

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Одеський національний політехнічний університет

ІНФОРМАТИКА ТА МАТЕМАТИЧНІ
МЕТОДИ В МОДЕЛЮВАННІ

INFORMATICS AND MATHEMATICAL
METHODS IN SIMULATION

Том 7, № 4

Volume 7, No. 4

Одеса – 2017
Odesa – 2017

Журнал внесений до переліку наукових фахових видань України
(технічні науки)
згідно наказу Міністерства освіти і науки України № 463 від 25.04.2013 р.

Виходить 4 рази на рік

Заснований Одеським національним
політехнічним університетом у 2011 році

Свідоцтво про державну реєстрацію
КВ № 17610 - 6460Р від 04.04.2011р.

Головний редактор: *Г.О. Оборський*

Заступник головного редактора:

А.А. Кобозєва

Відповідальний редактор:

О.Ю. Козлова

Редакційна колегія:

*Т.О. Банах, П.І. Бідюк, Н.Д. Вайсфельд,
А.Ф. Верлань, Г.М. Востров, В.Б. Дудикевич,
М.Б. Копитчук, С.В. Ленков, І.І. Маракова,
С.А. Нестеренко, М.С. Никитченко,
С.А. Положаєнко, О.В. Рибальський,
Х.М.М. Рубіо, В.Д. Русов,
І.М. Ткаченко-Горський, А.В. Усов,
С.В. Філіппова, В.О. Хорошко, М.Є. Шелест,
М.С. Яджак*

Published 4 times a year

Founded by Odessa National Polytechnic
University in 2011

Certificate of State Registration

КВ № 17610 - 6460Р of 04.04.2011

Editor-in-chief: *G.A. Oborsky*

Associate editor:

A.A. Kobozeva

Executive editor:

O.Yu. Kozlova

Editorial Board:

*T. Banakh, P. Bidyuk, V. Dudykevich,
V. Khoroshko, N. Kopytchuk,
S. Lenkov, I. Marakova, S. Nesterenko,
N. Nikitchenko, S. Philippova, S. Polozhaenko,
J. Rubio, V. Rusov, O. Rybalsky, M. Shelest,
I. Tkachenko Gorski, A. Usov, N. Vaysfeld,
A. Verlan, G. Vostrov, M. Yadzhak*

Друкується за рішенням редакційної колегії та Вченої ради Одеського національного
політехнічного університету

Оригінал-макет виготовлено редакцією журналу

Адреса редакції: просп. Шевченка, 1, Одеса, 65044, Україна

Телефон: +38 048 705 8506

Web: <http://immm.opu.ua>

E-mail: immm.ukraine@gmail.com

Editorial address: 1 Shevchenko Ave., Odessa, 65044, Ukraine

Tel.: +38 048 705 8506

Web: <http://immm.opu.ua>

E-mail: immm.ukraine@gmail.com

© Одеський національний політехнічний університет, 2017

ЗМІСТ / CONTENTS

РОЗРОБКА СТЕГANOГРАФІЧНОГО МЕТОДУ ДЛЯ БІНАРНИХ КОНТЕЙНЕРІВ-ЗОБРАЖЕНЬ А.А. Кобозева, М.В. Ворнікова, Г.В. Ахмаметьєва, О.А. Пенко	255	DEVELOPMENT OF THE STEGANOGRAPHIC METHOD FOR BINARY COVER-IMAGES Kobozeva A., Vornikova M., Akhmametieva A., Penko O.
ІНФОРМАТИЗАЦІЯ ПРОТИАВАРІЙНОГО КЕРУВАННЯ СКЛАДНИМИ ТЕХНІЧНИМИ СИСТЕМАМИ С.М. Коновалов, В.В. Вичужанін	265	INFORMATIZATION OF EMERGENCY CONTROL OF COMPLEX TECHNICAL SYSTEMS Konovalov S., Vichuzhanin V.
МЕТОД ВІДОКРЕМЛЕННЯ КЛОНУ ВІД ПРООБРАЗУ В ЦИФРОВОМУ ЗОБРАЖЕНІ В УМОВАХ ВІДСУТНОСТІ ВІДМІННОСТЕЙ ПРИ ЇХ ПОСТОБРОБЦІ І.І. Бобок	276	METHOD FOR SEPARATING OF CLONE AND PROTOTYPE IN A DIGITAL IMAGE IN THE CONDITIONS OF THE ABSENCE OF DIFFERENCES IN THEIR POST-PROCESSING Bobok I.
СПРОЩЕННЯ МАТЕМАТИЧНИХ МОДЕЛЕЙ ОБ'ЄКТІВ З РОЗПОДІЛЕНИМИ ПАРАМЕТРАМИ НА ОСНОВІ МЕТОДУ РОЗМІЩЕННЯ А.А. Верлань, В.А. Іванюк	285	SIMPLIFICATION OF MATHEMATICAL MODELS OF OBJECTS WITH DISTRIBUTED PARAMETERS BY THE METHOD OF SPLITTING Verlan A., Ivanyuk V.
ПАРАМЕТРИЧНИЙ СИНТЕЗ ДИСКРЕТНО-АДАПТИВНОЇ ОДНОКАНАЛЬНОЇ З ДОПЛЕРОВСЬКОЇ ЧАСТОТИ СИСТЕМИ ОБРОБКИ СИГНАЛУ В УМОВАХ ГАУССОВИХ ЗАВАД І.В. Цевух, А.В. Соколов, А.А. Сакович	291	PARAMETRIC SYNTHESIS OF ONE-CHANNEL ON DOPPLER FREQUENCY OF DISCRETE- ADAPRIVE PROCESSING SYSTEM FOR SIGNAL IN GAUSSIAN NOISES Tsevkukh I., Sokolov A., Sakovich A.
МОДЕЛЮВАННЯ ТА ОПТИМІЗАЦІЯ ДОСТУПУ ДО СТОРІНОК WEB-САЙТУ ДЛЯ РІЗНИХ ЗАКОНІВ РОЗПОДІЛУ ЙМОВІРНОСТЕЙ ЗВЕРТАННЯ ДО СТОРІНОК М.І. Юськів, Г.Г. Цегелик	300	MODELING AND OPTIMIZATION OF ACCESS TO THE PAGES OF WEB-SITE FOR DIFFERENT DISTRIBUTION LAWS OF PROBABILITY OF ACCESSING TO THE PAGE Yuskiv M., Tsegelik G.

АНАЛІЗ МЕТОДІВ
КОНСЕНСУСНОГО АГРЕГУВАННЯ
РАНГОВИХ ПЕРЕВАГ

В.О. Болтьонков, В.І. Куваєва,
О.В. Позняк

307

ANALYSIS OF MEDIAN METHODS
FOR CONSENSUS RANK
PREFERENCES AGGREGATION
Boltenkov V., Kuvayeva V., Pozniak A.

ОПТИМІЗАЦІЯ ВІДБОРУ ТА
АНАЛІЗУ ІНФОРМАЦІЇ В
РІЗНОСТРУКТУРНИХ СХОВИЩАХ
ДАНИХ

Д.С. Шибасєв, В.В. Вичужанін,
Н.О. Шибасєва, Н.Д. Рудніченко

318

OPTIMIZATION OF SELECTION AND
ANALYSIS OF INFORMATION IN
DATA WAREHOUSES
RESTRUCTURING
Shibaiev D., Vichuzhanin V.,
Shibaieva N., Rudnichenko N.

АЛГОРИТМ ФУНКЦІОНУВАННЯ
СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ
РІШЕНЬ В ЗЕМЕЛЬНИХ
ВІДНОСИНАХ

М.А. Кухар

325

ALGORITHM OF THE DECISION
SUPPORT SYSTEM IN LAND
RELATIONS
Kukhar M.

МЕТОДИ І АЛГОРИТМИ ПОКРИТТЯ
(Частина 2)

О.М. Паулін

333

COVERING METHODS AND
ALGORITHMS (Part 2)
Paulin O.

МЕТОД РОЗПІЗНАВАННЯ
ОПТИЧНИХ ЗОБРАЖЕНЬ У
ХЕММІНГОВОМУ ПРОСТОРІ НА
ОСНОВІ КРИТЕРІЮ СТРУКТУРНОЇ
СКЛАДНОСТІ

Н.Я. Возна

339

METHOD FOR RECOGNITION OF
OPTICAL IMAGES IN THE HAMMING
SPACE BASED ON THE STRUCTURAL
COMPLEXITY CRITERION
Vozna N.

DEVELOPMENT OF THE STEGANOGRAPHIC METHOD FOR BINARY COVER-IMAGES

A.A. Kobozeva¹, M.V. Vornikova¹, A.V. Akhmametieva¹, O.A. Penko²

¹ Odesa National Polytechnic University,
1, Shevchenko Ave., Odesa, 65044, Ukraine; e-mail: alla_kobozeva@ukr.net

² Richelieu Lyceum,
5, Elisavetinska Str., Odesa, 65026, Ukraine

A new steganographic method for binary digital images has been developed. During the development of the method, the expediency of choosing the spatial domain of the container for introducing hiding information is justified, which avoids additional accumulation of computational error and increase in computational complexity, as compared with methods working in the field of container transformation. The developed method provides for the decoding of hiding information, its restoration in full, the increase in the efficiency of steganographic transformation of binary digital images covers, quantitatively estimated by the peak signal-to-noise ratio (PSNR) for the generated stego, in comparison with existing analogues for comparable volumes of submerged information, than by 30%.

