирования интеллектуальной надстройки в сетях последующего поколения (NGN) для повышения качества управления интеллектуальными услугами. Работа посвящена анализу существующих принципов построения интеллектуальной надстройки, разработке интеллектуальной надстройки с децентрализованным принципом управления, построению ее аналитической модели и результирующего критерия качества управления интеллектуальными услугами. Приведены результаты исследований интеллектуальных надстроек с централизованным и децентрализованным принципами управления, указаны области их эффективного применения.*

**Ключевые слова:** *сеть последующего поколения, интеллектуальная надстройка, критерий качества.*

**Ключевые слова классификации АСМ:** *C.2. Computer-communication networks, H. Information Systems - H.1 Models and Principles, K. Computing Milieux - K.6 Management of computing and information system*

---

### Введение

В XXI веке в Украине наблюдается стремительное развитие сферы телекоммуникаций. Операторы предлагают широкий диапазон услуг, которые способны удовлетворить потребности любого пользователя. Значительно вырос спрос на интеллектуальные услуги. Предоставлять подобные услуги способна новая платформа, в основу которой входит интеллектуальная надстройка. Учитывая рост спроса на интеллектуальные услуги, одним из важнейших вопросов в сфере телекоммуникаций можно считать улучшение качества управления услугами за счет разработки принципов создания интеллектуальной надстройки в NGN.

Анализ научно-технической литературы показывает, что исследования принципов создания интеллектуальной надстройки непосредственно связаны с анализом архитектуры сети, принципом построения системы управления и разработкой результирующего критерия качества управления интеллектуальными услугами. Архитектуре NGN посвящены работы Б.С. Гольдштейна, О.Б. Гольдштейна [1,2], О.О. Атцика, О.В. Пинчука, Антоняна, Е.М. Скуратовской, С.И. Остроха, О.И. Ефремова. Свои архитектурные NGN-решения разработали такие лидеры телекоммуникационного рынка как Alcatel, Ericsson, Lucent Technologies, Siemens, российский производитель – НТЦ "Протей". Вопросы относительно управления сетями, услугами, оценивания эффективности функционирования систем управления, разработки результирующего критерия качества поднимаются в работах В.К. Стеклова [3], В.Г. Кривуцы, Л.Н. Беркман, Н.О. Князевой, Е.В. Кильчицкого, Костика, Е. Штейнберга, Н. Стародуба, Н.Я. Паршенкова. Качество услуг, поддерживаемых сетью NGN, рассматривали такие авторы как П. Фергюсон, Г. Хастон. Исследованию параметров разных трафиков посвящены работы Д. Хеймана, Х. Елгебали.

---

В последнее время в NGN, которые уже функционируют в Италии, Японии, достаточно успешно использовалась интеллектуальная надстройка с централизованным принципом управления (ИНЦПУ). Однако при росте спроса на интеллектуальные услуги ИНЦПУ не сможет качественно выполнять свои функции: будет возникать проблема, связанная с ограниченными пропускной способностью сети сигнализации и производительностью центров управления услугами, вырастет задержка предоставления услуги и так далее. При этом почти не проводятся исследования, посвященные совершенствованию интеллектуальной надстройки за счет децентрализации. Среди нерешенных проблем остается задача разработки принципов создания интеллектуальной надстройки с децентрализованным принципом управления (ИНДПУ). Недостаточно исследовано проблему разработки критерия качества управления интеллектуальными услугами. Учитывая ситуацию, которая сложилась в настоящий момент в сфере телекоммуникаций и растущий спрос на самые современные услуги, можно утверждать, что работа, которая посвящена исследованию и разработке принципов создания интеллектуальной надстройки в NGN, что будет способствовать повышению качества управления интеллектуальными услугами, является актуальной.

---

### **Исследование принципов создания интеллектуальной надстройки NGN и критерия качества. Постановка задачи**

---

В современных сетях NGN используется ИНЦПУ. При такой архитектуре существует единственный центр управления. В интеллектуальную надстройку, которая отвечает за управление интеллектуальными услугами, входит часть Softswitch, которая выполняет функцию коммутации услуги SSF, и сервер, который выполняет функцию управления услугой SCF. Допускается, что существует несколько территориально разнесенных районов. Управление районной сетью осуществляет установленный Softswitch. При росте спроса на интеллектуальные услуги может случиться, что ИНЦПУ не будет способна качественно выполнять свои функции, поскольку будут возникать проблемы, связанные с ограниченной пропускной способностью сети сигнализации и увеличением задержки предоставления услуги. Поэтому нужно искать более эффективное решение. Для управления услугами в IN, которая была предшественницей NGN и довольно существенно на нее повлияла, иногда предлагалась интеллектуальная надстройка с распределенным принципом управления (ИНРПУ). Архитектура сети с ИНРПУ позволяет решить часть проблем, существующих при ИНЦПУ. Прежде всего - уменьшить общее время обслуживания заявки на интеллектуальную услугу за счет выноса блоков управления сервисом ближе к пользователю. Однако остается очень много других нерешенных проблем, для решения которых необходимо использовать ИНДПУ.

Отдельно стоит вопрос относительно качества управления и разработки критерия качества. Качество управления интеллектуальными услугами определяется качеством функционирования интеллектуальной надстройки и зависит от ее возможностей обеспечить оптимальные значения критериев качества обслуживания заявок на интеллектуальные услуги.

На сегодняшний день используются векторные и скалярные критерии качества управления услугами. Сравнение систем с помощью векторного критерия не всегда возможно. Для сравнения необходимо, чтобы все частичные критерии, что входят в один векторный критерий, были не больше, чем частичные критерии, которые входят в другой. В противном случае для сравнения необходимо вводить дополнительные условия. Чаще всего сначала формируется векторный критерий, а затем осуществляется переход к скалярному критерию методами, которые предоставляют возможность количественной оценки качества управления услугами.

Для оценки качества управления интеллектуальными услугами ИНЦПУ и ИНДПУ необходимо определить принципы их построения, определить частичные критерии качества управления интеллектуальными услугами и методы их объединения в результирующий критерий.

Целью работы является решение комплекса научно-технических вопросов, связанных с исследованием и разработкой принципов создания интеллектуальной надстройки в NGN для повышения качества управления интеллектуальными услугами.

### Разработка принципов создания интеллектуальной надстройки NGN и результирующего скалярного критерия качества

На основе концепции IPCC [2] предложена архитектура NGN с ИНДПУ (рисунок 1). В сети следующего поколения с ИНДПУ существует несколько узлов коммутации услуги (Softswitch) и несколько узлов управления услугой (серверов). Предлагаются два подхода к реализации ИНДПУ. При первом подходе каждый сервер содержит логику обслуживания всех классов услуг (универсальный сервер). Второй подход предусматривает использование специализированных серверов (сервер содержит логику обслуживания лишь некоторых определенных классов услуг). Предусмотрена возможность предоставления одного класса услуг на нескольких серверах на случай выхода из строя любого сервера. Учитывая меньшую стоимость специализированных серверов и меньшее время обслуживания заявки, в дальнейшем считается, что при построении NGN с ИНДПУ используются специализированные серверы.

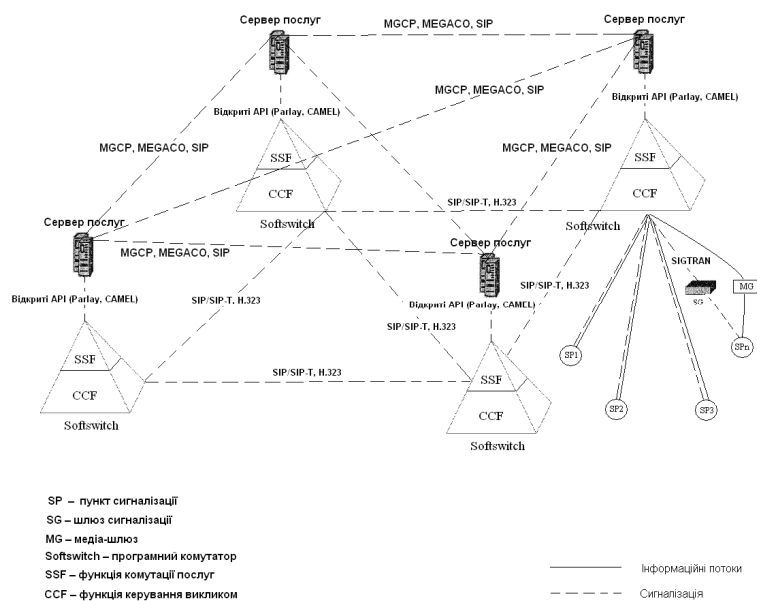


Рис. 1. Архитектура NGN с ИНДПУ

Предусмотрено существование нескольких территориально разнесенных районов. В каждом из них установлен Softswitch, который осуществляет управление районной транспортной сетью и выполняет функцию коммутации услуги SSF. В данном случае рядом с Softswitch расположен сервер, который содержит логику сложных интеллектуальных услуг. Интеллектуальная надстройка – практически отдельная сеть, которая необходима для обмена сообщениями и управляющими сигналами между серверами. В нее входят серверы и части Softswitch, которые выполняют функцию SSF. При обмене



информацией между серверами могут применяться протоколы MGCP, MEGACO, SIP. Связь между Softswitch и сервером осуществляется через открытые API (Parlay, Camel) [4].

Для предоставления преимущества архитектуре NGN с ИНЦПУ или ИНДПУ необходимо сравнить их качество управления интеллектуальными услугами.

Для этого следует определить критерии качества управления. Учитывая проведенные исследования [5], как частичные критерии качества управления предоставлением интеллектуальных услуг предложены: общее время обслуживания заявки на интеллектуальную услугу интеллектуальной надстройкой  $\bar{T}_{ii}$ , вероятность блокировки заявки  $\bar{P}_A$ , количество заявок, которые ожидают обслуживания  $L$  (далее эти критерии – технические), стоимость интеллектуальной надстройки  $\bar{C}$  (далее этот критерий – экономический). Подобные критерии используются для систем управления в IP-сетях и телекоммуникационных сетях. Однако в данном случае они имеют содержание, соответствующее объекту применения, и соответствующий способ расчета.

Как уже отмечалось, сравнить интеллектуальные надстройки с разными принципами управления с помощью частичных критериев, которые являются элементами векторного критерия, не всегда возможно. Целесообразно применить результирующий скалярный критерий. Для его создания использована аддитивная функция полезности:

$$F_p = \sum_{i=1}^I K_i v_i \quad (1)$$

где  $v_i$  – весовые коэффициенты,  $K_i$  –  $i$ -й частичный критерий,  $I$  – количество частичных критериев ( $I = 4$ ).

Для определения значений весовых коэффициентов частичных критериев использованы экспертные оценки приоритета  $\gamma_{ij}$  частичных критериев и сформирована матрица приоритетов [5].

Основой для расчета технических частичных критериев качества являются аналитические модели ИНЦПУ и ИНДПУ.

Для создания аналитических моделей интеллектуальных надстроек можно применить аппарат теории массового обслуживания и теории Марковских случайных процессов. Проанализированы характеристики входного потока заявок и интенсивности обслуживания заявок. При стационарном режиме интеллектуальные надстройки можно описывать не с помощью систем дифференциальных уравнений Колмогорова, а в виде алгебраических систем уравнений.

Предложен алгоритм построения аналитической модели ИНЦПУ, который представляет собой следующие этапы: кодировка состояний системы, формирования размеченного графа переходов системы, формирования системы уравнений и определения стационарных вероятностей, расчет частичных критериев ИНЦПУ. Предложены методы расчета частичных критериев ИНЦПУ.

Рассмотрена следующая аналитическая модель ИНЦПУ.

ИНЦПУ представлена в виде одноканальной системы массового обслуживания (СМО). Входной поток заявок – неоднородный: в систему поступают два класса заявок. Накопитель для заявок – с ограниченной емкостью:  $r = 2$ . Дисциплина обслуживания – с относительными приоритетами: заявки первого класса имеют приоритет по отношению к заявкам второго класса. Заявки двух классов, которые поступают в систему, образуют пуассоновские потоки с интенсивностями  $\lambda_1$  и  $\lambda_2$ , соответственно. Длительность обслуживания заявок каждого класса распределена по экспоненциальному закону с интенсивностями

$\mu_1 = 1/b_1$ ,  $\mu_2 = 1/b_2$ , где  $b_1$  и  $b_2$  – средняя длительность обслуживания заявок класса 1 и 2, соответственно. В СМО всегда существует стационарный режим, поскольку не может быть бесконечных очередей.

Кодирование состояний случайного процесса выполняется следующим образом:  $(\Pi, Ч)$ , где  $\Pi = \{0, 1, 2\}$  – состояние обслуживающего прибора, которое задается классом заявки находящейся на обслуживании («0» – прибор свободен; «1» или «2» – на обслуживании в приборе находится заявка класса 1 или 2, соответственно). Состояние накопителя:  $Ч = \{0, 1, 2, 11, 12, 22\}$ , где «0» – означает отсутствие заявок в накопителе; «1» – присутствие в накопителе только одной заявки класса 1; «2» – присутствие в накопителе только одной заявки класса 2; «11» – присутствие в накопителе двух заявок класса 1; «12» – присутствие в накопителе одной заявки класса 1 и одной заявки класса 2; «22» – присутствие в накопителе двух заявок класса 2. Закодировав состояния, построив граф переходов (рисунок 2), составив систему уравнений баланса для каждого из состояний системы и решив ее, можно определить вероятности нахождения системы в каждом из состояний  $p_i$ .

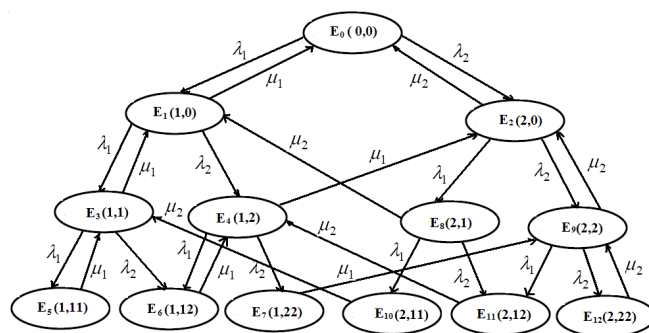


Рис. 2. Размеченный граф переходов Марковского процесса для ИНЦПУ

Характеристики предложенной ИНЦПУ при найденных значениях стационарных вероятностей состояний случайного процесса рассчитываются при помощи следующих выражений:

Среднее число заявок в очереди:

$$\bar{L}_{\text{оч}} = p_3 + p_4 + 2p_5 + 2p_6 + 2p_7 + p_8 + p_9 + 2p_{10} + 2p_{11} + 2p_{12} \quad (2)$$

Среднее число заявок в системе:

$$\bar{M}_{\text{сист}} = p_1 + p_2 + 2p_3 + 2p_4 + 3p_5 + 3p_6 + 3p_7 + 2p_8 + 2p_9 + 3p_{10} + 3p_{11} + 3p_{12} \quad (3)$$

Вероятность потери заявок:

$$\bar{P}_A = p_5 + p_6 + p_7 + p_{10} + p_{11} + p_{12} \quad (4)$$

Производительность системы:

$$\lambda' = (\lambda_1 + \lambda_2)(1 - \bar{P}_A) \quad (5)$$

Среднее общее время обслуживания заявки на интеллектуальную услугу:

$$\bar{T}_{\text{общ}} = \bar{M}_{\text{сист}} / \lambda' \quad (6)$$

Стоит отметить, что при увеличении количества классов услуг и длины очереди при сервере, сложность аналитической модели существенно возрастает, что усложняет ее практическое применение.