Keywords: binary digital image, steganographic method, spatial domain, reliability of perception of stego

Introduction

The high level of development of information technologies and computer sciences has both positive and negative consequences today one of which is the unresolved adequately task of information security of widespread and everywhere used digital media [1, 2]. Traditional ways for ensuring the protection of data and resources from the possibility of illegal access, for ensuring the reliability of received data and messages are: the encryption of confidential information that is realized by means of cryptography methods and algorithms [3], the concealment of the fact of existence of confidential information by using steganographic methods [4, 5]. The combination of the two approaches when the encryption of the concealed information precedes its concealment in an inconspicuous information content-container, in particular in a digital image, followed by transmission over a hidden (steganographic) communication channel increases the security of the transmitted information, and therefore it is, as a rule, used to ensure of preventing of illegal access now.

In the present work binary images are used as a container, a binary sequence formed by a random (pseudorandom) way and considered as a result of preliminary encoding (encryption) of really transferred message is used as a confidential information. Binary digital images not being widespread in everyday life have important applied significance: various schemes, graphs, tables, some medical images are presented in the form of binary images in order to ensure their high sharpness, contrast [6, 7]. Steganographic transformation of such digital image is the difficult and unresolved task related to their specificity – they have only two gradations of brightness. The change in the brightness of a pixel by one unit can be seen visually in the presence of the original image (fig. 1). And although the developments in this direction are conducted by modern scientists [8-11] the task of ensuring the preservation of the reliability of the perception of stego that is the result of embedding of additional information into binary image-container remains actual. This is due to the fact that when

organizing a steganographic channel, one of its required properties should be to provide a significant capacity of the hidden communication channel, and the proposed existing developments, as a rule, provides an acceptable quality of stego (reliability of perception), an increase in this quality that quantitatively evaluated by various difference indices, in particular, the peak signal-to-noise ratio (PSNR) [4], by reducing the amount of transmitted information [12-15] which is extremely undesirable.



Fig. 1. The result of embedding of additional information into binary digital image: a – digital image-container; b – digital image-stego

The aim of the article and setting of research problems

The *aim* of the work is to increase in efficiency of steganographic transformation of binary digital images by developing a new steganographic method that ensures the reliability of perception of the created stego.

The efficiency of a steganographic transformation in the context of this work will be quantitatively estimated by PSNR value of the stego formed by the corresponding steganographic method taking into account the provision of the required capacity of the hidden communication channel.

Taking into account the fact that all steganographic methods are divided into those that embeds additional information into the spatial domain of the digital image and those that use different transformation domains for steganographic transformation, and also taking into account the features of the formal representation of the binary digital image (the presence of only two brightness gradations) for achievement of the aim it is necessary to solve the following *problems*:

1. To define an area for embedding of additional information into a binary digital image (spatial, transform);
2. To determine the formal parameters of the image in the selected area which are the least sensitive to the perturbing influence consisting in embedding of additional information.

Main part

For embedding of additional information into the image-container, both the spatial and the transform domain can be used, which can be the frequency domain of the digital image, domain of various matrix expansions: singular, spectral, etc. [16]. In [17] it was shown that in order to provide certain required properties of stego, including the reliability of its perception, it does not matter in principle into which domain of the container the embedding of additional

information occurs, but it is important to which perturbations of singular numbers and singular vectors of the matrix (blocks of the matrix) of a container it will lead, where exactly within the formal parameters of singular expansion of a matrix (blocks of the matrix) these perturbations will be localized. At the same time, in case of the organization of steganographic transformation into transform domain of the digital image additional computational expenses and the accumulation of additional computational errors arise due to the execution of "transitions" from the spatial to the transform domain, and then in the reverse order [16].

For binary digital images the selection of the transform domain for embedding of additional information is connected to additional difficulties arising owing to rounding of the brightness values of the pixels when inverse transformation carrying out (return to the spatial domain of the image). If brightness change in the field of real numbers is very insignificant but it is more than 0.5 it can lead to inverse of color: 0 will be replaced by 1 or 1 will be replaced by 0. The illustration to told is given in fig. 2 in case of the discrete cosine transform (DCT) of one 8×8 block of the digital image shown in Fig. 1, a. Minor changes in the matrix of DCT coefficients (the spectral norm of the perturbation matrix was only 2.3 here) led to visually discernable changes in the spatial domain of the image (fig. 2).

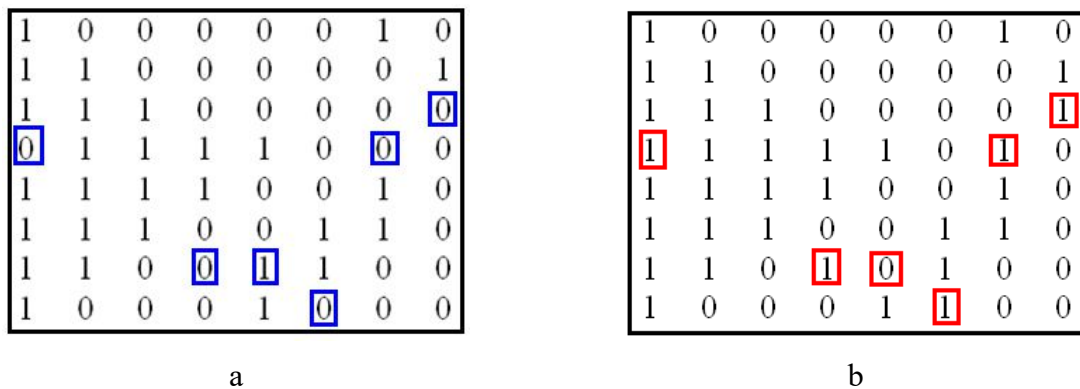


Fig. 2. The result of distortion of the binary digital image due to insignificant perturbations of the DCT coefficients: a – the block of the original binary image; b – the corresponding block of the binary digital image whose matrix norm is 2.3, after a perturbation carried out in the DCT domain

Thus, the spatial domain of a binary digital image-container is preferable in comparison with the transform domain of the container for embedding of additional information both from the point of view of computational complexity estimation, both from the point of view of estimation of computational error, and therefore it is used to organize steganographic transformation in the method being developed in the work.

Traditionally [5] to reduce the visual distortion of a color or grayscale image when embedding of additional information, a steganographic path is used which contains pixels belonging to the contours (significant differences in brightness values) presented in the image. The specificity of a binary digital image does not make this embedding method such which guarantees the reliability of the perception of a stego because here due to steganographic transformation a contour shift and a change in its shape can occur which will be visually obviously visible even without the presence of the original image-container.

For practical confirmation of the last statement in the Matlab environment a computational experiment was carried out in which 400 binary images of size 1000×1500 were used, forming a set which is called experimental below. During the experiment embedding of additional information into the pre-selected contour of digital images additively modulo 2 was carried out. By subjective ranging violation of reliability of perception of some stegos (the example is given in fig. 3) was revealed. The mean value of PSNR for all received

stegos was only 17.68 dB that is much lower than the value considered acceptable for the organization of the steganographic communication channel – 37 dB [4].

It should be noted that the use of contour pixels as a steganographic path can significantly limit the capacity of a hidden communication channel that is extremely unwanted, and also increases the probability of detection the presence of embedded additional information by means of steganalysis, obviously localizing within the boundaries of the digital image its location area expected on the basis of practical experience.

All of the above makes it reasonable to choose as steganographic path those pixels that are not directly related to the scene of the image.



Fig. 3. Visual distortions when embedding of additional information into the contour of the container: a – digital image-container; b – digital image-stego

As it appears from above the use of binary digital images as containers cannot provide a significant capacity of a hidden communication channel in principle. Taking this into account, as well as the fact that during the organization of steganographic transformation for reduction of computational complexity of the corresponding steganographic algorithms, the digital image-container matrix is usually broken up into not crossed blocks, and the embedding of a certain fixed bit count of additional information occurs block-by-block, it is proposed in the developed method after the preliminary standard partitioning of the matrix on 8×8 -blocks to embed additional information into the block no more than one bit of information.

By subjective ranking during the computational experiment it was found that the change of even one pixel in a monochrome block, i.e. block, the matrix of which contained only zeros or only ones, does not preserve the reliability of image perception therefore further such blocks do not participate in the process of steganographic transformation.

For localization in 8×8 -block of the pixel used for steganographic transformation which perturbation least of all will visually affect on digital image-stego the computational experiment was performed on the basis of digital images from the experimental set. During the experiment the original image-container was broken into 8×8 -blocks, into all blocks of the image, except those that consisted completely of zeros or ones, information was embedded by inverse of brightness value of the pixel located at position (1,1). For each digital image-stego the PSNR value was defined. Similar actions were carried out for accounting of all possible options for selecting a pixel within the block to embed the information bit: the pixel at position (1,2), (1,3), ..., (8,1), ..., (8,8). As a result of the experiment, it was found that when embedding information in the pixels of the category 1 of the block (fig. 4), the image quality taking into account PSNR, where the average value was 26 dB, was significantly better than in the case of pixels for the categories 2 and 3: the average value of PSNR here was 21 dB and 17 dB, respectively.

1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	→ Category 3
2,1	2,2	2,3	2,4	2,5	2,6	2,7	2,8	
3,1	3,2	3,3	3,4	3,5	3,6	3,7	3,8	→ Category 2
4,1	4,2	4,3	4,4	4,5	4,6	4,7	4,8	
5,1	5,2	5,3	5,4	5,5	5,6	5,7	5,8	→ Category 1
6,1	6,2	6,3	6,4	6,5	6,6	6,7	6,8	
7,1	7,2	7,3	7,4	7,5	7,6	7,7	7,8	
8,1	8,2	8,3	8,4	8,5	8,6	8,7	8,8	

Fig. 4. The categories of pixels of 8×8 -blocks of binary digital images used for embedding of additional information

During the experiment described above it was found that visually changes in digital image as a result of embedding of additional information occur not only when it embeds into a monochrome block pixel, but can also occur when the block contains a small quantity of ones/zeros, and the embedding is carried out by inversion of the zero/one presented there. For accounting of it in the developed steganographic method two parameters p_0 and p_1 are introduced which values allow controlling the process of selecting the next block for embedding the next bit of the transmitted information. Each of parameters accepts values from 1 to 63, herewith p_0 sets minimum possible quantity of zeros, and p_1 - minimum possible quantity of ones in the block used in the process of a steganographic transformation, i.e. if the number of zeros/ones in the block appears less than p_0 / p_1 such block is not used for embedding of additional information. Thus, the more there will be values p_0, p_1 , the less there will be a capacity of the organized hidden communication channel as with increase of p_0, p_1 the number of blocks which are not used in the process of a steganographic transformation will grow, but at the same time the probability of ensuring the reliability of perception of the created stego will increase due to increasing PSNR value.

It should be noted that in the process of a steganographic transformation of the block it is possible when the block after embedding of bit of additional information (therefore values of brightness of pixels/pixel will change) will be already such for which the conditions imposed by parameter values p_0, p_1 on the number of zero and ones, may not be executed. I.e. when reviewing such block in a stego the quantity of ones/zeros in it can appear less than p_0 / p_1 that will make more difficult to recognize the block participating in the process of steganographic transformation of the digital image when decoding additional information and will demand the organization of additional accounting of blocks containing in themselves bits of the transmitted data.

The embedding of the additional information bit into the 8×8 -block of the binary digital image in the developed method depends on the parity/oddness of the sum of all the elements of the block, i.e. the quantity of ones in the block: the additional information, which embeds in the block bit and is equal to 1/0 must correspond to the odd/even sum of the elements values of the block.

Let F is a $n \times m$ -matrix of a binary digital image-container, $d_1, d_2, \dots, d_t, d_i \in \{0,1\}, i = \overline{1,t}$ - the binary sequence which is result of encoding of the transferred confidential message, $\bar{F}$ - $n \times m$ -matrix of a binary digital image-stego. For accounting of the blocks participating in the steganographic transformation in the developed method a binary $\left[\frac{n}{8} \right] \times \left[\frac{m}{8} \right]$ - matrix T with elements t_{ij} each of which corresponds to the corresponding block of digital image is created. The matrix T formed during the embedding

of additional information is a part of the secret key for the developed method in case of its algorithmic implementation and is transmitted over a secure communication channel.

The main steps of the proposed steganographic method further called by SMBI look as follows.

Embedding of additional information

Step 1. Split the matrix F of digital image-container in a standard way into uncrossed 8×8 -

blocks $B^{(kl)}$, $k = 1, \left\lceil \frac{n}{8} \right\rceil, l = 1, \left\lceil \frac{m}{8} \right\rceil$, where $\lceil \bullet \rceil$ is the integer part of the argument. Form the original form of the $\left\lceil \frac{n}{8} \right\rceil \times \left\lceil \frac{m}{8} \right\rceil$ -matrix T , which is part of the secret key:

$$T = 0.$$

Step 2. Set the values of the parameters $p_0, p_1 \in \{1, 2, 3, \dots, 63\}$ (this can be done taking into account the required size of the transmitted confidential information, the minimum/maximum permissible distortion requirement of the digital image while steganographic transformation, the totality of these requirements, etc.).

Step 3. For the next block $B^{(kl)}$ with elements $b_{ij}^{(kl)}$, $i, j = \overline{1, 8}$, of the matrix F :

3.1. (*Conclusion about the suitability of the block for steganographic transformation*).

Find

$$s_1 = \sum_{i,j=1}^{64} b_{ij}^{(kl)}; s_0 = 64 - s_1.$$

If

$$(s_1 \geq p_1) \& (s_0 \geq p_0),$$

then

the block $B^{(kl)}$ is used for embedding the next bit of additional information;

$$t_{kl} = 1,$$

else

the block $B^{(kl)}$ does not used for embedding the next bit of additional information.

3.2. (*Embedding the next bit of additional information into the container block*).

If

the block $B^{(kl)}$ is used for steganographic transformation and d_i is the next bit of additional information,

then

if

$$(s_1 = 2n) \& (d_i = 1) \vee (s_1 = 2n - 1) \& (d_i = 0), \text{ where } n \text{ is a natural number,}$$

then

$$b_{5,5}^{(kl)} = |b_{5,5}^{(kl)} - 1|.$$

3.3. Go to the next block.

Decoding of additional information

Step 1. Split the matrix F of digital image-stego in a standard way into uncrossed 8×8 -

blocks $\bar{B}^{(kl)}$, $k = 1, \left\lceil \frac{n}{8} \right\rceil, l = 1, \left\lceil \frac{m}{8} \right\rceil$.

Step 2. For the next block $\bar{B}^{(kl)}$ with elements $\bar{b}_{ij}^{(kl)}$, $i, j = \overline{1, 8}$, of the matrix $\bar{F}$:

2.1. (*Conclusion about the presence of additional information bit in the block*).

If

the corresponding to the block $\bar{B}^{(kl)}$ element of the secret key matrix T

$$t_{kl} = 1,$$

then

block $\bar{B}^{(kl)}$ contains additional information,

else

block $\bar{B}^{(kl)}$ does not contain additional information.

2.2. (Extraction of the next bit of additional information from the stego block).

If

block $\bar{B}^{(kl)}$ contains additional information,

then

extract $\bar{d}_i$ - the next bit of additional information.

Find

$$\bar{s}_1 = \sum_{i,j=1}^{64} \bar{b}_{ij}^{(kl)} ;$$

If

$\bar{s}_1 = 2n$, where n is a natural number,

then

$$\bar{d}_i = 0,$$

else

$$\bar{d}_i = 1.$$

2.3. Go to the next block.

When testing of the proposed method when carrying out a computing experiment in the environment of Matlab its algorithmic implementations with different parameter values p_0 , p_1 were used. The results of the experiment in which digital images from the experimental set were used, characterized by the PSNR value for the created stego-images and the length of the embedded message (in bits) depending on the values of the parameters p_0 , p_1 are shown in Table 1, 2, respectively. The numbers presented in tables correspond to the mean value of the corresponding parameter for all digital images from the experimental set at fixed values p_0 , p_1 .

Since the stego was saved without loss when decoding additional information all the bits of the embedded message were restored correctly.

Table 1.

PSNR value (dB) for stego generated by algorithmic implementations of the SMBI method for various parameter values p_0 , p_1

$p_0 \backslash p_1$	1	5	10	20	30
1	31.71	32.13	32.36	35.86	33.54
5	32.83	33.16	33.50	34.16	35.14
10	33.78	34.41	34.57	35.44	36.52
20	33.38	35.93	36.50	37.79	39.89
30	34.34	37.73	38.59	40.79	46.98

Table 2.

Length of the embedded message (in bits) in stego generated by algorithmic implementations of the SMBI method for various parameter values p_0 , p_1

$p_0 \backslash p_1$	1	5	10	20	30
1	8815	8310	7833	6969	6940
5	7151	6645	6169	5305	4376
10	5914	5409	4933	4068	3139
20	4274	3769	3292	2428	1499
30	3089	2583	2107	1243	314

From these results it is visible that if both parameters have values not less than 10 then PSNR values are that that provide reliability of perception of a stego [4]. The length of the embedded message naturally decreases with increasing values p_0 , p_1 remaining considerable for a binary digital image-container when at least one of the parameters during steganographic transformation is less than 30.