Представить ИНДПУ с помощью обычной системы массового обслуживания уже нельзя. ИНДПУ можно рассматривать как несколько соединенных между собой систем массового обслуживания, то есть как сеть массового обслуживания (СМО).

Предложен алгоритм построения аналитической модели ИНДПУ, которая представляет собой следующие этапы: кодировка состояний системы, формирования размеченного графа переходов системы, формирования системы уравнений и определения стационарных вероятностей, расчет частичных критериев ИНДПУ. Предложены методы расчета частичных критериев ИНДПУ. Рассмотрена следующая аналитическая модель ИНДПУ.

Представлено ИНДПУ в виде разомкнутой экспоненциальной СМО с двумя одноканальными узлами, в которую из внешней среды поступают заявки с интенсивностями  $\lambda_{ij}$ , где  $i = \overline{1, 2}$  – номера серверов,  $j = \overline{1, 2}$  – классы заявок. Заявки первого класса имеют более высокий приоритет, чем заявки второго. Накопители в обоих узлах имеют ограниченную емкость  $r_i$ ; зададим:  $r_1 = r_2 = 1$ . Заявка, которая поступила в узел и застала накопитель заполненным, теряется. Длительность обслуживания  $b_{ij}$  в узлах распределена по экспоненциальному закону со средними значениями  $b_{11}, b_{12}, b_{21}, b_{22}$ , соответственно. Считается, что применяются два специализированных сервера. Первый способен обслуживать лишь заявки класса 1, а заявки класса 2 отправляет на второй сервер. Второй сервер способен обслуживать лишь заявки класса 2, а заявки класса 1 перенаправляет к первому серверу. В таком случае значения  $b_{12}, b_{21}$  будут отвечать не среднему времени обслуживания, а среднему времени дешифрации класса заявки и перенаправления в другой сервер.

Поскольку заявки в сети могут теряться, представленная разомкнутая СМО является нелинейной, то есть интенсивности потоков заявок, которые поступают в узлы СМО, не связаны между собой линейной зависимостью и не могут быть рассчитаны путем решения системы линейных алгебраических уравнений. Длительность обслуживания заявок в узлах СМО распределена по экспоненциальному закону с интенсивностями обслуживания:  $\mu_{11} = 1/b_{11}, \mu_{12} = 1/b_{12}, \mu_{21} = 1/b_{21}, \mu_{22} = 1/b_{22}$ . В разомкнутой СМО при любой нагрузке существует стационарный режим, поскольку в узлах сети не могут быть бесконечные очереди.

Для описания состояний Марковского случайного процесса использовано распределение заявок между узлами. Состояния закодированы таким образом:  $(M_1/M_2)$  (где  $M_i = (П, Ч)$ ; П – прибор,  $П = \{0, 1, 2\}$ ; Ч – очередь,  $Ч = \{0, 1, 2\}$ ). Расшифруем обозначение:

1. Для прибора П: «0» – узел свободный; «1» – на обслуживании в узле находится заявка класса 1; «2» – на обслуживании в узле находится заявка класса 2.
2. Для очереди Ч: «0» – нет заявок в очереди; «1» – в очереди находится заявка класса 1; «2» – в очереди находится заявка класса 2.

Построив граф переходов аналогично ИНДПУ, составив систему уравнений баланса для каждого из состояний системы и решив ее, можно определить вероятности нахождения системы в каждом из отмеченных состояний  $p_i$ .

После этого можно рассчитать все основные характеристики системы.

Среднее число заявок в очереди первого сервера:

$$\bar{L}_1 = p_2 + p_3 + p_5 + p_6 + \sum_{k=19}^{30} p_k + \sum_{l=37}^{48} p_l \quad (7)$$

Среднее число заявок в очереди второго сервера:

$$\begin{aligned} \bar{L}_2 = & p_8 + p_9 + p_{11} + p_{12} + p_{14} + p_{15} + p_{17} + p_{18} + p_{20} + p_{21} + p_{23} + \\ & p_{24} + p_{26} + p_{27} + p_{29} + p_{30} + p_{32} + p_{33} + p_{35} + p_{36} + p_{38} + p_{39} + p_{41} + \\ & p_{42} + p_{44} + p_{45} + p_{47} + p_{48} \end{aligned} \quad (8)$$

Тогда общая длина очереди:

$$\bar{L}_{i \text{ И } o} = \bar{L}_1 + \bar{L}_2 \quad (9)$$

Загрузки серверов соответствующим классом заявок определяются как суммы вероятностей состояний, в которых соответствующий сервер занят обслуживанием заявок соответствующего класса:

$$\begin{aligned} \rho_{11} = & p_1 + p_2 + p_3 + \sum_{k=13}^{30} p_k; \quad \rho_{12} = p_4 + p_5 + p_6 + \sum_{l=31}^{48} p_l \\ \rho_{21} = & p_7 + p_8 + p_9 + p_{13} + p_{14} + p_{15} + p_{19} + p_{20} + p_{21} + p_{25} + p_{26} + p_{27} + \\ & p_{31} + p_{32} + p_{33} + p_{37} + p_{38} + p_{39} + p_{43} + p_{44} + p_{45} \\ \rho_{22} = & p_{10} + p_{11} + p_{12} + p_{16} + p_{17} + p_{18} + p_{22} + p_{23} + p_{24} + p_{28} + p_{29} + p_{30} + \\ & p_{34} + p_{35} + p_{36} + p_{40} + p_{41} + p_{42} + p_{46} + p_{47} + p_{48} \end{aligned} \quad (10)$$

Производительности узлов (интенсивности обслуженных заявок соответствующего класса на выходе узлов) определяются по формулам:

$$\lambda'_{11} = \rho_{11}\mu_{11}; \quad \lambda'_{12} = \rho_{12}\mu_{12}; \quad \lambda'_{21} = \rho_{21}\mu_{21}; \quad \lambda'_{22} = \rho_{22}\mu_{22} \quad (11)$$

Вероятность потери заявок соответствующего класса на соответствующем сервере:

$$\bar{P}_{A11} = 1 - \frac{\lambda'_{11}}{\lambda_{11} + \lambda_{21}}; \quad \bar{P}_{A12} = 1 - \frac{\lambda'_{12}}{\lambda_{12}}; \quad \bar{P}_{A22} = 1 - \frac{\lambda'_{22}}{\lambda_{22} + \lambda_{12}}; \quad \bar{P}_{A21} = 1 - \frac{\lambda'_{21}}{\lambda_{21}} \quad (12)$$

Вероятность потери заявок соответствующего класса в ИНДПУ:

$$\bar{P}_{A1} = 1 - \frac{\lambda'_{11}}{\lambda_{11} + \lambda_{21}}; \quad \bar{P}_{A2} = 1 - \frac{\lambda'_{22}}{\lambda_{22} + \lambda_{12}} \quad (13)$$

Среднее число заявок соответствующего класса:

$$\begin{aligned} \bar{M}_{11} = & p_1 + 2p_2 + p_3 + p_5 + \sum_{k=13}^{18} p_k + 2\sum_{l=19}^{24} p_l + \sum_{m=25}^{30} p_m + \sum_{n=37}^{42} p_n; \quad \bar{M}_{12} = p_3 + p_4 + p_5 + 2p_6 + \sum_{k=25}^{42} p_k + 2\sum_{l=43}^{48} p_l \\ \bar{M}_{21} = & p_7 + 2p_8 + p_9 + p_{11} + p_{13} + 2p_{14} + p_{15} + p_{17} + p_{19} + 2p_{20} + p_{21} + p_{23} + p_{25} + 2p_{26} + p_{27} + \\ & p_{29} + p_{31} + 2p_{32} + p_{33} + p_{35} + p_{37} + 2p_{38} + p_{39} + p_{41} + p_{43} + 2p_{44} + p_{45} + p_{47} \\ \bar{M}_{22} = & p_{10} + p_{11} + 2p_{12} + p_{15} + p_{16} + p_{17} + 2p_{18} + p_{21} + p_{22} + p_{23} + 2p_{24} + \end{aligned} \quad (14)$$

$$2p_{30} + p_{33} + p_{34} + p_{35} + 2p_{36} + p_{39} + p_{40} + p_{41} + 2p_{42} + p_{45} + p_{46} + p_{47} + 2p_{48}$$

Определена весомость потока заявок соответствующего класса для каждого из серверов:

$$f_1 = \lambda_{11}(1 - P_{A11}); f_2 = \lambda'_{21}(1 - P_{A11}); f_3 = \lambda_{22}(1 - P_{A22}); f_4 = \lambda'_{12}(1 - P_{A22})$$

$$\varepsilon_{11} = \frac{f_1}{f_1 + f_2}; \varepsilon_{12} = \frac{f_2}{f_1 + f_2}; \varepsilon_{21} = \frac{f_3}{f_3 + f_4}; \varepsilon_{22} = \frac{f_4}{f_3 + f_4} \quad (15)$$

Время обслуживания заявки соответствующего класса на каждом из серверов:

$$\bar{O}_{11} = \frac{\bar{M}_{11}}{\lambda'_{11}}; \bar{O}_{12} = \frac{\bar{M}_{12}}{\lambda'_{12}}; \bar{O}_{21} = \frac{\bar{M}_{21}}{\lambda'_{21}}; \bar{O}_{22} = \frac{\bar{M}_{22}}{\lambda'_{22}} \quad (16)$$

Тогда общее время обслуживания заявки соответствующего класса в ИНДПУ:

$$\bar{O}_{i \in \Lambda} = \varepsilon_{i1} \bar{O}_{11} + \varepsilon_{i2} (\bar{O}_{21} + \bar{O}_{11}); \bar{O}_{i \in \Lambda} = \varepsilon_{i2} \bar{O}_{22} + \varepsilon_{i1} (\bar{O}_{12} + \bar{O}_{22}) \quad (17)$$

Следует отметить, что при росте количества классов услуг, количества серверов, длины очереди при серверах, учитывая нелинейность систем, с помощью которых моделируются сервера, и объединения их в сеть, сложность аналитической модели существенно растет, что усложняет ее практическое использование.

Одним из частичных критериев качества ИНЦПУ и ИНДПУ является их стоимость, точнее говоря, чистая приведенная стоимость NPV. Формулы расчета NPV для ИНЦПУ и ИНДПУ будут иметь одинаковый вид. Для расчета введены следующие обозначения:  $P$  – объем доходов от внедрения интеллектуальной надстройки;  $B_u$  – инвестиционный капитал;  $B_e$  – текущие расходы на интеллектуальную надстройку;  $A$  – годовая амортизация основных средств, в которые инвестирован капитал;  $T$  – период реализации и эксплуатации интеллектуальной надстройки (количество лет);  $t$  – порядковый номер каждого года;  $r$  – годовая дисконтная ставка.

NPV можно определить, как общую разницу между приведенными доходами и расходами на интеллектуальную надстройку за все годы реализации и эксплуатации:

$$NPV = \sum_{t=1}^T \frac{(P_t - B_{ut} - P_{et})}{(1+r)^{t-1}} \quad (18)$$

Как уже отмечалось, при исследовании интеллектуальных надстроек, которые содержат больше двух специализированных серверов и обслуживают несколько классов заявок на интеллектуальные услуги, практическое использование аналитических моделей не целесообразно через их сложность. Предлагается в таких случаях использовать имитационные модели.

### Разработка имитационных моделей ИНЦПУ и ИНДПУ и определение сферы их эффективного применения

На основании разработанных аналитических моделей для исследования ИНЦПУ и ИНДПУ предложены их имитационные модели. Представлено алгоритм функционирования имитационной модели ИНЦПУ. Для достижения определенной точности результатов в модель должно поступить соответствующее количество заявок. Происходит генерирование заявок двух классов с разными приоритетами, которые поступают в систему. Отдельные блоки моделируют пребывание заявки в очереди и обслуживания заявок сервером, рассчитывают частичные критерии.

Алгоритм функционирования имитационной модели ИНДПУ представлен на рис. 3. Для достижения определенной точности результатов в модель должно поступить соответствующее количество заявок. Происходит генерирование заявок двух классов с разными приоритетами, которые на следующем шаге поступают в два сервера. Два сервера и очереди моделируются отдельными блоками, которые выполняют всю необходимую логику обслуживания и рассчитывают частичные критерии. Предусмотрена передача заявки класса 2, что поступила в первый сервер, на второй сервер. Подобная процедура происходит и с заявками класса 1, что поступили во второй сервер [6].

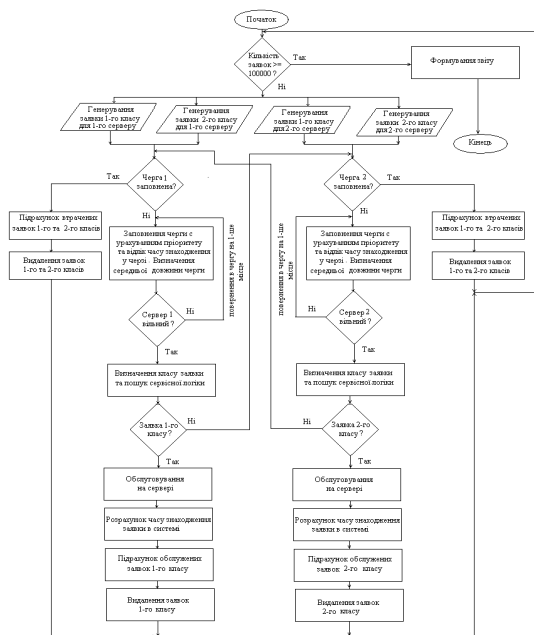


Рис.3. Алгоритм функционирования имитационной модели ИНДПУ

Для моделирования описанной системы на ЭВМ использовано GPSS (General Purpose Simulation System). В результате тестирования отмеченных моделей было установлено, что погрешность результатов, полученных с помощью аналитической и имитационной моделей, для ИНЦПУ не превышает 0,85%, для ИНДПУ – 4%.

Большее значение погрешности при сравнении имитационной и аналитической моделей ИНДПУ в первую очередь связано со сложностью ее архитектуры и, соответственно, ее аналитической и имитационной моделей. Полученные результаты подтверждают возможность использования имитационных моделей интеллектуальных надстроек в тех случаях, когда создание их аналитических моделей практически невозможно.

Допускалось, что серверы имеют постоянное время обслуживания заявки соответствующего класса. Для сравнения качества функционирования интеллектуальных надстроек использованы следующие значения. Для серверов ИНДПУ:  $\mu_{11} = 0.4$ ,  $\mu_{12} = 0.5$ ,  $\mu_{21} = 0.5$ ,  $\mu_{22} = 0.4$ . Для сервера ИНЦПУ:  $\mu_1 = 0.4$ ,  $\mu_2 = 0.4$ .

Изменяя значение  $\lambda$ , можно сравнить качество управления обслуживанием заявок. Считалось, что,  $\lambda_{11} = \lambda_{12} = \lambda_{21} = \lambda_{22}$  а,  $\lambda_1 = \lambda_{11} + \lambda_{21}$ ,  $\lambda_2 = \lambda_{12} + \lambda_{22}$ . В расчетах принята стоимость ИНДПУ 10 000 у.е., а стоимость ИНЦПУ – 7 000 у.е. Коэффициент нормирования для стоимости составляет 1000, для других показателей – 1. Выполнив соответствующие расчеты, получены результаты, представленные в таблице 1.

Таблица 1. Результаты расчетов

ИН	$\lambda$	L	$\bar{T}_{oo1}$	$\bar{T}_{oo2}$	$\bar{P}_{Б1}$	$\bar{P}_{Б2}$	C	F <sub>p</sub>
ДПУ	0,05	0,159	4,375	4,373	0,115	0,117	10 000	3,22
ЦПУ	0,1	0,274	3,785	4,208	0,069	0,067	7 000	2,74
ДПУ	0,08	0,313	4,657	4,662	0,222	0,218	10 000	3,49
ЦПУ	0,16	0,563	4,215	5,073	0,174	0,17	7 000	3,21
ДПУ	0,1	0,402	4,753	4,743	0,283	0,283	10 000	3,62
ЦПУ	0,2	0,76	4,522	5,617	0,253	0,25	7 000	3,53
ДПУ	0,2	0,793	5,081	5,094	0,518	0,517	10 000	4,08
ЦПУ	0,4	1,334	5,01	7,169	0,534	0,53	7 000	4,42
ДПУ	0,4	1,209	5,19	5,21	0,724	0,730	10 000	4,42
ЦПУ	0,8	1,7	5,275	8,516	0,754	0,758	7 000	5,10
ДПУ	0,6	1,406	5,12	5,16	0,8	0,8	10 000	4,49
ЦПУ	1,2	1,803	5,124	8,796	0,833	0,832	7 000	5,23
ДПУ	0,8	1,536	5,092	5,149	0,858	0,856	10 000	4,56
ЦПУ	1,6	1,86	5,136	9,026	0,875	0,872	7 000	5,34

Как видно с таблицы 1 при малых значениях интенсивности поступления заявки  $\lambda$ , когда  $\lambda < \mu$ , лучше применять ИНЦПУ. При росте  $\lambda$  и приближении ее к  $\mu$  преимущество ИНЦПУ над ИНДПУ становится значительно меньше и при определенных значениях уже лучше применять ИНДПУ. Особенно существенно преимущество ИНДПУ ощутимое при  $\lambda \geq \mu$ . Однако такие результаты касаются лишь интеллектуальных надстроек с простой структурой (обслуживают незначительное количество классов заявок, ИНДПУ имеет два сервера).

Проанализировано поведение более сложных интеллектуальных надстроек.

Рассмотрено ИНЦПУ, обслуживающую четыре класса заявок.

ИНЦПУ – одноканальная СМО. Входной поток заявок – *неоднороден*: в систему поступает 4 класса заявок. Накопитель для заявок – *ограниченной емкостью*  $r = 16$ . Дисциплина буферизации – *без вытеснения* заявок: если при поступлении в систему заявки любого класса накопитель заполнен до конца, то заявка теряется. Дисциплина обслуживания – *с относительными приоритетами*: чем меньше значение класса, тем выше приоритет заявок. Всякий раз из накопителя на обслуживание выбирается заявка с наивысшим приоритетом. При этом во время поступления в систему высокоприоритетной заявки обслуживания низкоприоритетной не прерывается. Предположения и допущения: заявки классов, которые поступают в систему, образуют *простые* потоки с интенсивностями  $\lambda_j$ , где  $j = \overline{1, 4}$ . Длительность обслуживания заявок каждого класса распределена по экспоненциальному закону с интенсивностями  $\mu_j = 1/b_j$ , где  $b_j$  – средняя длительность обслуживания заявок  $j$ -го класса.

Проанализировано поведение ИНДПУ с четырьмя серверами, которая обслуживает четыре класса заявок. ИНДПУ представлено в виде следующей СЕМО: ИНДПУ – разомкнута экспоненциальная СЕМО с 4 узлами, которые отвечают серверам ИНДПУ. Узлы СЕМО – одноканальные. Накопители в узлах имеют

ограниченную емкость  $r_i$ . Установим  $r_i = 4$ , где  $i = \overline{1,4}$ . Входной поток заявок – *неоднороден*: в систему поступает 4 класса заявок с интенсивностями  $\lambda_{ij}$ , где  $i = \overline{1,4}$  – номер серверов,  $j = 4$  – класс заявок.

Заявки класса 4 не поступают к четвертому серверу, а перенаправляются еще на стадии коммутации программным коммутатором. Дисциплины буферизации в узлах – с потерями заявок, если накопители заполнены. Дисциплина обслуживания – с *относительными приоритетами*: чем меньше значение класса, тем высший приоритет заявок. Всякий раз из накопителя на обслуживание выбирается заявка с наивысшим приоритетом. При этом во время поступления в систему высокоприоритетной заявки обслуживания низкоприоритетной не прерывается. Заданы матрицы  $Q^j = \|q_{ik}^j\|$  вероятностей передач заявки из текущего сервера на другие серверы или обслуживание текущим сервером, где  $i, k = \overline{1,4}$  – номер серверов, для классов заявок  $j = 4$ . При  $i \neq k$   $q^j$  отвечает вероятности передачи, при  $i = k$   $q^j$  отвечает вероятности обслуживания текущим сервером.

Предположение и допущение:

Длительность обслуживания заявок в узлах СЕМО распределена по экспоненциальному закону с интенсивностями обслуживания:  $\mu_{ij} = 1/b_{ij}$ , где  $i = \overline{1,4}$  – номер сервера,  $j = 4$  – класс заявок.  $b_{ij}$  – средняя длительность обслуживания заявок  $j$ -го класса на  $i$ -м сервере. В таком случае  $b_{13}, b_{14}, b_{23}, b_{24}, b_{31}, b_{32}, b_{41}, b_{42}$  будут соответствовать не среднему времени обслуживания, а среднему времени дешифрования класса заявки и перенаправления в другой сервер.

Для описанных интеллектуальных надстроек разработка аналитических моделей связана с большими сложностями, потому для исследования таких систем целесообразно прибегнуть к имитационному моделированию.

Алгоритм функционирования имитационной модели рассмотренной ИНЦПУ является подобным алгоритму имитационной модели ИНДПУ, которая обслуживает два класса заявок. Отличие заключается лишь в количественных характеристиках.

Представлен алгоритм функционирования имитационной модели ИНДПУ. Модель ИНДПУ должна работать в течение определенного времени или к ней должно поступить соответствующее количество заявок. Происходит генерирование заявок четырех классов с разными приоритетами. Четыре сервера и очереди моделируются отдельными блоками, которые выполняют всю необходимую логику обслуживания и рассчитывают характеристики. Если к первому и второму серверам поступили заявки классов 1 и 2, то они обслуживаются. Если же поступили заявки классов 3 и 4, то они передаются на другие сервера, способные их обслужить, в соответствии с  $Q^j = \|q_{ik}^j\|$ . Подобная процедура происходит и с заявками классов 1 и 2, которые поступили в третий и четвертый сервер.

Проведено сравнение результатов, полученных при использовании предложенных моделей. Считалось, что сервера имеют постоянное среднее время обслуживания заявки соответствующего класса, а время обслуживания и время дешифрации для всех заявок на всех серверах одинаково. Для сравнения качества функционирования интеллектуальных надстроек использовались следующие значения интенсивностей обслуживания и дешифрации: серверы ИНДПУ –  $\mu_{i\dot{a}\dot{n}} = 0.4$ ,  $\mu_{\dot{a}\dot{a}\dot{o}} = 0.4$ , сервер ИНЦПУ –  $\mu = 0.4$ .

Изменяя значение интенсивностей поступления заявок в систему, можно сравнить качество управления обслуживанием заявок. Считалось, что все  $\lambda_{ij}$  для ИНДПУ одинаковые, где  $i = \overline{1,4}$  – номер серверов,



$j = \overline{1,4}$  – класс заявок, а  $\lambda_j$  для ИНЦПУ  $\lambda_j = \sum_{i=1}^4 \lambda_{ij}$  при ИНДПУ. Кроме того  $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$ . В

расчетах принята стоимость ИНДПУ 25 000 у.е., а стоимость ИНЦПУ – 7 000 у.е. Коэффициент нормирования для стоимости составляет 1000, для других показателей – 1. В результате использования моделей и выполнения соответствующих расчетов для входных параметров получены значения, представленные в таблице 2.

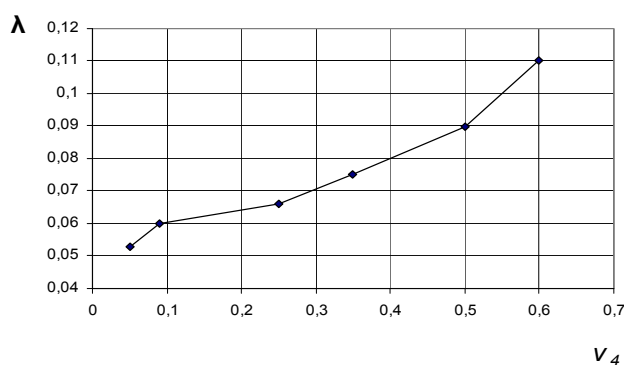
Таблица 2. Результаты расчетов

ИН	$\lambda$	$L$	$\overline{T}_{ii1}$	$\overline{T}_{ii2}$	$\overline{T}_{ii3}$	$\overline{T}_{ii4}$	$\overline{P}_{A1}$	$\overline{P}_{A2}$	$\overline{P}_{A3}$	$\overline{P}_{A4}$	$C$	$F_p$
ИНДПУ	0,0062	0,138	3,833	3,867	3,864	4,386	0	0	0	0	25000	6,26
ИНЦПУ	0,025	0,084	3,179	3,287	3,371	3,556	0	0	0	0	7000	3,99
ИНДПУ	0,007	0,05	3,874	3,89	3,891	4,464	0	0	0	0	25000	6,29
ИНЦПУ	0,028	0,108	3,251	3,402	3,52	3,7	0	0	0	0	7000	4,12
ИНДПУ	0,01	0,106	4	4,07	4,09	4,79	0	0	0	0	25000	6,66
ИНЦПУ	0,04	0,262	3,644	3,869	4,251	4,859	0	0	0	0	7000	4,83
ИНДПУ	0,015	0,411	4,299	4,394	4,469	5,484	0	0	0	0	25000	6,98
ИНЦПУ	0,06	0,907	4,286	5,061	6,512	9,392	0	0	0	0	7000	7,10
ИНДПУ	0,025	1	4,908	5,18	5,325	7,475	0,0093	0,0069	0,0097	0,01753	25000	8,15
ИНЦПУ	0,1	7,562	5,514	7,899	14,629	63,533	0,0585	0,0555	0,0577	0,05899	7000	24,85

Как уже отмечалось, чем значение критерия меньше, тем лучше качество функционирования интеллектуальной надстройки. Как видно из таблицы 2, даже при незначительных интенсивностях поступления заявок  $\lambda = 0,06$  качество управления интеллектуальными услугами ИНДПУ лучше, чем ИНЦПУ заданной сложности.

Проанализировав таблицы 1 и 2, можно сделать вывод, что при усложнении структуры интеллектуальной надстройки (росте количества классов интеллектуальных услуг, количества серверов при ИНДПУ, разных вариантах размещения логики обслуживания заявок на услугу и тому подобное) значения интенсивности поступления заявок на услугу, при которой ИНДПУ функционирует качественнее чем ИНЦПУ, уменьшается и выбор архитектуры зависит от ряда факторов. Причем одним из важнейших является стоимость интеллектуальной надстройки. Однако при совсем малых значениях интенсивности поступления заявок на услугу преимущество ИНДПУ сводится на нет в результате учета времени передачи заявок от сервера к серверу в случае отсутствия необходимой логики обслуживания для заявки.

Следует отметить, что достаточно существенно на результат сравнения влияют весовые коэффициенты в обобщенном скалярном критерии качества. В таком случае при ограниченных инвестиционных возможностях весовой коэффициент при частичном критерии  $\overline{C}$  будет достаточно большим. В результате расчетов при разных значениях весового коэффициента стоимости интеллектуальной надстройки получен график зависимости интенсивности поступления заявок в систему, при которой ИНДПУ функционирует уже качественнее за ИНЦПУ, от значения весового коэффициента при стоимости (рисунок 4).



**Рис.4.** График зависимости интенсивности поступления заявок в систему, при которой ИНДПУ функционирует уже качественнее за ИНЦПУ, от значения весового коэффициента при стоимости

## Литература

1. Гольдштейн А.Б. Подводная часть айсберга по имени NGN / А.Б. Гольдштейн, Н.А. Соколов // Технологии и средства связи. – 2006. – № 2. – С. 12-21.
2. Гольдштейн А.Б. Построение NGN: IPCC vs. TISPAN / А.Б. Гольдштейн, А. Атцик // Connect! Мир связи. – 2006. – № 4. – С. 90-95.
3. 3.Управління телекомунікаціями із застосуванням новітніх технологій/ [В.Г. Кривуца, В.К. Стеклов, Л.Н. Беркман, Б.Я. Костік, В.Ф. Олійник]. – Техніка, 2007. – 384 с.
4. Шестопапов С.В. Підвищення якості управління послугами при застосуванні децентралізованої системи управління / С.В. Шестопапов, Н.о. Князева // Вісник ДУІКТ. – т.8. – №1 – К.: ДУІКТ, 2010р. – С. 21-28.
5. Шестопапов С.В. Результуючий критерій якості системи управління додатковими послугами в NGN / С.В. Шестопапов // Науковий Вісник Донбаської державної машинобудівної академії. – №2 (5E). – Краматорск: Издательство ДГМА, 2009р. – С. 185-189.
6. Шестопапов С.В. Моделювання процесів функціонування систем управління / С.В. Шестопапов, Н.о. Князева // Холодильная техника и технология. – №2 (124) – Одеса: Видавництво ОДАХ. – 2010. – С. 55-62.

## Благодарности

*Настоящая работа была выполнена при поддержке интернационального проекта ITHEA XXI Института информационных теорий и их приложений FOI ITHEA и Ассоциации ADUIS Украина (Ассоциация разработчиков и пользователей интеллектуальных систем), в связи с чем автор и выражает свою искреннюю благодарность.*

## Информация об авторах

**Сергей Шестопапов** – Факультет информационных технологий и кибербезопасности ОНАПТ, ассистент кафедры информационных систем и сетей ; ул. Дворянская, 1/3, Одесса-26, 65026, Украина; моб. 097-625-96-77; e-mail: shestopalov\_s@mail.ru

*Главные области научного исследования: Качество управления интеллектуальными услугами.*

---

## ПОВЫШЕНИЕ СТРУКТУРНОЙ ЖИВУЧЕСТИ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ

Нина Князева

**Аннотация:** Представлен метод повышения структурной живучести телекоммуникационной сети путем реконфигурации сети, выполняемой на основе оценки относительного изменения показателей структурной живучести – величин средневзвешенного максимального потока и средневзвешенного кратчайшего пути в сети, служащих для формирования средневзвешенной комплексной оценки относительного изменения показателей структурной живучести. Задача реконфигурации сети сформулирована как задача линейного программирования.