For algorithmic implementation of the developed method (in case of $p_0 = 5$, $p_1 = 5$) the comparative analysis of its efficiency (from the point of view of ensuring the reliability of perception of the created stego) with the modern analogs [1, 12-15] was carried out. The values of the parameters p_0 , p_1 were chosen in such a way as to ensure the correctness of the comparison, i.e. analogous to the conditions of testing given in [1, 12-15] (steganographic transformation occurred due to the change of slightly more than 300 pixels in binary digital images-containers which were used in [1, 12-15]). The results of the experiment speaking about increase in efficiency of a steganographic transformation of binary digital images from the point of view of increase in PSNR value, other things being equal, are given in Table 3.

Table 3.

PSNR values (dB) of the developed method and existing analogs

Steganographic method	SMBI (2017)	Do Van Tuan (2012)	CTL (2005)	TCP (2000)	Ki-Hyun Jung (2009)	Venkatesan et al.'s (2007)
Stego						
Baboon	27.49	-	-	-	16.32	16.36
Aircraft	32.09	-	-	-	21.50	21.25
Lena	29.92	22.90	22.16	22.88	20.51	20.45
Boat	30.76	-	-	-	20.45	20.56

As can be seen from Tabl. 3 we managed to increase the PSNR value from 30.7% (for digital image “Lena” in comparison with the DoVanTuan method (2012)) up to 68% for digital image “Baboon” in comparison with Venkatesan et al.'s (2007) in case of comparable amounts of the embedded message.

Conclusions

The result of this work is an increase in the efficiency of steganographic transformation of binary digital images-containers by developing a new steganographic method SMBI which provides an increase in the difference in PSNR for estimating the reliability of perception of

the created stego in comparison with existing analogs for comparable amounts of embedded information by a minimum of more than 30%.

During development of the SMBI method expediency of a choice of spatial domain of binary digital images for embedding of additional information is justified that allows to avoid additional accumulation of a computing error and increase in computing complexity in comparison with the methods working in the transformation domain of a container.

As a result of the computational experiment, it was set that at the maximum amount of embedded information ($p_0 = 1$, $p_1 = 1$), the average value of the PSNR for the digital images from the experimental set was 32 dB. When decoding additional information was restored correctly in full.

The capacity of the organized hidden communication channel is still insufficient when working with binary digital images, search of methods to increase of it is an object of research by the authors at the present time.

References

1. Chang, Chin-Chen. Hiding Data in Binary Images / Chin-Chen Chang, Chun-Sen Tseng, Chia-Chen Lin. // ISPEC 2005: Information Security Practice and Experience. — 2005. — Pp. 338-349.
2. Wu, M.Y. A Novel Data Embedding Method for Two-color Facsimile Images / J.H. Lee, M.Y. Wu. // Proc. Int. Symp. On Multimedia Information Processing. — Chung-Li (Taiwan), 1998. — Pp. 138-149.
3. Баричев, С. Г. Основы современной криптографии. — 3-е изд. / С.Г. Баричев, В.В. Гончаров, Р.Е. Серов. — М.: Диалог-МИФИ, 2011. — 176 с.
4. Конахович, Г.Ф. Компьютерная стеганография. Теория и практика / Г.Ф. Конахович, А. Ю. Пузыренко. — К.: МК—Пресс, 2006. — 288 с.
5. Грибунин, В.Г. Цифровая стеганография / В.Г. Грибунин, И.Н. Оков, И.В. Туринцев. — М.: Солон-Пресс, 2002. — 272 с.
6. Федоров, А. Бинаризация черно-белых изображений: состояния и перспективы развития [Электронный ресурс] / А. Федоров. // CLAIM – научно-образовательный кластер. Режим доступа: <http://it-claim.ru/Library/Books/ITS/wwwbook/ist4b/its4/fyodorov.htm/> (Дата обращения 08.08.2017).
7. Фурман, Я.А. Цифровые методы обработки и распознавания бинарных изображений / Я.А. Фурман, А.Н. Юрьев, В.В. Яншин. — Красноярск: Краснояр. ун-та, 1992. — 248 с.
8. Guo, Fu Gui. A New Asymmetric Watermarking Scheme for Copyright Protection / Gui Fu Guo, Ling Ge Jiang, Chen He. // IECE Trans. Fundamentals. — 2006. — Vol. E89-A, No. 2. — Pp. 147-153.
9. Lee, Y.K. Capacity Image Steganographic Model / Y.K. Lee, L.H. Chen. // Proc. of IEEE International Conference on Vision, Image and Signal Processing. — 2000. — Vol. 147, No. 3. — Pp. 288-294.
10. Smitha, B. Spatial Domain – High Capacity Data Hiding in ROI Images / B. Smitha, K.A. Navas. // IEEE – ICSCN 2007. — Chennai, India, Feb, 2007. — Pp. 528-533.
11. Tzeng, C.H. Adaptive Data Hiding in Palette Image by Color Ordering and Mapping with Security Protection / C.H. Tzeng, Z.F. Yang, W.H. Tsai. // IEEE Transactions on Communications. — 2004. — Vol. 52, No. 5. — Pp. 791-800.
12. Tseng, Y.C. A secure Data Hiding Scheme for Binary Images / Y.C. Tseng, Y.Y. Chen, K.H. Pan. // IEEE Transactions on Communications. — 2002. — Vol. 50, No. 8. — Pp. 1227-1231.
13. Do Van Tuan. A Novel Data Hiding Scheme for Binary Images / Do Van Tuan, Tran Dang Hien, Pham Van At. // International Journal of Computer Science and Information Security. — 2012. — Vol. 10, No. 8. — Pp. 33-38.
14. Venkatesan, M. A new data hiding scheme with quality control for binary images using block parity / M. Venkatesan, P. Meenakshidevi, K. Duraiswamy, K. Thiagarajah. // 3rd Inter. Symposium on Information Assurance and Security. — 2007. — Pp. 468-471.
15. Jung, Ki-Hyun. Data Hiding Method with Quality Control for Binary Images / Ki-Hyun Jung, Kee-Young Yoo. // Journal of Software Engineering & Applications. — 2009. — Vol. 2. — Pp. 20-24.

16. Костырка, О.В. Анализ преимуществ пространственной области цифрового изображения-контейнера для стеганопреобразования / О.В. Костырка. // Информатика та математичні методи в моделюванні. — 2013. — Т. 3, № 3. — С. 275–282.
17. Кобозева, А.А. Аналіз захищеності інформаційних систем [Текст]: підруч. для студ. вищ. навч. закл., які навч. за напр. «Інформаційна безпека» та «Системні науки та кібернетика» / А.А. Кобозева, І.О. Мачалін, В.О. Хорошко. — К.: ДУІКТ, 2010. — 316 с.

РОЗРОБКА СТЕГАНОГРАФІЧНОГО МЕТОДУ ДЛЯ БІНАРНИХ КОНТЕЙНЕРІВ-ЗОБРАЖЕНЬ

А.А. Кобозєва¹, М.В. Ворнікова¹, Г.В. Ахматєєва¹, О.А. Пенко²

¹ Одеський національний політехнічний університет,
просп. Шевченка, 1, Одеса, 65044, Україна; e-mail: alla_kobozeva@ukr.net

² Рішельєвський лицей,
вул. Єлісаветинська, 5, Одеса, 65026, Україна

Розроблено новий стеганографічний метод для бінарних цифрових зображень (ЦЗ). В ході розробки методу обґрунтовано доцільність вибору просторової області контейнера для вбудови додаткової інформації, що дозволяє уникнути додаткового накопичення обчислювальної похибки і збільшення обчислювальної складності, в порівнянні з методами, які працюють в області перетворення контейнера. Розроблений метод забезпечує при декодуванні додаткової інформації її відновлення в повному обсязі, підвищення ефективності стеганоперетворення бінарних ЦЗ-контейнерів, яка кількісно оцінюється піковим відношенням «сигнал-шум» (PSNR) для формованих стеганоповідомлень, в порівнянні з існуючими аналогами при порівнянних обсягах вбудованої інформації, мінімально більш, ніж на 30%.

Ключові слова: бінарне цифрове зображення, стеганографічний метод, просторова область, надійність сприйняття стеганоповідомлення

РАЗРАБОТКА СТЕГАНОГРАФИЧЕСКОГО МЕТОДА ДЛЯ БИНАРНЫХ ИЗОБРАЖЕНИЙ-КОНТЕЙНЕРОВ

А.А. Кобозева¹, М.В. Ворникова¹, А.В. Ахметьева¹, Е.А. Пенко²

¹ Одесский национальный политехнический университет,
просп. Шевченко, 1, Одесса, 65044, Украина; e-mail: alla_kobozeva@ukr.net

² Ришельевский лицей,
ул. Елисаветинска, 5, Одесса, 65026, Украина

Разработан новый стеганографический метод для бинарных цифровых изображений (ЦИ). В ходе разработки метода обоснована целесообразность выбора пространственной области контейнера для внедрения дополнительной информации, что позволяет избежать дополнительного накопления вычислительной погрешности и увеличения вычислительной сложности, по сравнению с методами, работающими в области преобразования контейнера. Разработанный метод обеспечивает при декодировании дополнительной информации ее восстановление в полном объеме, повышение эффективности стеганопреобразования бинарных ЦИ-контейнеров, количественно оцениваемой пиковым отношением «сигнал-шум» (PSNR) для формируемых стеганосообщений, по сравнению с существующими аналогами при сравнимых объемах погруженной информации, минимально более, чем на 30%.