**Ключевые слова:** структурная живучесть, сеть, реконфигурация, максимальный поток, кратчайший путь, комплексная оценка.

**Ключевые слова классификации ACM:** C.2. Computer-communication networks, H. Information Systems - H.1 Models and Principles, K. Computing Milieux - K.6 Management of computing and information system

---

### Введение

Эффективность функционирования телекоммуникационных сетей (ТКС) зависит от многих свойств ТКС, среди которых одним из важнейших является живучесть сети. Обеспечение живучести ТКС становится все более актуальным в связи с интенсивным развитием телекоммуникаций в последнее десятилетие.

Применительно к Future Networks – будущим сетям (БС) – в документах ИТУ [1] рекомендуется, чтобы проектирование, эксплуатация и развитие БС осуществлялись таким образом, чтобы обеспечить надежность, безопасность, а также живучесть, определяющую возможность системы своевременно выполнять свою миссию – обеспечивать приемлемый уровень обслуживания, даже если нормальная эксплуатация сети осложняется различными проблемами. Под живучестью ТКС понимают способность сохранять и восстанавливать выполнение основных функций в заданном объеме и на протяжении заданного времени в случае изменения структуры ТКС и условий ее функционирования вследствие некоторых неблагоприятных внешних воздействий (НВВ) [2, 3]. Живучесть ТКС анализируют и оценивают на разных уровнях проектирования, моделирования и функционирования. Проблеме оценки живучести ТКС посвящен ряд работ (Белоцерковский Д.Л., Вишневский В.М., Громов Ю.Ю., Додонов А.Г., Кузнецова М. Г., Мельников Ю.Е., Стекольников Ю.И., Малашенко Ю.Е., С. J. Colbourn, K.Sekine, H.Imai, S.Tarn, и др.). В настоящее время существуют различные аналитические модели и методы оценки и повышения живучести систем, основанные на разнообразных подходах. Так, в работе [2] представлены методы оценки функциональной и структурной живучести, основанные на использовании теоретико-игрового, энтропийного подходов, а также логико-вероятностных моделей. В работе [3] рассматриваются детерминированные и стохастические модели невозстанавливаемых и восстанавливаемых систем, которые позволяют исследовать процессы изменения свойства живучести во времени и условиях, которые непрерывно или дискретно изменяются; приведены методы оценки живучести разных систем – ассоциативных, структурных, ассоциативно-структурных. В работе [4] представлены методы вычисления

общей живучести сетевой информационной системы в полиномиальной форме, а также методы, которые основаны на использовании искусственного интеллекта; предложены модели и алгоритмы для решения задач анализа и оценки живучести сетей. В работе [5] предложено использование минимаксных критериев и экспертного выбора структуры системы, устойчивой к НВВ. Отметим, что рассмотренные общие проблемы оценки и повышения живучести систем различного назначения относятся и к ТКС.

### Постановка задачи исследования

Современные тенденции развития телекоммуникаций, переход к сетям нового поколения – NGN (Next Generation Network – сети следующего поколения), FGN (Future Generation Network – сети будущего поколения (БС)) – обуславливают важность вопросов оценки и повышения живучести ТКС с учетом тенденций их развития. При этом особое значение приобретают вопросы обеспечения структурной живучести ТКС, решение которых возможно на основе реконструкции, реорганизации или реконфигурации сети, что позволит создать структуру, обеспечивающую выполнение критического подмножества функций для достижения цели функционирования.

Целью данной работы является разработка метода оценки и повышения структурной живучести ТКС путем реконфигурации сети, выполняемой на основе оценки относительного изменения показателей структурной живучести.

### Показатели структурной живучести

В математической теории графов показатели структурной живучести интерпретируются как количественные меры связности для структуры графа: узловая связность, обобщенная связность, минимальное сечение, максимальный поток, длина пути и т.д., то есть задачи оценки структурной живучести сети можно свести к задачам анализа связности графов, оценки величины кратчайших путей и максимальных потоков, оценки вероятности формирования работоспособной структуры в случае НВВ и т.д. [2]. При оценке структурной живучести возникает необходимость в определении кратчайших путей и максимальных двухполюсных потоков в сети, а также в определении их изменения при влиянии НВВ. Кроме того, существует необходимость в определении количества путей, которые остаются в сети при НВВ, которое выражается в удалении некоторой линии связи (ветви) из сети, а также в определении вероятности безотказного функционирования связи произвольной пары пунктов ( $i-j$ ) и структурной надежности всей сети.

В работе [6] приведена оценка вероятности исключения некоторого заданного пути при удалении некоторой произвольной ветви. Получено выражение для количества путей ранга  $r$ , которые остаются после удаления из полносвязной сети  $l$  ветвей (в сети остается  $L = L_{\max} - l$  ветвей):

$$M_{r,L} = \frac{n(n-1)}{2} C_{n-2}^{r-1} \left( 1 - \frac{2m_{r,L_{\max}}}{n(n-1)C_{n-2}^{r-1}} \right)^l. \quad (1)$$

Здесь  $m_{r,L_{\max}}$  – количество путей ранга  $r$ , которые приходятся на одну связь ( $i-j$ ) в сети с  $n$  пунктами и  $L_{\max}$  ветвями ( $L_{\max}$  – максимально возможное количество ветвей сети,  $i, j = \overline{1, n}$ ,  $i \neq j$ ,  $n$  – количество пунктов сети. Ранг  $r$  – количество ветвей, составляющих путь).

Количество путей ранга  $r$ , которые приходятся на одну связь ( $i-j$ ) в сети с  $n$  пунктами и  $L$  ветвями, определяется по формуле:

$$m_{(ij)r,L} \frac{M_{r,L}}{g} = \frac{M_{r,L}}{n(n-1)}, \quad (2)$$

где  $g$  — общее количество тяготеющих пар.

Вероятность безотказного функционирования связи произвольной тяготеющей пары  $(i-j)$ , которая характеризует структурную надежность всей сети, рассчитывается по формуле:

$$P_{(ij)} = 1 - \prod_{r=1}^R (1 - p_{xy}^r)^{m_{(ij)r,L}}, \quad (3)$$

где  $p_{xy}$  — вероятность безотказной работы ветви между произвольными пунктами  $x$  и  $y$ . Использование выражений (1) и (2) позволяет определить основные характеристики живучести ТКС –  $K$ -связность графа (сети), диаметр графа (сети) и др.

В табл. 1 приведены результаты расчетов числа путей в сети со 100 пунктами любой случайной (неопределенной) структуры, имеющей заданные структурные характеристики – число пунктов и число ветвей. Рис. 1 представляет графики зависимости числа путей  $M$  от числа ветвей  $L$  для сети неопределенной структуры размерности  $n = 100$ . Здесь  $M_a$  – пути, полученные на основе аналитических выражений (1, 2),  $M_{э+и}$  – пути, полученные на основе результатов расчетов программы, моделирующей структурные характеристики сети.

На основе выражений (1) и (2) могут получены графики для сетей заданной размерности с любым заданным числом ветвей, аналогичные графику, приведенному на рис.1, что позволяет выполнить расчет безотказного функционирования связи  $(i-j)$  в сетях с заданными структурными характеристиками ( $i, j = \overline{1, n}, i \neq j, n$  – количество пунктов сети).

Таблица 1 – Число путей в сети неопределенной структуры со 100 пунктами

Число ветвей $L$	Число путей (аналитика) $M_a$	Число путей (эмпирика, интерполяция) $M_{э+и}$	Логарифмическое отклонение $M_a$ от $M_{э+и}$ , $\delta_{lg}$ , %
130	4,82E+12	1,86E+11	12,5%
160	9,48E+17	9,04E+17	0,1%
190	1,46E+23	1,11E+24	-3,7%
220	1,19E+28	4,18E+29	-5,2%
...			
625	3,38E+69	5,14E+75	-8,2%
...			
1000	3,04E+89	4,65E+96	-7,4%
2000	1,48E+119	1,42E+124	-4,0%
3000	3,82E+136	6,45E+138	-1,6%
4000	8,80E+148	2,99E+149	-0,36%

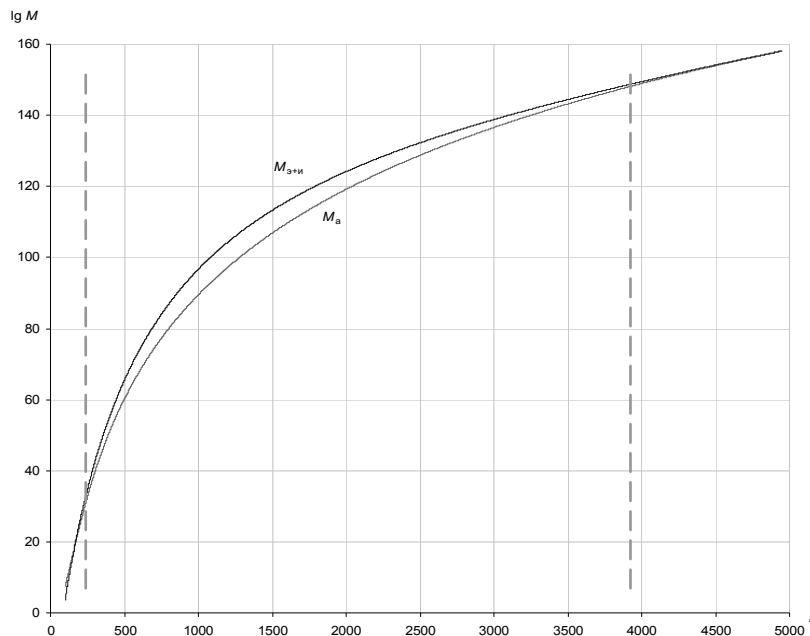


Рисунок 1 – Зависимости числа путей  $M$  от числа ветвей  $L$  для сети неопределенной структуры размерности  $n = 100$

Для решения задачи поиска максимальных двухполюсных потоков в работе [7] представлен эффективный вычислительный алгоритм, который на основе использования процедуры двоичного кодирования сечений [8] позволяет сформировать не только матрицу максимальных междупольсных потоков  $V$ , но также и матрицу  $CS$  кодов сечений, каждый элемент которой  $c_{Sst}$  представляет двоичный код минимального сечения, определяющего величину максимального потока  $v_{st}$ . Размер квадратных матриц  $V$  и  $CS$  определяется числом пунктов сети. Сечению, которое разделяет, например, вершины 1, 2, 5 и 3, 4, соответствуют строки 1, 2, 5 и столбцы 3 и 4 матрицы  $C$  пропускных способностей ветвей сети, на пересечении которых расположены элементы  $c_{13}, c_{14}, c_{23}, c_{24}, c_{53}, c_{54}$ , определяющие набор ветвей, входящих в указанное сечение. Вычисление значения пропускной способности сечения сводится к суммированию значений  $c_{xy}$  ветвей, которые входят в сечение.

Рассмотрим как пример ориентированный граф, изображенный на рис. 2, матрица пропускных способностей ребер графа представлена в табл. 2.

В приведенном примере величина пропускной способности сечения  $S$ , которое отделяет вершины 1, 2, 5 от вершин 3, 4, составляет:  $C(S) = c_{13} + c_{14} + c_{23} + c_{24} + c_{53} + c_{54} = 0 + 0 + 9 + 0 + 0 + 7 = 16$  ед.

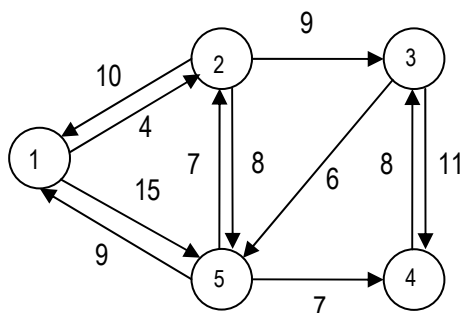


Рисунок 2– Ориентированный граф

Таблица 2

$$C = \begin{bmatrix} \infty & 4 & 0 & 0 & 9 \\ 10 & \infty & 9 & 0 & 8 \\ 0 & 0 & \infty & 11 & 6 \\ 0 & 0 & 8 & \infty & 0 \\ 15 & 7 & 0 & 7 & \infty \end{bmatrix}$$

Данный алгоритм позволяет одновременно с вычислением величин всех сечений определить максимальные потоки между всеми парами вершин графа. Алгоритм основан на том, что индексы всех элементов матрицы  $C_{xy}$ , которые находятся на пересечениях строк и столбцов, соответствующих коду сечения, совпадают не только с индексами всех ребер  $(x, y)$ , которые входят в сечение, но и с индексами всех междуполюсных потоков  $\varphi_{st}$ , которые разделяются этим сечением.

Для графа, который приведен на рис.2, матрица максимальных междуполюсных потоков и матрица кодов минимальных сечений приведены в таблицах 3 и 4, соответственно.

Таблица 3

$$V = \begin{bmatrix} \infty & 11 & 13 & 13 & 13 \\ 24 & \infty & 16 & 16 & 23 \\ 6 & 6 & \infty & 17 & 6 \\ 6 & 6 & 8 & \infty & 6 \\ 22 & 11 & 16 & 16 & \infty \end{bmatrix}$$

Таблица 4

$$CS = \begin{bmatrix} - & 11101 & 00001 & 00001 & 00001 \\ 01110 & - & 10011 & 10011 & 01111 \\ 01100 & 01100 & - & 00100 & 01100 \\ 01100 & 01100 & 01000 & - & 01100 \\ 11100 & 11101 & 10011 & 10011 & - \end{bmatrix}$$

Так, например, максимальный поток  $\varphi_{42} = 6$  ед., код соответствующего минимального сечения – 01100. Сечение составляют ребра (3, 1), (3, 2), (3, 5), (4, 1), (4, 2), (4, 5) (в табл. 4 нумерация разрядов кода и соответственно – номеров вершин графа – осуществлена справа налево). Величина сечения 01100 составляет:

$$C(01100) = C_{31} + C_{32} + C_{35} + C_{41} + C_{42} + C_{45} = 0 + 0 + 6 + 0 + 0 + 0 = 6 \text{ ед.}$$

При НВВ, в результате которого некоторая ветвь сети  $\beta_{xy}$  выйдет из строя, использование матрицы **CS** дает возможность получить необходимую информацию относительно того, какие максимальные потоки изменят свои величины при удалении ветви  $\beta_{xy}$  сети.

Представленные в работах автора [6, 7, 9] методы определения числа путей в сетях различной структуры, определения кратчайших путей и определения минимальных разрезов для всех максимальных потоков используются для определения относительного изменения показателей структурной живучести ТКС, предложенных в данной работе, – средневзвешенной величины максимального потока и средневзвешенной величины кратчайших путей, служащих для формирования средневзвешенной комплексной оценки относительного изменения показателей структурной живучести по всем требованиям на передачу информации.

### Метод повышения структурной живучести

Основной задачей ТКС является обслуживание потоков требований различных классов в соответствии с требуемым качеством обслуживания QoS (Quality of Service) для соответствующих классов потоков.

Имеется ТКС, модель которой описывается взвешенным ориентированным графом с заданными матрицами пропускных способностей ветвей сети (ребер графа)  $\mathbf{C} = [c_{ij}]$  и требований на передачу потоков информации  $\Phi = [\varphi_{ij}]$  ( $i, j = \overline{1, n}$ ,  $i \neq j$ ,  $n$  – количество пунктов сети). Каждый поток  $\varphi_{ij}$  принадлежит соответствующему классу  $k$  ( $k = \overline{1, K}$ ,  $K$  – число классов потоков). Каждому классу  $k$  присваивается приоритет исходя из заданного показателя качества обслуживания QoS. Для потоков соответствующего класса задается их значимость – «вес». Задаются допустимые относительные изменения:

а) комплексного показателя структурной живучести для сети в целом  $\Delta \overline{VP}_{срвз}^{\partial on}$ , учитывающего изменения средневзвешенных величин максимальных потоков и минимальных путей;

б) величины максимальных потоков  $\Delta^k_{st\ max}$  для соответствующих классов  $k$  потоков ( $s, t = \overline{1, n}$ ,  $s \neq t$ ,  $n$  – число пунктов сети).

Предложенный метод оценки и повышения структурной живучести ТКС основан на реконфигурации сети после НВВ, выражающемся в удалении некоторой ветви  $\beta_{xy}$  из сети. Реконфигурация сети осуществляется на основе определения трасс некоммутируемых каналов для передачи потоков требований с целью обеспечения заданного показателя качества обслуживания QoS. Метод состоит в выполнении следующих действий.

1. Определение средневзвешенных величин – максимального потока  $\overline{V}_{maxcp}$  (с использованием матриц максимальных междуполюсных потоков  $V$  и кодов сечений **CS**) и кратчайших путей  $\overline{P}_{mincp}$  (с использованием таблицы путей ТП) [7, 9].

2. Расчет средневзвешенных величин максимального потока в сети  $\overline{V}_{maxcp}^{нев}$  и кратчайшего пути в сети  $\overline{P}_{mincp}^{нев}$  после НВВ (удаления некоторой ветви  $\beta_{xy}$  из сети). Данный расчет осуществляется на основе скорректированной матрицы  $V^{нев}$ , получаемой из исходной  $V$  после определения значений тех максимальных потоков, которые изменились в результате НВВ, и скорректированной таблицы путей  $ТП^{нев}$ , получаемой из исходной ТП после определения значений тех кратчайших путей, которые изменились в результате НВВ.



3. Определение относительного изменения показателей структурной живучести – для максимальных потоков (4):

$$\Delta \bar{V}_{\max cp}^{HBB} = \frac{\bar{V}_{\max cp} - \bar{V}_{\max cp}^{HBB}}{\bar{V}_{\max cp}} \quad (4)$$

и для минимальных путей (5):

$$\Delta \bar{P}_{\min cp}^{HBB} = \frac{\bar{P}_{\min cp} - \bar{P}_{\min cp}^{HBB}}{\bar{P}_{\min cp}} \quad (5)$$

4. Формирование средневзвешенной комплексной оценки относительного изменения показателей структурной живучести:

$$\Delta \bar{VP}_{\text{срез}}^{HBB} = \Delta \bar{V}_{\max cp}^{HBB} * W_f + \Delta \bar{P}_{\min cp}^{HBB} * W_r,$$

где  $W_f$  и  $W_r$  – весовые характеристики оценок для потоков и путей, соответственно, определяемые исходя из условия:  $W_f + W_r = 1$ .

Данные весовые характеристики наиболее целесообразно получать экспертными методами.

5. Сравнение полученного значения относительного изменения комплексного показателя структурной живучести  $\Delta \bar{VP}_{\text{срез}}^{HBB}$  с заданным допустимым изменением:

$$\Delta \bar{VP}_{\text{срез}}^{HBB} \leq \Delta \bar{VP}_{\text{срез}}^{\text{дон}} \quad (6)$$

Здесь  $\Delta \bar{VP}_{\text{срез}}^{\text{дон}}$  – заданное допустимое относительное изменение комплексного показателя структурной живучести.

Если условие (6) выполнено, это означает, что HBB не оказало существенного влияния на живучесть сети, следовательно – никаких действий, связанных с изменением структуры ТКС, предпринимать не следует. Иначе – переход к п. 6.

6. Сравнение максимального относительного изменения максимальных потоков класса  $k$  с заданным изменением максимальных потоков  $\Delta_{st \max}^k$  соответствующего класса  $k$ , или приоритета:

$$\max_{(s,t)} \Delta V_{st \max}^{HBB}(k) \leq \Delta_{st \max}^k \quad (7)$$

7. Формирование перечня требований (потоков, для которых условие (7) не выполняется) на реконфигурацию сети:

$$V_{st \max}^{mp}(k) = V_{st \max}^{HBB}(k) + \Delta'_{st}(k),$$

где  $\Delta'_{st}(k)$  – то значение, которое обеспечивает требуемую величину  $V_{st \max}^{mp}(k)$  потока соответствующего класса  $k$ .

8. Построение матрицы допустимых путей  $M$  для всех требований  $V_{st \max}^{mp}(k)$  на реконфигурацию.

9. Формулирование задачи реконфигурации сети как задачи линейного программирования:

максимизировать взвешенную целевую функцию – суммарную емкость путей сети, которые представлены в матрице  $M$ , (или суммарную пропускную способность):

$$F = \sum_{\mu \in M} f_i c_i \rightarrow \max \quad (8)$$

(здесь  $i$  – номер пути,  $c_i$  – пропускная способность  $i$  –го пути,  $f_i$  – вес (ценность) пути  $\mu_i$ , который определяет относительную целесообразность использования пути и может быть выбран обратно пропорционально его длине, рангу, стоимости и т.п., т.е. выбор  $f_i$  зависит от критерия, по которому осуществляется оптимизация распределения каналов. При оптимизации по длине (рангу) в качестве  $f_i$  можно, например, взять отношение самого длинного пути множества  $M$  к длине (рангу) рассматриваемого  $i$ -го пути:

$$f_i = b \frac{l_{max}}{l_i} \text{ или } f_i = b \frac{r_{max}}{r_i},$$

где  $b$  – коэффициент, выбираемый таким образом, чтобы все  $f_i$  были целыми числами, что обеспечивает возможность получения целочисленного решения)

при следующих ограничениях:

- 1). Пропускные способности путей не могут принимать негативных значений, т.е.  $c_i \geq 0$  для всех  $\mu_i \in M$ ;
- 2). Суммарная емкость совокупности путей  $m_{st} = \{ \mu_i(s, t) \}$  между парой  $s$  и  $t$  пунктов сети, для которых существует требование  $V_{stmax}^{mp}$ , должна равняться (если это возможно) количеству каналов  $V_{stmax}^{mp}(k)$ , которые необходимо организовать в пучке между указанной парой пунктов, то есть

$$\sum_{\mu_i(s,t) \in m_{st}} c_i(s, t) \leq V_{stmax}^{mp}(k) \quad (9)$$

(здесь  $\mu_i(s, t)$  –  $i$ -й путь множества  $M$ , используемый для распределения требования  $V_{stmax}^{mp}(k)$ ,  $c_i(s, t)$  – емкость этого пути);

- 3). Для любой ветви (ребра графа)  $a(x, y)$  сети суммарная емкость всех путей, включающих эту ветвь, не может превышать емкости  $c(x, y)$  этой ветви:

$$\sum_{\forall \mu_i \in a(x,y)} c_i(s, t) \leq c(x, y). \quad (10)$$

10. Реконфигурация сети – формирование новой структуры с учетом полученных в результате решения задачи (в соответствии с целевой функцией (8) и ограничениями (9) и (10)) трасс прямых каналов для распределения требований  $V_{stmax}^{mp}(k)$ .

11. Для реконфигурированной сети (ркфс) – расчет средневзвешенных величин максимального потока  $\bar{V}_{maxcp}^{ркфс}$  и кратчайшего пути  $\bar{P}_{mincp}^{ркфс}$ .

12. Определение относительного изменения показателей структурной живучести реконфигурированной сети:

- для максимальных потоков (11):

$$\Delta \bar{V}_{maxcp}^{ркфс} = \frac{\bar{V}_{maxcp} - \bar{V}_{maxcp}^{ркфс}}{\bar{V}_{maxcp}} \quad (11)$$

- и для минимальных путей (12):

$$\Delta \bar{P}_{mincp}^{ркфс} = \frac{\bar{P}_{mincp} - \bar{P}_{mincp}^{ркфс}}{\bar{P}_{mincp}}. \quad (12)$$

13. Формирование средневзвешенной комплексной оценки относительного изменения показателей структурной живучести реконфигурированной сети:

$$\Delta \overline{VP}_{срез}^{ркфс} = \Delta \overline{V}_{maxcp}^{ркфс} * W_f + \Delta \overline{P}_{mincp}^{ркфс} * W_r \quad (13)$$

14. Сравнение полученного значения (13) относительного изменения комплексного показателя структурной живучести реконфигурированной сети  $\Delta \overline{VP}_{срез}^{ркфс}$  с заданным допустимым изменением по условию (14):

$$\Delta \overline{VP}_{срез}^{ркфс} \leq \Delta \overline{VP}_{ср}^{доп} \quad (14)$$

Выполнение условия (14) означает, что получено обеспечение требуемых показателей живучести сети. В противном случае – выработка следующих возможных рекомендаций:

- увеличение пропускных способностей определенных ветвей сети на вычисленную в результате решения задачи величину;
- изменение приоритетов определенных требований;
- ограничение величины поступающих в сеть требований;
- изменения задаваемых предельно допустимых значений относительных изменений для потоков соответствующего класса  $\Delta^k_{st\ max}$  или для сети в целом  $\Delta \overline{VP}_{срез}^{доп}$ .

## Заключение

Представлен метод оценки и повышения структурной живучести ТКС путем реконфигурации сети, выполняемой на основе оценки изменения показателей структурной живучести в результате неблагоприятного внешнего воздействия. Предложены показатели структурной живучести – средневзвешенные величины максимального потока  $\overline{V}_{maxcp}$  и кратчайших путей  $\overline{P}_{mincp}$ , служащие для формирования средневзвешенной комплексной оценки относительного изменения показателей структурной живучести  $\Delta \overline{VP}_{срез}^{нвс}$ . Задача повышения структурной живучести ТКС путем реконфигурации сформулирована как задача линейного программирования, решение которой либо позволит обеспечить требуемые показатели живучести ТКС, либо обеспечит выработку рекомендаций по: увеличению пропускных способностей определенных ветвей сети на вычисленную в результате решения задачи величину; изменению приоритетов определенных требований; ограничению величины поступающих в сеть требований; изменению задаваемых предельно допустимых значений относительных изменений для потоков соответствующего класса  $\Delta^k_{st\ max}$  или для сети в целом  $\Delta \overline{VP}_{срез}^{доп}$ .

## Литература

1. Рекомендация ITU-T Y.3001: Глобальная информационная инфраструктура, аспекты протокола интернет и сети последующих поколений / ITU: Committed to connecting the world. – Режим доступа: <http://www.itu.int> / (дата обращения 10.03.2012 г.).
2. Додонов А.Г., Ландэ Д.В. Живучесть информационных систем. – К.: Наук. думка, 2011. – 256 с.
3. Стекольников Ю.И. Живучесть систем – СПб.: Политехника, 2002. – 155с.
4. Синтез и анализ живучести сетевых систем : монография / Ю.Ю. Громов, В.О. Драчев, К.А. Набатов, О.Г. Иванова. – М. : «Изд-во Машиностроение-1», 2007. – 152 с.

- 
5. Руденко Ю.Н., Ушаков И.А. Надежность систем энергетики / Ю.Н. Руденко, И.А. Ушаков – 2-е изд., перераб. и доп. – Новосибирск: Наука. Сиб. отд-ние. – 1989. – 328 с.
  6. Князева Н.А., Ненов А.Л. Метод оценки структурной надежности сети при изменении ее структуры. – К.: Вісник ДУІКТ. Т.9, №4, 2011 р. – С. 318–325.
  7. Алгоритмы оценки структурной живучести инфокоммуникационной сети 2012р., Київ, ДУІКТ. Сучасні інформаційно-комунікаційні технології. VIII наук.-техн. конф.: Збірник тез. С. 192-193.
  8. Кутасин Б.П., Ящук Л.Е. Некоторые вопросы применения ЭВМ для управления сетью связи//Труды ОЭИС им. А.С.Попова. Вып. 18. Одесса: изд ОЭИС, 1969. – С. 3–12.
  9. Князева Н.О. Теорія проектування комп'ютерних систем і мереж. Ч.2. Методи аналізу і синтезу комп'ютерних систем і мереж, Одеса: СПД, 2012. – 240 с.

---

### Благодарности

---

*Настоящая работа была выполнена при поддержке интернационального проекта **ITHEA XXI Института информационных теорий и их приложений FOI ITHEA и Ассоциации ADUIS Украина (Ассоциация разработчиков и пользователей интеллектуальных систем)**, в связи с чем автор и выражает свою искреннюю благодарность.*

---

### Информация об авторах

---

**Нина Князева** – Факультет информационных технологий и кибербезопасности Одесской национальной Академии пищевых технологий, доктор технических наук, профессор, заведующая кафедрой информационных систем и сетей; ул. Дворянская, 1/3, Одесса-26, 65026, Украина; тел. (048)-7209176; моб. (38067)-430-17-28; E – m: knyazeva@ukr.net

*Главные области научного исследования: Управление в телекоммуникационных сетях, проблемы повышения надежности и живучести телекоммуникационных сетей*

---

## РАЗРАБОТКА МОДЕЛИ ПОЛЬЗОВАТЕЛЯ ИНФОРМАЦИОННЫХ СЕТЕЙ КАК ПЛОХОФОРМАЛИЗОВАННОГО ОБЪЕКТА

Анастасия Смирнова

**Аннотация:** *предложен подход к созданию модели пользователя информационных сетей путем формализации одного из важнейших их входных параметров. Модель служит для повышения эффективности и достоверности моделирования информационных сетей. В предложенной модели критерием формирования является применимость ее для анализа и синтеза информационных сетей.*

**Ключевые слова:** *информационные сети, пользователь информационных сетей, плохоформализованные объекты, глобальная информационная инфраструктура*

**Ключевые слова классификации ACM:** *1.6 SIMULATION AND MODELING - 1.6.5 Model Development, C.2 COMPUTER-COMMUNICATION NETWORKS - C.2.1 Network Architecture and Design, K. COMPUTING MILIEUX - K.6 Management of computing and information system*

---

### Введение

Сейчас, когда общество меняется из индустриального к информационному, особенно остро стоит вопрос создания сети, в которой каждый человек сможет иметь возможность создавать информацию и знания, находить и распространять их, иметь к ним доступ, обмениваться ими так, чтобы каждый индивидум имел возможность получить их и полностью реализовать свой потенциал, следуя духу «Всеобщей декларации прав человека». Ведь для развития каждого человека основу составляют именно образование, знания и информация, свободный доступ к ним приведет к благосостоянию общества.