Ключевые слова: бинарное цифровое изображение, стеганографический метод, пространственная область, надежность восприятия стеганосообщения

ИНФОРМАТИЗАЦИЯ ПРОТИВОАВАРИЙНОГО УПРАВЛЕНИЯ СЛОЖНЫМИ ТЕХНИЧЕСКИМИ СИСТЕМАМИ

С.Н. Коновалов, В.В. Вычужанин

Одесский национальный морской университет,
ул. Мечникова, 34, Одесса, 65029, Украина; e-mail: konserge-serge@rambler.ru

В статье предложена информатизация задач метода противоаварийного управления сложными техническими системами на основе гибридной экспертной системы. Исследованы условия и критерии работоспособности элементов сложных технических систем. Разработанный метод противоаварийного управления основывается на использовании результатов диагностики и прогнозирования технического состояния систем. В качестве объекта исследования рассмотрена судовая система кондиционирования воздуха.

Ключевые слова: противоаварийное управление, сложная техническая система, гибридная экспертная система, нейронная сеть, нечёткая логика, база знаний, диагностика, прогнозирование

Введение

Эксплуатация сложных технических систем (СТС) связана с возникновением аварийных ситуаций, приводящих к потери работоспособности систем, а также к опасностям для жизни людей [1]. Одним из путей решения такой проблемы является использование систем противоаварийного управления (ПАУ) СТС.

ПАУ предназначено для предотвращения возникновения, развития аварий в СТС, их локализации [2,3]. Наиболее перспективная система ПАУ строится на основе гибридных интеллектуальных систем (ГИС) [4], разновидностью которых является гибридная экспертная система (ГЭС). Подобные системы объединяют внутри одного инструментального средства различные модели представления знаний и механизмы функционирования систем [5-7]. В ГЭС предъявляются требования к форме представления знаний, которые разнообразны (фреймы, семантические сети, базы данных (БД), базы знаний (БЗ), нейросети, нечёткая логика, генетические алгоритмы). Поэтому в ГЭС сложно объединить знания даже в рамках единого информационного пространства [8]. Кроме того, традиционные ГЭС имеют ряд недостатков [4]:

- трудность и неестественность реализации определённых условий для средств автоматизации и телемеханики;
- трудности в условиях неопределённости, нехватки знаний.

Уменьшить влияние недостатков при ПАУ СТС на основе ГЭС можно при помощи использования ГЭС, включающей нейронные сети (НС) и нечёткую логику (НЛ).

Таким образом, разработка метода ПАУ СТС и его информатизация на основе гибридной экспертной системы является актуальной задачей.

Целью работы является повышение надёжности СТС.

Для достижения поставленной цели необходимо решить следующие задачи:

- разработать и исследовать ГЭС;
- разработать метод ПАУ СТС;
- исследовать ГЭС ПАУ СТС.

Основная часть

Для ПАУ разработана структурная схема гибридной экспертной системы ГЭС объединённого типа на основе НС и НЛ (рис. 1).

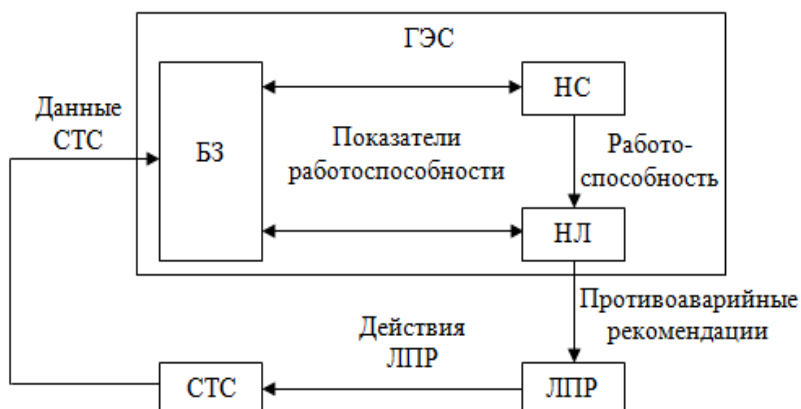


Рис. 1. Структурная схема ПАУ СТС

Информация из СТС поступает в БЗ, где структурируется и распределяется в виде показателей работоспособности (факторы и критерии). При наличии комбинации показателей работоспособности в БЗ имеются данные работоспособности по этим показателям и противоаварийные рекомендации для лица, принимающего решения (ЛПР). Данные поступают через блок НЛ к ЛПР. Если комбинация показателей работоспособности отсутствует в БЗ, то по полученным из БЗ факторам и критериям работоспособности, передаваемым в блок НС, находятся работоспособности отдельных элементов СТС. По значениям работоспособностей элементов СТС вычисляется общая работоспособность СТС. Между блоками НС и БЗ имеется обратная связь для передачи результатов расчётов по показателям работоспособности, не содержащихся в БЗ. Они поступают вместе с данными работоспособности элементов СТС, вычисляемых в НС. В НС прогнозируются состояния СТС, исходя из данных, поступивших в НС в различные временные интервалы. Результаты всех расчётов вместе с данными из БЗ и НС поступают в блок НЛ. Между блоками НЛ и БЗ существует обратная связь для того, чтобы данные по противоаварийным рекомендациям поступали в БЗ. В блоке НЛ производится сравнение данных работоспособности СТС с последующим выявлением причин того или иного результата по расчётам показателей работоспособности СТС. В блоке НЛ осуществляется прогнозирование технического состояния СТС на основе прогноза данных блока НС с учётом данных из блоков БЗ и НС по результатам анализа состояния СТС. ГЭС с помощью блока НЛ формирует рекомендации по предотвращению или ликвидации аварий для ЛПР. В целом, блок НС совместно с БЗ выполняет функцию обучения и адаптации ГЭС, а блок НЛ – функцию управления и влияет на качество прогноза состояния СТС.

На значения работоспособности СТС влияют различные факторы: риск отказа, критичность, оценка ресурса, функция погрешности, риск ошибки с сигналом [9,10].

Риск отказа СТС R описывается следующей зависимостью: $R = f(P, Y, Q)$, где P – вероятность выхода из строя СТС; Y – ущерб от выхода из строя СТС; Q – вес риска; вероятность выхода из строя СТС P – формулой теории надёжности:

$P(t) = 1 - \exp\left[-\int_0^t \lambda(\tau) d\tau\right]$, где $\lambda(t)$ – интенсивность отказов узлов, зависящая от времени.

Риск отказа СТС R определяется по сумме рисков для элементов r_i СТС:

$R = \sum_{i=1}^m r_i \cdot y_i$, где y_i – ущерб от выхода из строя i -го узла СТС; m – количество технических узлов СТС; r_i – риск отказа для узла i : $r_i = q_i \cdot p_i(t)$, где $p_i(t)$ – вероятность выхода из строя узла СТС в момент времени t , q_i – вес риска для узла i .

Критичность k_i (вероятность перехода системы в отказ при условии, что данный элемент вышел из строя) описывается следующей формулой: $k_i = P(S_n / S)$, где $P(S_n / S)$ – вероятность перехода СТС в неисправное состояние S_n из исправного S при условии, что узел i вышел из строя. При этом риск отказа с учётом критичности r_i^k определяется: $r_i^k = k_i \cdot r_i = k_i \cdot q_i \cdot p_i(t)$; риск отказа системы с учётом критичности R^k :

$$R^k = \sum_{i=1}^m r_i^k \cdot y_i.$$

Оценка ресурса C_i – оценка затрат на восстановление технического узла: $C_i = V(C_i - C_0)$; до полного восстановления: $C_0 = V(C_0 - C_0) = 0$, где C_i – текущее состояние узла СТС, C_0 – состояние узла на момент восстановления, $V(C_i - C_0)$ – процесс перехода из одного состояния в другое.

Погрешности нахождения параметров ГЭС $\varphi(\delta)$ определяются следующим образом: $\varphi(\delta) = \min \varphi(\delta_1, \delta_2)$, где δ_1 – погрешность преобразователей параметров устройств; δ_2 – погрешность из-за потери данных:

$$\delta_1 = \frac{\delta}{X \cdot 100}; \quad \delta_2 = \frac{y(t) - y'(t)}{U},$$

где X – значения СТС, δ – относительная погрешность, $y(t)$ – идеальный сигнал, $y'(t)$ – реальный сигнал, U – диапазон измерений, t – время наблюдения сигнала.

Риск ошибки с сигналом R^S при его искажении или некорректного принятия при передаче данных из одного блока к другому определится как:

$$R^S = \sum_{i=1}^z l_i \cdot p(S_j / S_i),$$

где S_j – сигнал с ошибкой, S_i – правильный сигнал, l_i – потери при ошибке, $p(S_j / S_i)$ – вероятность ошибки, z – число сигналов.

В результате зависимость работоспособности СТС W от факторов, влияющих на точность данных работы СТС, определяется следующим образом: $W = f(R^k, C, \varphi(\delta), R^S)$.

Из теории надежности известны критерии работоспособности СТС: прочность, жёсткость, устойчивость, теплоёмкость, износоустойчивость, виброустойчивость, надёжность. При этом зависимость работоспособности элемента СТС w_i от критериев работоспособности выражается как: $w_i = f(x_{i1}, x_{i2}, \dots, x_{in})$, где $x_{i1}, x_{i2}, \dots, x_{in}$ – значения критериев работоспособности для i -го элемента; n – количество критериев.

Вес работоспособности q_i^w каждого элемента системы оценивается с помощью метода парных сравнений Саати на основании мнений экспертов [11]. Работоспособность элемента рассчитывается на основании составляющих

вышеперечисленных факторов, влияющих на общую работоспособность. Для диагностики и прогнозирования технического состояния СТС применяются анализ временных рядов [12].

При учёте критериев работоспособность элемента СТС w_i в момент времени t в общем виде выражается следующей формулой:

$$w_i(t) = N(q_i^w, r_i^w, C_i, \varphi(\delta_i), r_i^s, (x_{i1}, x_{i2}, \dots, x_{im}))(t), \quad (1)$$

где $N()$ – нейросетевая функция; r_i^s – риск ошибки с сигналом для технического узла i .

С учётом (1) формула работоспособности СТС W в момент времени t с учётом всех факторов, влияющих на точность измерений, имеет вид:

$$W(t) = N \left(\begin{array}{l} (q_1^w, r_1^k, C_1, \varphi(\delta_1), r_1^s, (x_{11}, x_{12}, \dots, x_{1n})), \\ (q_2^w, r_2^k, C_2, \varphi(\delta_2), r_2^s, (x_{21}, x_{22}, \dots, x_{2n})), \dots, \\ (q_m^w, r_m^k, C_m, \varphi(\delta_m), r_m^s, (x_{m1}, x_{m2}, \dots, x_{mn})), \end{array} \right) (t) = N(w_1, w_2, \dots, w_m)(t),$$

где $w_1, w_2, \dots, w_m$ – работоспособности элементов СТС, m – количество элементов СТС.

Данные из СТС поступают в БЗ, которая объединяет в себе фреймовую и нечётко-продукционную модель [13], её удобно использовать как НС, так и НЛ. Фреймовая БЗ используется для диагностики и статистики состояния СТС, а нечётко-продукционная для описания причинно-следственных связей между состоянием работоспособности СТС и значением её различных критериев.

Фреймовая БЗ представлена в общем виде:

$$KB = \langle Fw, Frk, FC, F\varphi\delta, Frs, \{Fw_m\}, \{Frk_i\}, \{FC_j\}, \{F\varphi\delta_k\}, \{Frs_l\} \rangle,$$

где Fw – фрейм-прототип работоспособности элемента, Frk – фрейм-прототип риска с учётом критичности, FC – фрейм-прототип оценки ресурса, $F\varphi\delta$ – фрейм-прототип функции погрешности, Frs – фрейм-прототип риска ошибки с сигналом, $\{Fw_m\}$ – множество фреймов-экземпляров работоспособности элемента, $\{Frk_i\}$ – множество фреймов-экземпляров риска с учётом критичности, $\{FC_j\}$ – множество фреймов-экземпляров оценки ресурса, $\{F\varphi\delta_k\}$ – множество фреймов-экземпляров функции погрешности, $\{Frs_l\}$ – множество фреймов-комплексов риска ошибки с сигналом.

Нечётко-продукционное правило (НПП) представлено в общем виде:

$$FR = \langle NFR, \{ \langle LVZ_i, LTZ_i \rangle \} \rightarrow \langle LVAS, LTAS \rangle, qR \rangle,$$

где NFR – имя НПП, LVZ_i, LTZ_i – лингвистическая переменная (ЛП) i -го значения, а также её лингвистический терм (ЛТ) из предпосылки НПП, $LVAS, LTAS$ – ЛП аварийной ситуации и её ЛТ из заключения НПП, qR – весовой коэффициент НПП.

Лингвистическая переменная представлена в общем виде:

$$LV = \langle NLV, TSLV, ULV, GLV, MLV, TLV \rangle,$$

где NLV – имя ЛП, $TSLV$ – терм-множество ЛП, ULV – область определения каждого элемента $TSLV$, GLV – синтаксические правила, которые порождают название ЛТ, часто преподносятся в виде формальной грамматики, MLV – семантические правила,

которые задают функции принадлежности (ФП) ЛТ, порождённых синтаксическими правилами GLV , TLV – тип переменных ЛТ.

Лингвистический терм переменной представлен в общем виде: $LT = \langle NLT, MF \rangle$, где NLT – имя ЛТ, MF – функция принадлежности ЛТ.

НПП в БЗ представлены в табл. 1, где $\overline{1, k_i}$, – номер НПП для аварийной ситуации AS_i , $X_j, j = \overline{1, n}$, – входные ЛП признаков, z_j^{i, k_j} и $w_i^{k_j}$ – ЛТ входной переменной X_j и выходной переменной $Y = AS_i$ в НПП с номером k_i , $W_i, i = \overline{1, m}$, – обобщённый ЛТ выходной переменной $Y = AS_i$ в НПП группы i , $Y = AS_i, i = \overline{1, m}$, – выходные ЛП аварийных ситуаций, $qR_i^{k_j}$ – весовой коэффициент НПП с номером k_i , в соответствии с ЛТ переменной аварийной ситуации $Y = AS_i$.

В НС производятся расчёты работоспособностей, значения которых вместе с другими расчётами передаются в БЗ и блок НЛ. Структура НС, разработанной при помощи пакета Fuzzy Logic Toolbox, по которой будут проводиться расчёты, в общем виде представлена на рис. 2.

Таблица 1.

Представление НПП в БЗ

№ НПП	ЕСЛИ			ТО			
	X_i	$...X_j...$	X_n	Y			qR
[1,1]	$z_1^{1,1}$	$z_j^{1,1}$	$z_n^{1,1}$	w_1^1	W_1	AS_1	qR_1^1
[1,2]...	$z_1^{1,2}$	$z_j^{1,2}$	$z_n^{1,2}$	w_1^2			qR_1^2
[1, k_i]	z_1^{1,k_1}	z_j^{1,k_1}	z_n^{1,k_1}	$w_1^{k_1}$			$qR_1^{k_1}$
...							
[i ,1]	$z_1^{i,1}$	$z_j^{i,1}$	$z_n^{i,1}$	w_i^1	W_i	AS_i	qR_i^1
[i ,2]...	$z_1^{i,2}$	$z_j^{i,2}$	$z_n^{i,2}$	w_i^2			qR_i^2
[i , k_i]	z_1^{i,k_i}	z_j^{i,k_i}	z_n^{i,k_i}	$w_i^{k_i}$			$qR_i^{k_i}$
...							
[m ,1]	$z_1^{m,1}$	$z_j^{m,1}$	$z_n^{m,1}$	w_i^1	W_m	AS_m	qR_m^1
[m ,2]...	$z_1^{m,2}$	$z_j^{m,2}$	$z_n^{m,2}$	w_i^2			qR_m^2
[m , k_m]	z_1^{m,k_m}	z_j^{m,k_m}	z_n^{m,k_m}	$w_i^{k_m}$			$qR_m^{k_m}$

В блоке НЛ нечёткое множество данных показано через базовую шкалу H , а также функцию принадлежности $\mu(x_i), x \in H$, которая принимает значения на сегменте

$$[0,1]: H = \sum_{i=1}^n \frac{x_i}{\mu(x_i)}, \text{ где } x_i - i\text{-ое значение базовой шкалы.}$$