Раньше инфокоммуникационные и телекоммуникационные сети создавались, оптимизировались и изменялись путем аналитических оценок и статистических расчетов. Однако превращение их в информационные сети (ИС) не может быть проанализировано такими методами, а требует моделирования и формализации каждого из этих шагов [1, 3, 4, 6]. Следовательно, с целью анализа, синтеза и оптимизации нужно решить задачи условного разбиения ИС на составные части, и формализовать каждую из них для последующего моделирования их взаимоотношений и состояний в рамках сложной системы.

Концепция Глобальной информационной инфраструктуры (ГИИ) имеет дуальную сущность. С одной стороны глобальность информационной инфраструктуры обеспечивается созданием всемирной сети связи, которая сочетает в себе национальные, региональные и ведомственные сети коммуникаций. С другой, она предполагает факт персонализации связи для каждого отдельного пользователя [15-17].

В рамках концепции ГИИ пользователи являются источниками и потребителями информации, которые используют услуги ИС, и создают потоки сообщений родного вида и назначения. Именно пользователи предъявляют к сети требования по доставке и обработке информации с соблюдением отдельных качественных и количественных показателей. Информационная сеть предоставляет пользователям набор

---

различных видов связи и услуг, как по обработке информации, так и ряд других, связанных с облегчением пользования связью и получения различной информации.

Учитывая вышесказанное, можно утверждать, что центральным объектом ИС является пользователь.

Существующие модели, используемые для формализации ИС, зачастую являются уровневыми и основаны на семиуровневой модели взаимодействия открытых систем [2, 4], однако встречается и четырехуровневая модель, а также *Stratified Reference Model (SRM)*, описанная в [2].

Однако четкого определения – кто же он, пользователь информационной сети, до сих пор нет, поэтому актуальным заданием является повышение эффективности и достоверности моделирования ИМ путем формализации одного из наиболее важных входных параметров – модели пользователя ИС.

---

### Постановка задачи

---

На первом этапе формализации ИС может быть использована уровневая модель для определения процессов, протекающих в сети на каждом из ее уровней, и методы математического, имитационного или эвристического моделирования для описания каждого из них. Затем следует сформулировать (если это возможно) модель процессов, протекающих в ИС в целом. На основе формализованной модели процессов создается формализованная модель ИМ в целом, которая методами аналогий и предположений доказывает свою адекватность реальности. То есть все разработанные математические модели каждого из уровней формализуются и «вводятся» в надсистему – в данном случае ИС. Эта модель может тестироваться на ЭВМ, благодаря чему может быть выполнен ряд важных для разработки и создания ИС операций – от проверки ее адекватности реальности до синтеза конкретных ИС.

На данный момент существуют работы по концепции доступа пользователя, сформулировано понятие служб, услуг и платформ, которые представляют, проведен анализ вариантов организации линий доступа [1, 3, 6, и др.], и разработаны стандартизированные интерфейсы протоколы доступа пользователя (рекомендации МСЭ серии Y), а также методов распределения потоков информации в ИС, методы разработки структуры ИМ в зависимости от различных факторов [3], разработаны методы классификации услуг, нормативно-правовые акты и технические требования к базовым сетям. Относительно пользователя ИМ - его описания существуют с точки зрения сети и ее организации, расчета потока данных и нагрузку на сеть, что он создает.

Но проблема отсутствия формализованного описания главного фигуранта ИС - пользователя - остается открытой.

Пользователи являются источниками и потребителями информации, которые используют услуги ИС, и создают потоки сообщений различного вида и назначения. Именно пользователи предъявляют к сети требования по доставке и обработке информации с соблюдением отдельных качественных и количественных показателей.

Сложность объекта, который представляет собой пользователь ИС, определяет его многогранность, неявная взаимосвязь и взаимоотношение его характеристик, а также сложность в формализации этих параметров. Однако четких методов анализа плохоформализованных объектов для исследуемой предметной области еще не предложено.

Настоящая работа посвящена разработке формализованной модели пользователя ИС, которая в дальнейшем позволит произвести исследования, направленные на задачи изучения и моделирования ИС в целом.

---

## Разработка модели пользователя информационных сетей как плохоформализованного объекта

---

Пользователь может быть представлен как плохоформализованный объект по многим причинам. Ведь плохоформализованные объекты обладают свойствами, плохо известными априори и меняющимися в процессе функционирования [9]. Современный пользователь, как плохоформализованный объект, должен быть подвергнут системному анализу, который в данном случае выступает инструментом познания сложного объекта. Однако следует обратить особое внимание на то, что, несмотря на большое количество разработок в области системного анализа плохоформализованных объектов, нельзя сказать, что разработанные методы создали идеальный инструмент для проведения системного анализа. Методики различных ученых и аналитиков (Стенфорд Оптнер, Спартак Никаноров, Юрий Черняк, Анатолий Катренко) разнятся, и одна из причин этого – различные предметные области, в которых был применен этот системный анализ. Для информационных технологий такой методики пока не создано.

Для четкого понимания, что же собой представляет плохоформализованный объект, исследователи предлагают рассматривать с одной стороны как набор параметров, требований и критериев (которые предъявляет пользователь), а с другой – как закономерности, зависимости и механизмы взаимосвязи этих параметров с входными состояниями объекта (финансовые возможности пользователя и стоимость услуг, предоставляемых сетью). Классификация таких объектов представляет собой упорядочение и разработку схемы соотношения параметров объекта, данных о нем и его результирующего поведения в различных условиях.

Рассматривая объект с организационной, управляющей точки зрения следует понимать, что существовать он будет в активной среде (надсистеме), и взаимодействовать будет с ней, поэтому попытки охарактеризовать его автономно от нее являются неразумными, ведь, в конечном счете, именно надсистема влияет на стабильное функционирование и обеспечивает нормальную жизнедеятельность объекта [10]. Из вышесказанного вытекает необходимость изучения как поведения пользователя, так и требований, которые он предъявляет к ИС.

Для рассмотрения сложной системы, которой является плохоформализованный объект, можно применить асимптотический анализ – аппроксимацию сложных объектов более простыми. То есть пользователь будет представлен как набор дискретных параметров и характеристик.

Как сформулировать требования к формализации пользователей так, чтобы они касались только «жизненно важных», которые имеют непосредственное влияние на поведение пользователя, аспектов, отбросив малозначимые и взаимозаменяемые? Помня о дуальной структуре плохоформализованного объекта, нужно рассмотреть две взаимодействующих области – внутреннюю и внешнюю среду для объекта. В конкретном случае – характеристики пользователей ИС, их потребности и требования к информационным услугам (ИУ) и их характеристики. Первые определены с помощью проведенного социального исследования (анкетирования) достаточной выборки пользователей ИС, анализа рынка пользователей ИУ, а также анализа социологических и маркетинговых исследований, которые уже были выполнены с подобной целью. Вторые представляют собой набор характеристик ИУ, которые может предоставить современная информационная сеть, связанные с пользователем и его требованиями ИУ.

В работе предложен ряд характеристик пользователя, описывающих его поведение в ИС. Необходимость снижения размерности исходных данных является одной из основных вех в системном анализе, поэтому, рассматривая предложенный ряд характеристик пользователя ИС и характеристик современных ИП, было выдвинуто предположение о нечеткой, но логически обоснованной взаимосвязи некоторых характеристик. Заметим также, что все эти предположения являются продуктами эвристической

деятельности и имеют вероятностную природу. Полученные характеристики были минимизированы, и была разработана модель пользователя ИС, основанная на предложенных характеристиках, которая представлена на рис. 1.

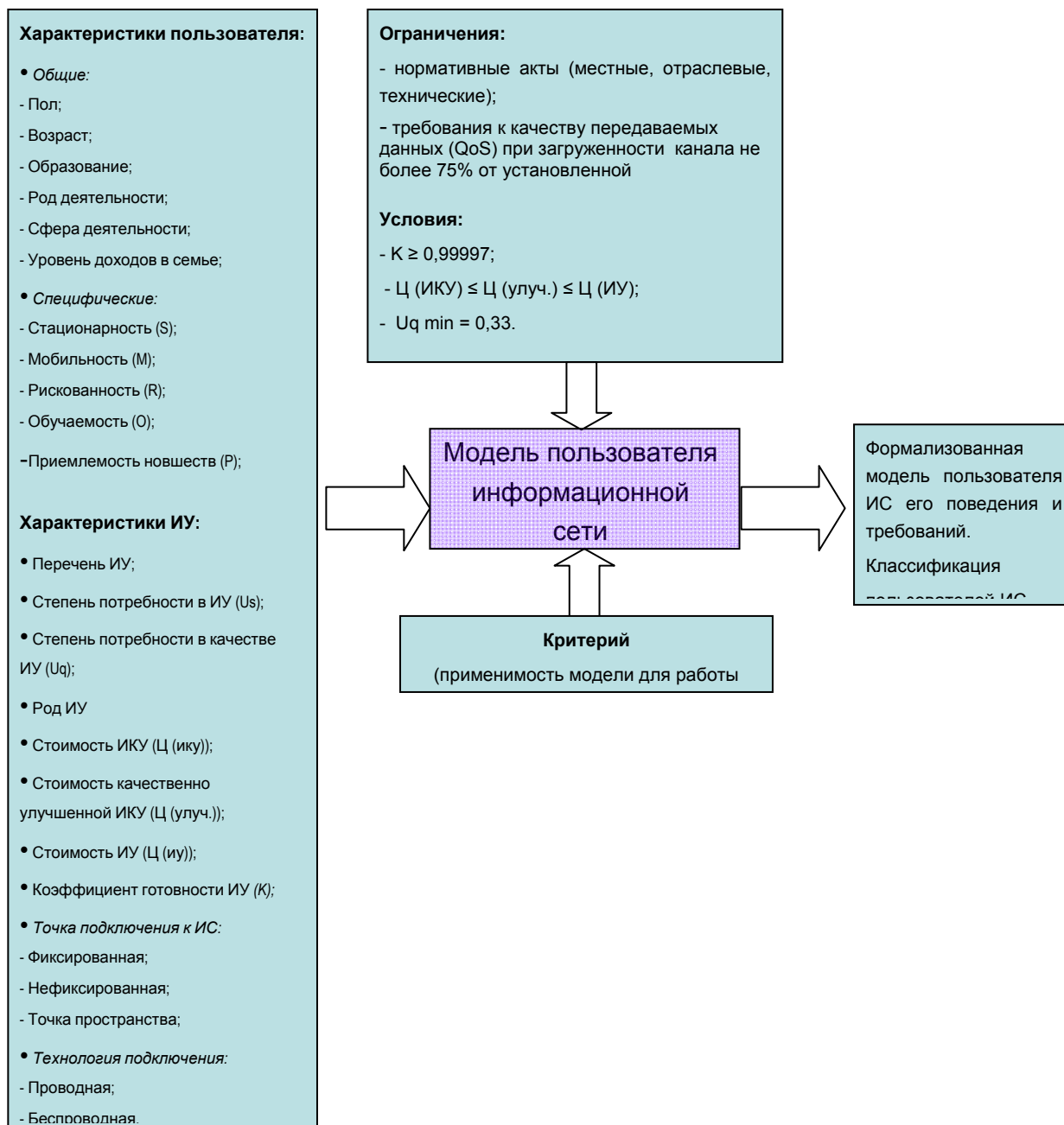


Рисунок 1. Модель пользователя информационных сетей

Выражение 1 описывает формализованную модель пользователя ИС:

$$U = \left\{ A, Sx, E, Ar, F, S, M, R, O, P, Us, Uq, \dots \right\} \quad (1)$$

$K \geq 0,99997, Ц_{ИКУ} < Ц_{улуч.} \leq Ц_{ИУ}$

где  $A$  – возраст,



*Sx* – пол,  
*E* – образование,  
*Ar* – сфера деятельности,  
*F* – вид деятельности,  
*S* – стационарность,  
*M* – мобильность  
*R* – рискованность  
*O* – обучаемость,  
*P* – приемлемость новшеств  
*Us* – требования к стоимости услуги,  
*Uq* – требования к качеству услуги,  
*Ц(ику)* – стоимость информационно-коммуникационной услуги,  
*Ц(улуч.)* – стоимость качественно улучшенной информационно-коммуникационной услуги,  
*Ц(иу)* – стоимость информационной услуги,  
*K* – коэффициент готовности услуги.

Каждая из характеристик описывает определенную сторону поведения пользователя, имеет определенную шкалу и значение ее ранжируется в пределах от 0 до 1.

Для того, чтобы доказать возможность применения и адекватности модели по нынешним реалиям и оценки характеристик соответствия создаваемой модели поведению современных реальных пользователей ИС, использующих ИУ, проведено социологическое исследование среди случайной группы пользователей информационной сети. На основании модели пользователя ИУ, предложенной выше, сделан вывод о необходимом для однозначного определения каждого конкретного пользователя набора характеристик. Предложено провести анкетирование пользователей с целью получения обучающей выборки входных данных, используя которые, в дальнейшем можно было бы получить конкретное представление о группах пользователей ИС и их характеристики, а также последующего выделения признаков или групп признаков, характерных для каждого класса, применив методы классификации.

Путем классификации характеристик пользователей с помощью методов кластерного анализа, используя различные методы объединения (метод одиночной связи, полной связи, невзвешенного попарного среднего, взвешенного попарного среднего, медианы, Варда) и различные меры расстояния (метрика Эвклида, квадрат метрики Эвклида, метрика Чебышева, Манхеттенская метрика), автор пришел к выводу, что формализованную модель можно уточнить описанием взаимосвязей характеристик, которые явно не коррелируют. Это позволило привести формализованную модель пользователя информационной сети, описанную в выражении 1, к дополненному выражению 2.

$$\begin{aligned}
 & K \geq 0.99997, \text{Ц(ику)} < \text{Ц(улуч.)} \leq \text{Ц(конв.)} \\
 U = & \overbrace{\{(M, Ar), (A, F), (Uq, R), (O, P), (Sx, E), S, Us, \dots\}}
 \end{aligned} \tag{2}$$

После этого разработан программный продукт на языке C++ в программной среде QT, в основе которого лежит метод растущих пирамидальных сетей [12-14]. Результаты его позволили выделить характерные для нескольких классов пользователей совокупности признаков, что описаны в выражении 3,

$$\begin{aligned}
 \text{Клас I} = \{ & Sx(m), Ar(\text{инженерная деятельность}), E(0,6\dots 0,79), M(0..0,19), S(0,4\dots 0,59), \\
 & P(0,4\dots 0,59), O(0,8\dots 1)\}
 \end{aligned} \tag{3}$$

---

Клас II = { $Sx(m)$ ,  $Ar$ (сфера услуг),  $E(0,6...0,79)$ ,  $M(0..0,19)$ ,  $S(2)$ ,  $Uq(0,2...0,39)$ }

Клас III = { $Sx(w)$ ,  $Ar$ (производство),  $E(0,6...0,79)$ ,  $M(0,2...0,39)$ ,  $S(0)$ ,  $O(0,8...1)$ ,  
 $P(0,4...0,59)$ }

Клас IV = { $Sx(m)$ ,  $Ar$ (инженерная деятельность),  $E(0,6...0,79)$ ,  $M(0..0,19)$ ,  $S(0,2...0,39)$ ,  
 $O(0,8...1)$ ,  $P(0,4...0,59)$ ,  $R(0,4...0,59)$ }

Клас V = {...}.