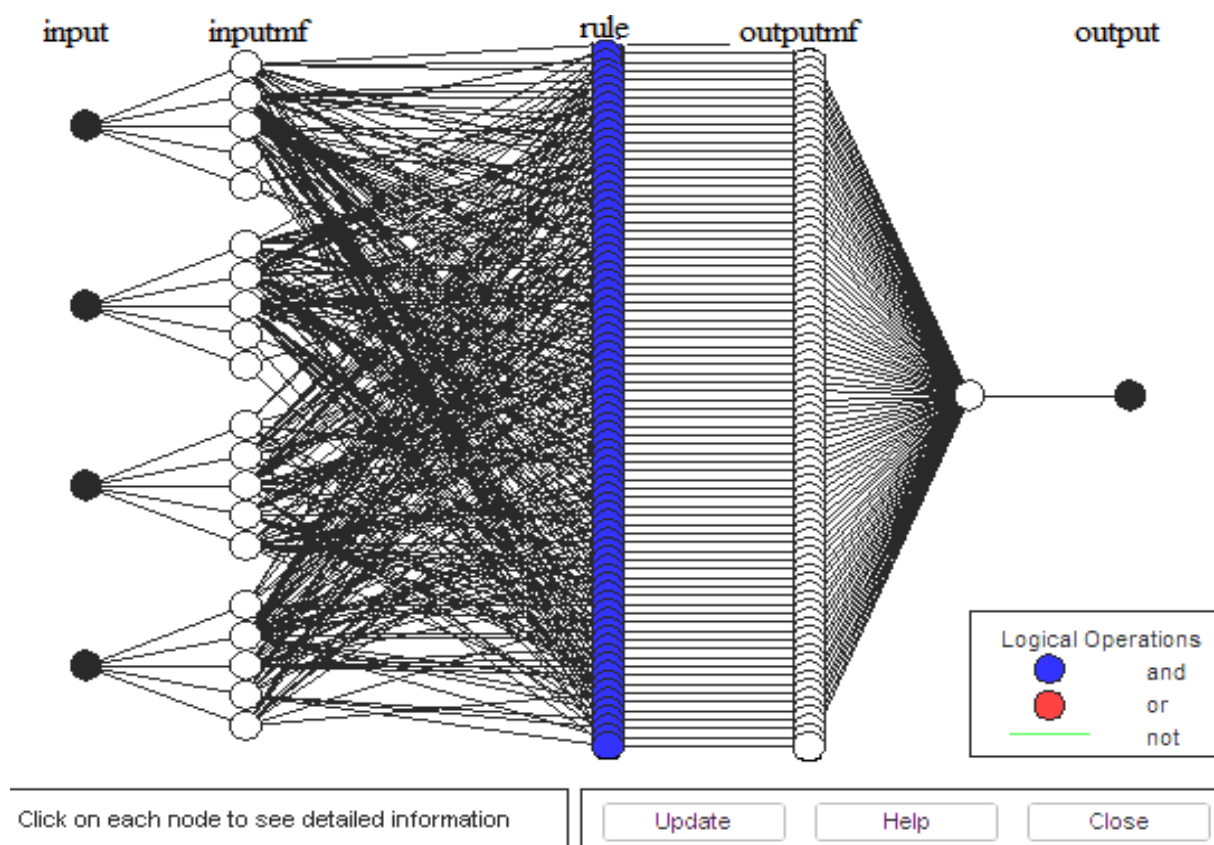


Рис. 2. Визуализация структурной схемы нечёткой НС

Основа для нечётких выводов выстраивается через нечёткое отношение, которое определяет связь нескольких субъективных понятий (правил) x и y , и является подмножеством декартового произведения $X \times Y$ [15]: $G = \int_{(x,y) \in (X \times Y)} \mu_G(x, y) / (x, y)$.

Большое число нечётких отношений используется в композиции правил вывода. Самое распространённое из них – операция взятия минимума. Для организации логического вывода используется композиционное правило, предложенное Л. Заде: $E = A \circ G, \mu_E(y) = \max_x \min(\mu_A(x), \mu_G(x, y))$.

По данным из блоков НС и БЗ, блок НЛ проводит диагностику технического состояния СТС и находит причину в БЗ.

Поиск необходимой информации в БЗ осуществляется путём поиска прямого вывода и поиска в ширину [14]. В системах с прямым выводом по известным данным находится те данные, которые им соответствуют. При поиске в ширину система НЛ проанализирует все данные за последний период времени, а потом, если не найдёт совпадений, пойдёт на следующий период. В итоге, при диагностике и прогнозировании аварийных и нештатных ситуаций находится их причина и способы их устранения, которые передаются ЛПР. Алгоритм ПАУ СТС представлен на рис. 3.

Результаты. Для исследований выбрана судовая СТС, а именно система кондиционирования воздуха, работающая в «зимнем» режиме. Она состоит из: электровентилятора, нагревателей 1-ой и 2-ой ступеней, увлажнителя воздуха [15,16]. Результаты диагностики СТС записаны в табл. 2 значений работоспособности СТС по некоторым временным рядам $v(t)$. В табл. 2 работоспособность: w_1 – электровентилятора, w_2 – нагревателя воздуха 1-й ступени, w_3 – увлажнителя воздуха, w_4 – нагревателя воздуха 2-й ступени, W – системы кондиционирования воздуха.

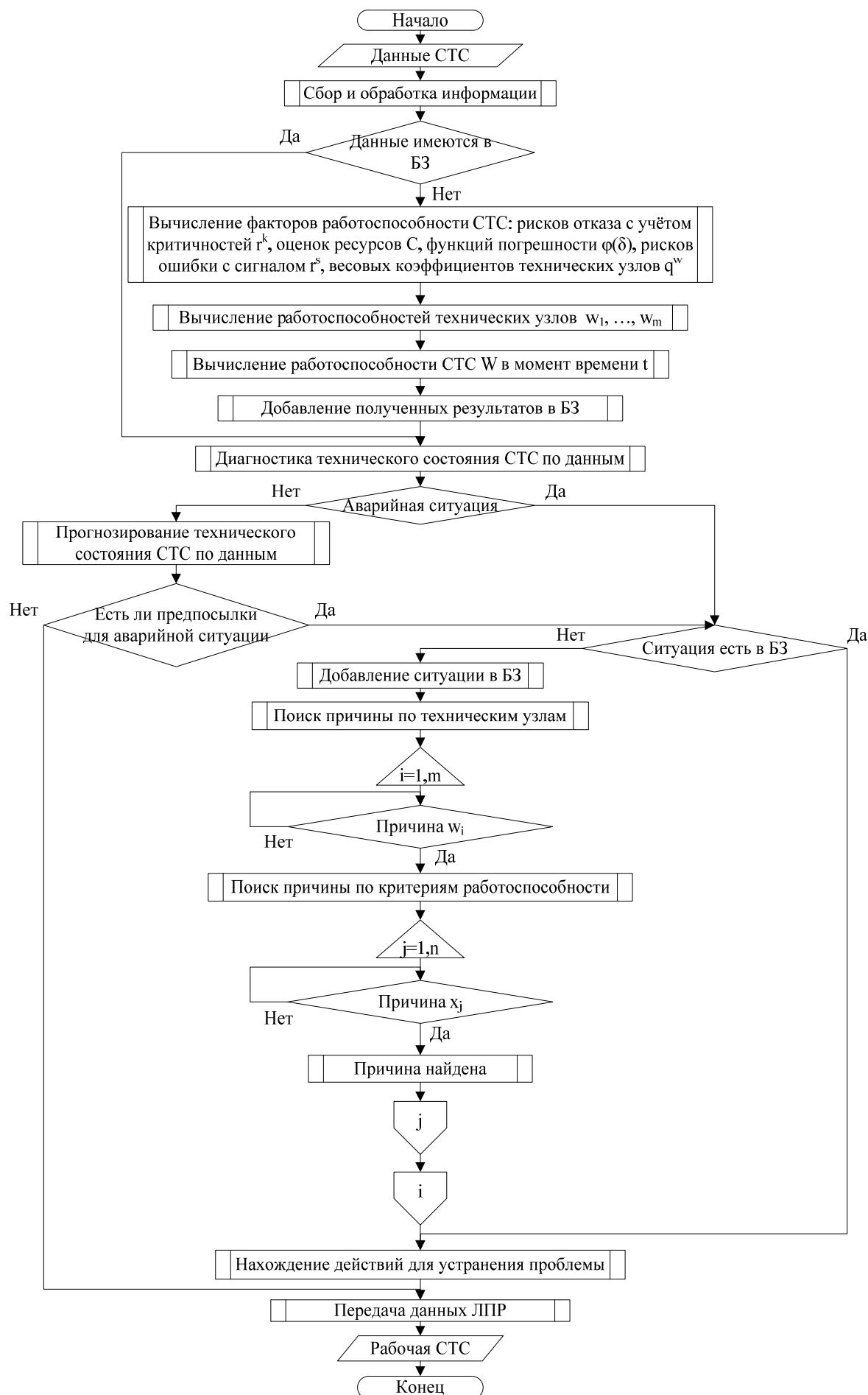


Рис. 3. Алгоритм ПАУ СТС

Таблица 2.

Значения работоспособности элементов СТС

Данные	Временной ряд $v(t)$			
	$v(t_1)$	$v(t_2)$	$v(t_3)$	$v(t_4)$
w_1	75	80	63	70
w_2	83	95	59	72
w_3	90	87	74	88
w_4	78	69	84	89
W	95.0151	94.2643	90.1223	96.3569

Используя среду Matlab – Network/Data Manager, проведен прогноз данных СТС на основе модели НС.

При диагностике технического состояния СТС произведено обучение НС (рис. 4).