Приводя это описание к предложенным до формализации значениям характеристик, можно сказать, что класс I может включать в себя мужчин, занимающихся инженерной деятельностью, имеющих высшее образование, имеющих необходимость в мобильности для подключения к ИС в пределах 100 км от областного центра, имеющих высокую потребность в стационарной точке подключения к ИС, готовых с энтузиазмом воспринимать новую информацию с сфере информационных технологий и желающих обучаться в использовании услуг ИС. Сходное описание дается по всем классам. Класс V включает в себя пользователей, не попавших в предыдущие четыре класса.

---

### Заключение

Предложенная в работе модель пользователя информационных сетей является гибкой и динамичной для пополнения ее информацией, и компактной для внедрения ее в модель информационной сети в целом, отображая внутренние и внешние процессы взаимодействия этих моделей, что поможет в оптимизации, синтезе и управлении информационными сетями. Более того, разработанная модель имеет практическую ценность, так как основанная на ней классификация пользователей ИС, использующая методы кластерного анализа и методы растущих пирамидальных сетей, может быть применима разработчиками информационных сетей, так как изучение пользовательского контингента и правильная его стратификация является одной из целевых задач при построении ИС.

---

### Библиография

1. Гайворонская Г.С. Сети и системы телекоммуникаций (т.1) / Г.С. Гайворонская, М.В. Захарченко, А.И. Ещенко и др. // К.: Техника. – 2000. – 304 с.
2. Гайворонская Г.С., Сетевая модель как метод анализа информационных сетей // VII Міжнародна науково-технічна конференція «Сучасні інформаційно-комунікаційні технології». COMINFO'2011-Livadia. 10-14.09.2011. – Київ. – С. 40.
3. Гайворонська Г.С., Інформаційна мережа як об'єкт аналізу і синтезу: навчальний посібник з дисципліни «Оптимальний синтез інформаційних мереж». ОДАХ – 2011. – С. 46-50.
4. Пасечников И.И., Методология анализа и синтеза предельно нагруженных информационных сетей. Москва. И.: «Машиностроение-1». – 2004.
5. Гайворонская Г.С.. Формализация модели эволюции телекоммуникационных сетей // *Applicable Information Models ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA and the Consortium FOI Bulgaria*. – С. 155-168
6. Гайворонская Г.С. Концепция пользовательского доступа. И.:ОГАХ. – Одесса. – 2008. – С. 26-34
7. Гайворонская Г. С. Уровневая модель представления инфокоммуникационных услуг // *Зв'язок* 2007. №1. – С. 49-55.
8. Лившиц Б.С., Теория телетрафика / Лившиц Б.С., Пшеничников А.П., Харкевич А.Д., – М.: «Связь». – 1979. – С. 19.

- 
9. Кобозева А.А., Анализ информационной безопасности / А.А.Кобозева, В.А.Хорошко. – К.: Изд. ГУИКТ. – 2009. – С. 251.
  10. Крисиллов, А.Д., и др., Краткий методологический меморандум, ч.1 // XV Интернациональная конференция "Knowledge-Dialogue-Solution" KDS-2 2009. материалы конференции *International Book Series "Information Science and Computing"*. – С. 257-267.
  11. Вятчинин Д.А., Нечеткие методы автоматической классификации. Изд.: УП «Технопринт». – Минск. – 2004. – С. 7-16.
  12. Гладун В.П., Партнерство с компьютером. Человеко-машинные целеустремленные системы, К.: *Port-Royal*, – 2000. – С. 17.
  13. Гладун В.П., Растущие пирамидальные сети // *Новости искусственного интеллекта*. – 2004. – С. 32-40.
  14. Гладун В.П., Ващенко Н.Д., Величко В.Ю., Прогнозирование на основе растущих пирамидальных сетей // *Программные продукты и системы*, 2002. Вып. 2. – С. 26.
  15. *ITU-T Y.100 – 06.1998. General overview of the Global Information Infrastructure standards development.*
  16. *ITU-T Y.101 – 03.2000. Global Information Infrastructure terminology: Terms and definitions.*
  17. *ITU-T Y.120 – 06.1998. Global Information Infrastructure scenario methodology.*
  18. Смирнова А.С. «Построение модели пользователя конвергентной сети с помощью метода растущих пирамидальных сетей» // VI Международная научно-техническая конференция и IV студенческая научно-техническая конференция «Проблемы телекоммуникаций». 27.04.12. – Киев. – С. 325-327.

---

### Информация об авторе

---

**Анастасия Смирнова** – Одесская государственная академия холода; студентка кафедры информационно-коммуникационных технологий – ул. Дворянская., 1/3, Одесса – 82, 65082, Украина; e-mail: [asya.smi@gmail.com](mailto:asya.smi@gmail.com)

Основные области исследований: информационные сети, телекоммуникационные сети, формализация сложных и слабоформализованных объектов

## ПРОГРАММНЫЙ КОМПЛЕКС АНАЛИЗА КОНЪЮНКТУРЫ ОПТОВОГО РЫНКА ЭЛЕКТРОЭНЕРГИИ И МОЩНОСТИ

Галина Старкова

**Abstract:** В статье приводится анализ текущего состояния рынка электроэнергии Российской Федерации, рассматриваются ключевые моменты построения программного комплекса анализа конъюнктуры оптового рынка электроэнергии и мощности, предназначенного для операционного и долгосрочного прогнозирования потребления электроэнергии, а также для поддержки принятия инвестиционных решений по приобретению или продаже активов, по размещению объектов строительства на новых или на уже действующих электростанциях. Расширение числа внешних связей топливно-энергетического комплекса, переход от создания изолированных экономических и энергетических моделей, приводят к необходимости создания программных комплексов моделей, позволяющих отражать процесс энергопотребления в отдельных секторах экономики с различным уровнем детализации и агрегации моделируемых показателей, позволяющих также проводить многовариантные сценарные расчёты. Созданный программный комплекс анализа конъюнктуры оптового рынка электроэнергии и мощности содержит модули моделирования и прогнозирования конечного потребления электроэнергии, потерь электроэнергии в сетях, числа часов использования мощности, а также локальных максимумов потребления электроэнергии на среднесрочную и долгосрочную перспективу. Полученные результаты могут быть использованы для формирования рациональной стратегии развития энергетики как страны в целом, так и отдельных её субъектов в частности. Программный комплекс региональной модели конъюнктуры оптового рынка электроэнергии и мощности реализован на базе PROGNOS Platform (интегрированная платформа для создания информационно-аналитических систем и систем поддержки принятия решений, объединяющая технологии построения хранилищ данных, средства формирования отчётности и оперативного анализа данных, инструменты моделирования и прогнозирования).

**Keywords:** конъюнктура оптового рынка электроэнергии и мощности РФ; моделирование электропотребления; эконометрическое моделирование; информационно-аналитические системы; системы поддержки принятия решений; хранилища данных.

**ACM Classification Keywords:** H. Information Systems. H.4 Information Systems Applications: H.4.2 Types of Systems – Decision support (e.g., MIS).

**Conference topic:** Business Informatics.

---

### Введение

Электроэнергия является одним из наиболее значимых продуктов, предназначенных для промежуточного потребления предприятиями, и составляет существенную долю в затратах многих видов экономической деятельности. Дефицит электроэнергии неизбежно влечёт к ограничению экономического роста. Кроме

того, недооценка ожидаемого потребления электроэнергии приводит к необходимости использования дорогих пиковых станций. Завышенный прогноз потребления электроэнергии приводит к увеличению издержек на поддержание в рабочем состоянии излишних резервных мощностей. Производители электроэнергии заинтересованы в прогнозах электропотребления с целью оперативного реагирования на спрос и с целью оптимального развития инфраструктуры. Эффективность мероприятий по управлению потреблением электроэнергии, качество планирования и экономичность режимов работы энергосистемы определяются достоверностью прогноза. Повышение точности прогноза потребления электроэнергии способствует экономии энергоресурсов, увеличению эффективности работы и соответственно увеличению прибыли предприятий энергетики.

Таким образом, прогнозирование динамики электропотребления в стране приобретает особую актуальность. Прогнозирование электропотребления является многоэтапным и многоуровневым процессом, результаты которого могут использоваться для формирования рациональной стратегии развития энергетики как страны в целом, так и отдельных её субъектов в частности.

### Анализ российского рынка электроэнергетики

В экономике энергетической отрасли по-прежнему сохраняются долгосрочные тенденции. Эти тенденции проявляются в увеличении объёма потребляемой энергии, повышении производства и потребления энергии, а также растущей диверсификации её источников [1]. Наиболее мощными движущими силами спроса на энергию выступают рост численности населения и рост доходов населения. По данным компании BritishPetroleum «за последние 20 лет численность населения мира увеличилась на 1,6 млрд. человек и, как ожидается, возрастет на 1,4 млрд. человек в предстоящие 20 лет. Реальный доход в мире увеличился на 87% за прошедшие 20 лет и, скорее всего, возрастет на 100% в течение следующих 20 лет» [2].

В настоящий момент в составе Единой Энергетической Системы (ЕЭС) России функционируют 69 региональных энергосистем, которые, в свою очередь, образуют 7 Объединённых энергосистем (ОЭС): Средней Волги, Урала, Центра, Северо-Запада, Юга и Сибири. Развитие Единой энергетической системы России происходило путём поэтапного объединения и организации параллельной работы региональных энергетических систем, формирования межрегиональных объединённых энергосистем (ОЭС) и их последующего объединения в составе Единой энергетической системы. Переход к такой форме организации был обусловлен необходимостью более рационального использования энергетических ресурсов, повышения экономичности и надёжности электроснабжения страны. Фактическое потребление электроэнергии по ЕЭС России в 2011 г. составило 1 038 250 тыс. МВт·ч, что выше фактического значения в 2012 г. на 1,7% и выше значения в 2010 г. на 2,8% (рис. 1, [3]).

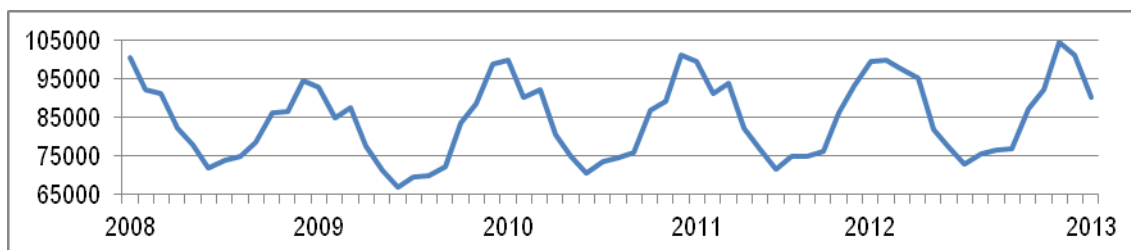


Рисунок 1. Объём потребления электроэнергии в РФ (тыс. МВт·ч/мес)

Потребление электроэнергии в России носит волатильный характер, тесно связанный с «утяжелённой» структурой экономики, которая характеризуется потреблением больших объёмов электроэнергии электроёмкими отраслями промышленности при относительно низкой диверсификации промышленности и незначительной доле неэлектроёмкой сферы услуг.

Динамика электропотребления РФ преимущественно зависит от динамики промышленного производства, что отражается в высокой эластичности электропотребления в промышленном производстве (табл. 1).

Таблица 1. Структура потребления электроэнергии в 2011 г. [4]

	млн.кВт.час	Доля, %
Добыча полезных ископаемых, обрабатывающие производства, производство и распределение электроэнергии, газа и воды	567 909,3	54,5
Сельское хозяйство, охота и лесное хозяйство	15 525,1	1,5
Строительство	11 810,9	1,1
Транспорт и связь	90 354,9	8,7
Другие виды экономической деятельности	119 630,5	11,5
Население	130 889,3	12,6
Потери в электросетях	105 002,1	10,1
Всего	1 041 122,1	100,0

Современный электроэнергетический комплекс России насчитывает около 700 электростанций единичной мощностью свыше 5 МВт. Структура выработки электроэнергии по типам электростанций ЕЭС России представлена на (рис. 2, [5]).

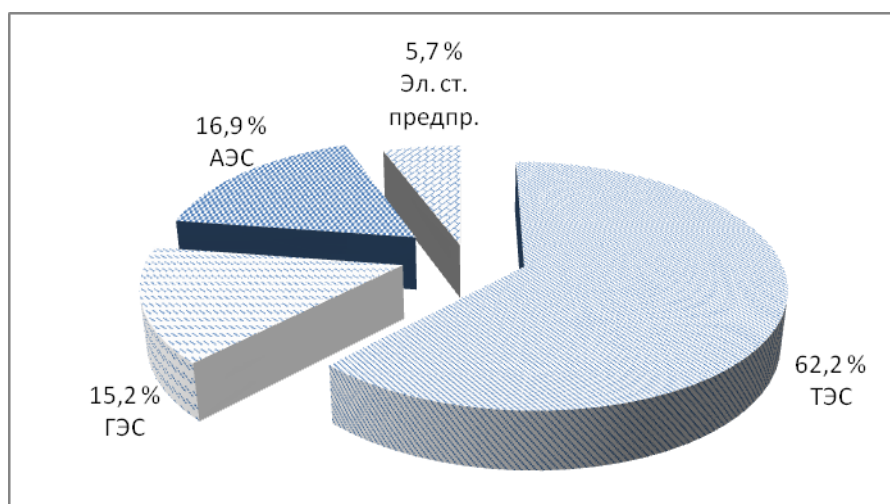


Рисунок 2. Структура выработки электроэнергии по типам электростанций ЕЭС России

Лидирующее положение тепловыработки является исторически сложившейся и экономически оправданной закономерностью развития российской энергетики. Наибольшее распространение получили тепловые электростанции общего пользования, работающие на органическом топливе (газ, уголь). Основной задачей развития тепловой генерации является обеспечение технического перевооружения и

---

реконструкции действующих электростанций, а также ввод новых генерирующих мощностей. Гидроэнергетика является ключевым элементом обеспечения системной надёжности Единой Энергосистемы страны, располагая более 90% резерва регулировочной мощности. ГЭС являются наиболее манёвренными и способны при необходимости быстро существенно увеличить объёмы выработки, покрывая пиковые нагрузки. В настоящий момент на территории РФ эксплуатируется 10 атомных станций, рассоложенных преимущественно в Европейской части России и на Северо-западе. В стадии строительства находятся ещё 5 атомных электростанций.

---

### **Программный комплекс анализа конъюнктуры оптового рынка электроэнергии и мощности**

---

Изучив существующие методы и модели прогнозирования энергопотребления в целом, и электропотребления в частности, были выявлены следующие тенденции:

- расширение числа внешних связей топливно-энергетического комплекса;
- переход от создания изолированных экономических и энергетических моделей к их синтезу;
- создание комплексов моделей, позволяющих отражать процесс энергопотребления в отдельных секторах экономики с различным уровнем детализации и агрегации моделируемых показателей, позволяющих также проводить многовариантные сценарные расчёты.

Так возникла идея создания программного комплекса анализа конъюнктуры оптового рынка электроэнергии и мощности РФ. Целями создания данного программного комплекса является операционное и долгосрочное прогнозирование потребления электроэнергии, а также поддержка принятия инвестиционных решений по приобретению или продаже активов, по размещению объектов строительства на новых или на уже действующих электростанциях.

Объектом моделирования данного программного комплекса выступает конъюнктура регионального рынка электроэнергии в части спроса на электроэнергию и мощность с учётом влияния конечных цен для основных групп потребителей: промышленных предприятий, домохозяйств, органов муниципального управления и предприятий, предоставляющих жилищно-коммунальные услуги и других групп конечных потребителей. Согласно 3 статьи ФЗ № 35 от 23.03.2003 в редакции от 29.06.2012 под оптовым рынком электроэнергии будем понимать сферу обращения особых товаров – электрической энергии и мощности в рамках Единой энергетической системы России в границах единого экономического пространства РФ с участием крупных производителей и крупных покупателей электрической энергии и мощности, а также иных лиц, получивших статус субъектов оптового рынка и действующих на основе правил оптового рынка [6].

Программный комплекс анализа конъюнктуры оптового рынка электроэнергии и мощности включает в себя четыре основных блока:

1. Блок прогнозирования конечного потребления электроэнергии (рис. 3).
2. Блок прогнозирования потерь электроэнергии в сетях.
3. Блок прогнозирования числа часов использования мощности (ЧЧИМ).

## 4. Блок прогнозирования локальных максимумов потребления.

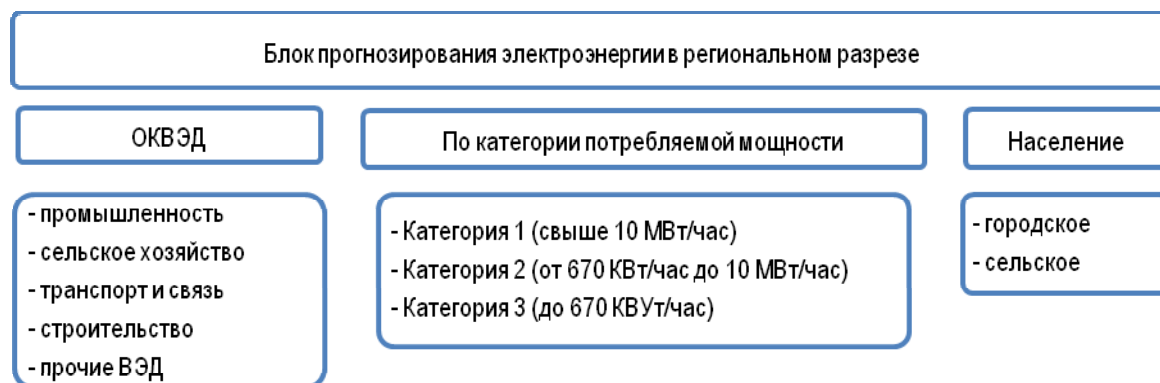


Рисунок 3. Классификация потребителей, для которых строятся прогнозы конечного электропотребления

Прогнозирование электропотребления на год вперёд в программном комплексе осуществляется в динамике по месяцам, на следующие 4 года – в квартальной динамике и на следующие 15 лет – в годовой динамике. Для решения поставленных задач было принято решение о моделировании электропотребления первые 5 лет в ежемесячной динамике с последующей её агрегацией до квартальной динамики, соответственно со 2 по 5 годы включительно.

Для обеспечения программного комплекса актуальными данными, необходимыми для расчёта результирующих показателей, были использованы данные таких официальных источников, как Росстат, некоммерческое партнёрство «Совет рынка», Системный оператор Единой энергетической системы, Федеральная служба по тарифам, данные метеорологических служб, IEA – Electricity Information, IEA – World Energy Outlook, IEA – World Energy Statistics and Balances и других.

Трудности, связанные с получением длинных временных рядов потребления электроэнергии РФ в региональном разрезе, связаны с изменениями в статистическом учёте (с 2005 г. электробалансы РФ и регионов стали формироваться в соответствии с Общероссийским классификатором видов экономической деятельности; с 2010 г. промышленность перешла с классификатора ОКП на классификатор ОКПЭД и другие), с неточностью сведений, получаемых от субъектов рынка электроэнергетики.

В настоящий момент основные источники данных по потреблению электроэнергии – Системный оператор Единой Энергетической системы и Росстат. Однако существуют определённые различия в этих данных. Системный оператор не учитывает потребление электроэнергии «внесистемной» электроэнергетикой (энергоустановки, не имеющие выход в электросеть «общего пользования»), а Росстат, в свою очередь, учитывает полное потребление электроэнергии в регионе. Динамика электропотребления в регионах по данным Системного оператора ЕЭС более логична и отражает динамику изменений региональной экономики. Но данные Системного оператора не структурированы по видам экономической деятельности. Поэтому на основании данных Росстата были рассчитаны доли видов экономической деятельности в электробалансе регионов, после чего, в силу относительной устойчивости полученных значений долей, был произведён расчёт потребления электроэнергии по ВЭД на основании данных Системного оператора:



$$\begin{aligned} \text{Потребление электроэнергии ВЭД}_{\text{расч}} &= \\ &= \text{Потребление электроэнергии регионом всего}_{\text{Сист.оператор ЭЭС}} \times \\ &\times \text{Доля потребления электроэнергии ВЭД}_{\text{Росстат}} \end{aligned}$$

Для моделирования показателей потребления электроэнергии использовались преимущественно два класса моделей: экстраполяционные и факторные. Согласно проведённому исследованию, приемлемую точность в прогнозировании электропотребления на год вперёд дают экстраполяционные модели. Предпосылкой экстраполяционных моделей является то, что исходный модельный ряд содержит всю необходимую информацию для построения прогноза по выбранному показателю.

Для построения среднесрочных и долгосрочных прогнозов необходимо учитывать влияние внешних факторов, например изменение объёмов промышленного производства. Таким образом, среднесрочное и долгосрочное моделирование опираются на факторные модели.

В свою очередь, построение факторных моделей подразумевает:

- наличие длинных временных рядов моделируемых и объясняющих показателей, на основе которых можно идентифицировать модели;
- наличие значений сценарных показателей на периоде прогнозирования для построения прогнозов моделируемых переменных.

Отличительной особенностью построения факторных моделей в годовой динамике и динамике по месяцам является предварительная очистка данных в динамике по месяцам от сезонных колебаний, которая осуществлялась с помощью процедуры Census II. Данная процедура позволяет разложить исходный ряд на две составляющие: тренд-циклическую и сезонную. В силу изменения амплитуды колебания сезонного фактора и отсутствия относительно постоянной сезонной вариации, применялась мультипликативная модель. При этом исходный ряд рассчитывается с помощью обратных преобразований после моделирования тренд-циклической составляющей временного ряда.

На основе данных Росстата (с 2005 по 2007 гг.) о потреблении электроэнергии различными ВЭД в ежегодной динамике по региону были рассчитаны доли электропотребления по отдельным категориям потребителей: категория-1 (потребление свыше 10 МВт/час); категория-2 (потребление от 670 КВт/час до 10 МВт/час) и категория-3 (потребление менее 670 КВт/час). Доли «пограничных» ВЭД определялись экспертно, согласно наличию на территории региона крупных и средних предприятий соответствующего вида деятельности. После умножения соответствующих долей на потребление электроэнергии были найдены значения потребления электроэнергии для отдельных категорий потребителей в соответствии с выбранной классификацией (см. рисунок 3).

После получения всех модельных значений, был проведён бэк-тестинг с периодом идентификации 1993-2010 гг. В итоге было показано, что количество регионов с существенным отклонением от фактических данных незначительно, а отклонение по всей России не превышает 5%. Таким образом, было принято решение не изменять методику расчёта, кроме того, при расширении обучающей выборки (до 2011 г. включительно) удаётся корректнее учитывать фактор инерционности и погрешности по «критическим» регионам.

В блоке *прогнозирования потерь электроэнергии в сетях* был введён вспомогательный показатель – коэффициент потерь, значения которого моделировались с помощью применения экстраполяционной модели ARIMA  $(p,q,d)$ . При этом итоговая величина потерь определяется по формуле:

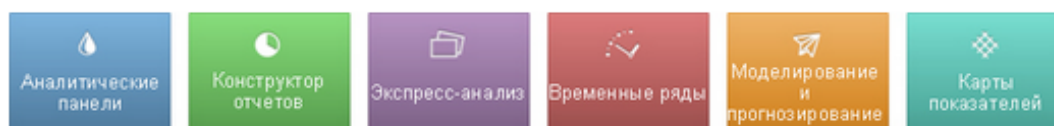
$$\begin{aligned} & \text{Потери электроэнергии в сетях} = \\ & = \text{Коэффициент потерь} \times \\ & \times (\text{Потребление электроэнергии населением} + \text{Потребление электроэнергии ВЭД}) \end{aligned}$$

Помимо расчёта значений абсолютных потерь программный комплекс даёт возможность разделения потерь электроэнергии в сетях на две группы: технологические и коммерческие потери, с учётом соответствующих долей.

Программный комплекс анализа конъюнктуры оптового рынка электроэнергии и мощности позволяет прогнозировать число часов использования мощности для трёх категорий потребителей, а также для городского и сельского населения. Число часов использования мощности – отношение суммарного заявленного годового электропотребления каждой категорией потребителей в региональном разрезе к фактической потребляемой максимальной мощности по каждой категории потребителей в часы максимальных нагрузок по каждому региону.

Программный комплекс региональной модели конъюнктуры оптового рынка электроэнергии и мощности реализован на базе PROGNOZ Platform (рис. 4), созданной ведущим российским разработчиком систем бизнес-анализа компанией «Прогноз». PROGNOZ Platform – интегрированная платформа для создания информационно-аналитических систем и систем поддержки принятия решений, объединяющая современные технологии построения хранилищ данных, средства формирования отчётности и оперативного анализа данных, инструменты моделирования и прогнозирования [7].

#### Инструменты аналитической обработки и предоставления информации



#### Средства разработки и интеграционные компоненты



#### Единая инфраструктура метаданных

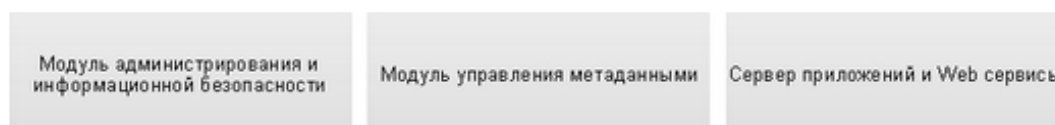


Рисунок 4. Структура PROGNOZ Platform

---

## Заключение

---

Использование подобных моделей и реализованных на их основе программных комплексов позволяет осуществлять прогнозирование развития и функционирования электроэнергетики, что во многом определяет энергетическую безопасность страны в целом. В сочетании со сценарным подходом и использованием современных информационных технологий они представляют эффективный инструментарий для поддержки принятия управленческих решений на государственном уровне в сфере энергетики. Результаты прогнозирования электропотребления могут быть использоваться для формирования рациональной стратегии развития энергетики как страны в целом, так и отдельных её субъектов в частности.

---

## Библиографический список

---

1. Андрианов Д.Л., Науменко Д.О., Старкова Г.С. Анализ методов и моделей энергопотребления на макроуровне // Научно-технические ведомости СПбГПУ. Экономические науки. 2012. № 4. С. 215 - 219.
2. BP: прогноз развития мировой энергетики до 2030 г. [Электронный ресурс]. URL: [http://www.bp.com/liveassets/bp\\_internet/russia/bp\\_russia\\_russian/STAGING/local\\_assets/downloads\\_pdfs/s/bp\\_energy\\_outlook\\_2030\\_rus.pdf](http://www.bp.com/liveassets/bp_internet/russia/bp_russia_russian/STAGING/local_assets/downloads_pdfs/s/bp_energy_outlook_2030_rus.pdf) (дата обращения 25.03.2013).
3. Автоматизированная информационная система «Рынки электроэнергии и мощности» [Электронный ресурс]. URL: [http://www.ais.np-sr.ru/information/IASE\\_0V\\_R0\\_CONSUMPTION#0/0/CONSUMPTION](http://www.ais.np-sr.ru/information/IASE_0V_R0_CONSUMPTION#0/0/CONSUMPTION) (дата обращения 25.03.2013).
4. Официальный сайт Федеральной службы государственной статистики [Электронный ресурс]. URL: <http://www.gks.ru> (дата обращения 25.03.2013).
5. Официальный сайт Системного оператора Единой энергетической системы России [Электронный ресурс]. URL: [http://so-ups.ru/fileadmin/files/company/reports/disclosure/2012/ues\\_rep2011.pdf](http://so-ups.ru/fileadmin/files/company/reports/disclosure/2012/ues_rep2011.pdf) (дата обращения 25.03.2013).
6. Федеральный закон РФ «Об электроэнергетике» от 23.03.2003г. № 35-ФЗ (в редакции от 29.06.2012г.).
7. Официальный сайт компании Прогноз [Электронный ресурс]. URL: [http://www.prognoz.ru/ru/prognoz\\_platform.php](http://www.prognoz.ru/ru/prognoz_platform.php) (дата обращения 25.03.2013).

---

## Сведения об авторах

---



**Галина Старкова** – Пермский государственный национальный исследовательский университет, аспирант, ассистент кафедры информационных систем и математических методов в экономике; Россия, г. Пермь, 614990, ул. Букирева, д. 15; e-mail: [Starkova@prognoz.ru](mailto:Starkova@prognoz.ru).

*Major Fields of Scientific Research: математические и инструментальные методы экономики; информационно-аналитические системы и системы поддержки принятия решений.*

---



---

## TABLE OF CONTENT

<i>Classification of Incomplete Data</i>	
Vasily V.Ryazanov, Vladimir V.Ryazanov .....	203
<i>On a Logical Regularities Based Method of Definite Quality Object Synthesis</i>	
Alexander Dokukin .....	212
<i>Modeling of an Intellectual Problem Solver by Transformation of Semantic Models</i>	
Elena Shalfeeva .....	217
<i>Modeling of the Organizational and Expert Activities of the Scientists at the Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, in 2005-2010</i>	
J. Tabov, S. Hristova .....	223
<i>Methods and Models for Selection of Rational Solutions in Decision-Making Systems</i>	
Oleksandr Kuzomin, Mikhail Shtukin .....	228
<i>Storing Dictionaries and Thesauruses Using NL-Addressing</i>	
Krassimira B. Ivanova, Koen Vanhoof, Krassimir Markov, Vitalii Velychko .....	239
<i>Задача выбора топологической структуры сети доступа</i>	
Галина Гайворонская, Антон Бондаренко .....	252
<i>Качество управления интеллектуальными услугами в сетях последующего поколения</i>	
Сергей Шестопапов .....	262
<i>Повышение структурной живучести телекоммуникационной сети</i>	
Нина Князева .....	275
<i>Разработка модели пользователя информационных сетей как плохоформализованного объекта</i>	
Анастасия Смирнова .....	285
<i>Програмный комплекс анализа конъюнктуры оптового рынка электроэнергии и мощности</i>	
Галина Старкова .....	292
<i>Table of content</i> .....	300