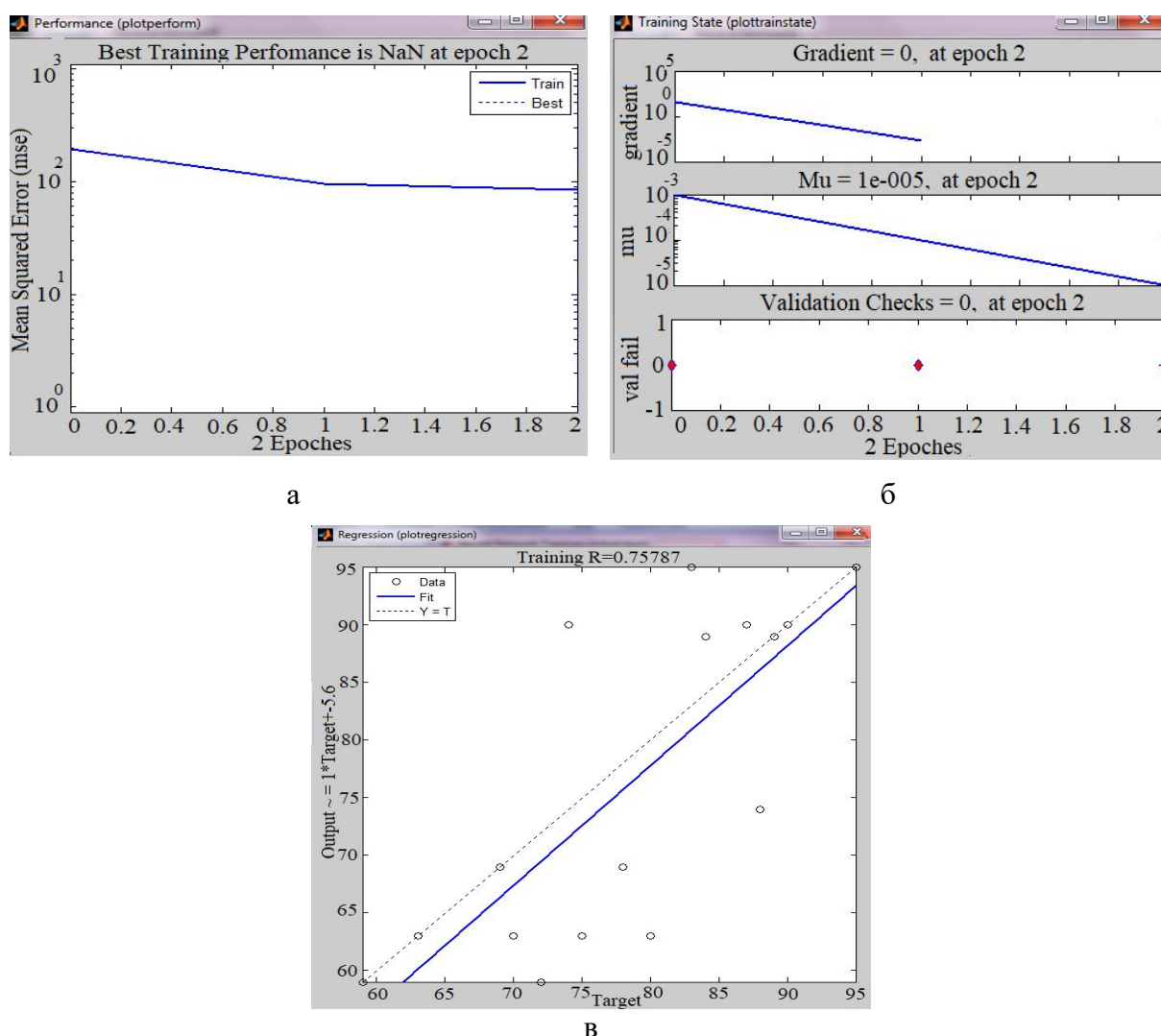


Рис. 4. Результаты обучения сети: а – производительность; б – состояние обучения; в – регрессия

На рис. 5 и 6 представлены таблицы полученных данных и ошибок в результате обучения соответственно.

	1	2	3	4
1	63	63	63	63
2	95	95	59	59
3	90	90	90	74
4	69	69	89	89

Рис. 5. Таблица полученных данных

	1	2	3	4
1	12	17	0	7
2	-12	0	0	13
3	0	-3	-16	14
4	9	0	-5	0

Рис. 6. Таблица ошибок

Схема модели нейронной сети в Simulink приведена на рис. 7.

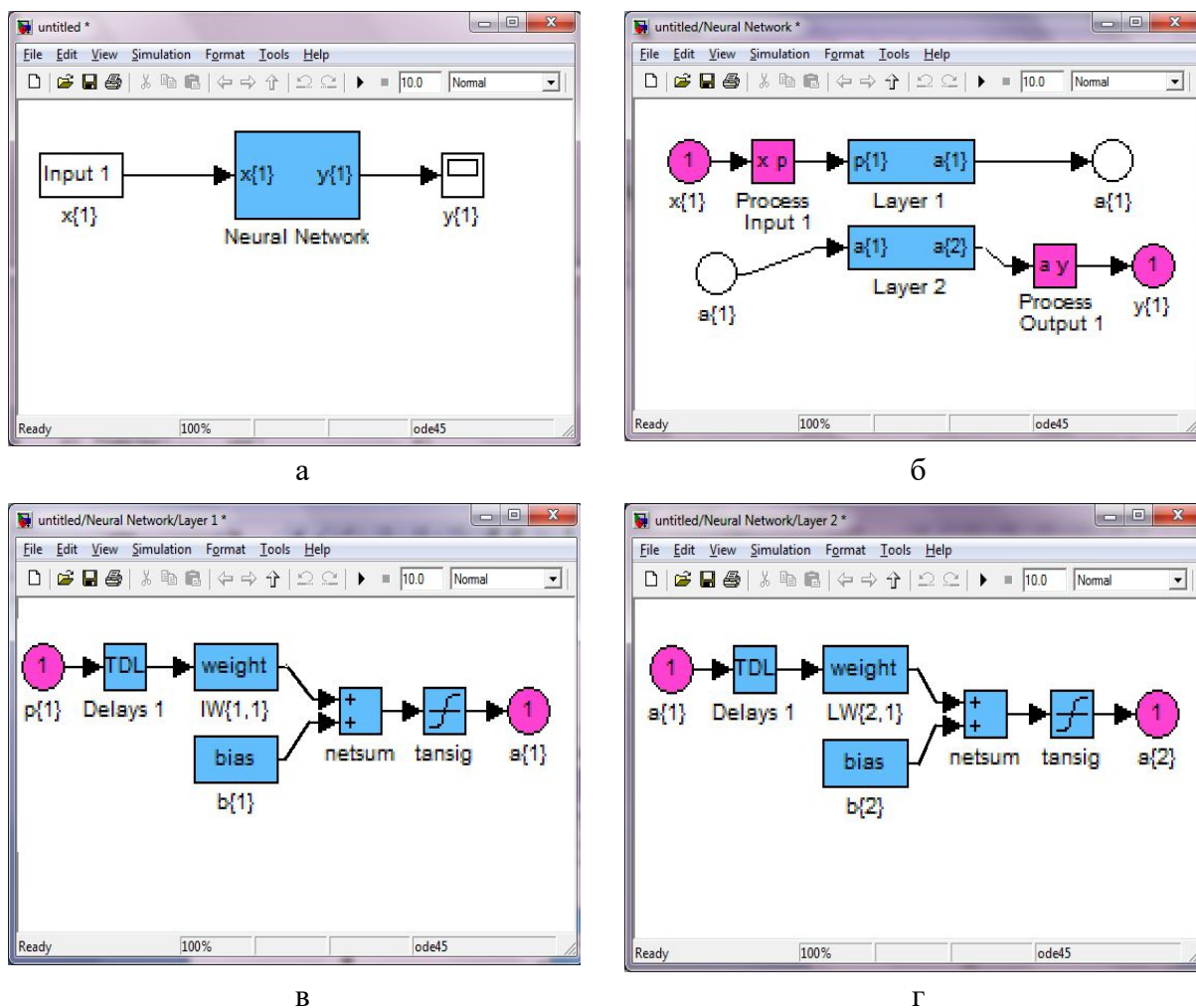


Рис. 7. Схема модели нейронной сети: а – общая модель НС в Simulink; б – модель блока Neural Network; в – модель первого уровня блока Neural Network; г – модель второго уровня блока Neural Network

В итоге была вычислена работоспособность системы на основе прогноза данных (табл. 3). Среднее значение ошибки нейронной сети мало и равно 2,25. Таким образом, точность прогноза высока, потому что скорость реагирования ПАУ также высока.

Таблица 3.

Результаты прогноза и работоспособность СТС

Данные	w_1	w_2	w_3	w_4	W
Временной ряд $v(t)$	63	59	74	89	90.8509

Выводы

При нахождении работоспособности элементов и всей СТС, при диагностике и прогнозировании их технического состояния учтено множество факторов и критериев работоспособности СТС. При моделировании НС для диагностики и прогнозирования работоспособности судовой СТС кондиционирования воздуха получены результаты обучения сети за 2 эпохи. Среднее значение ошибки нейронной сети относительно небольшое: 2,25, что подтверждает точность проведенных вычислений. В целом, полученные результаты свидетельствуют об эффективности ПАУ СТС.

Список литературы

1. Рудниченко, Н.Д. Оценки структурного и функционального рисков сложных технических систем / Н.Д. Рудниченко, В.В. Вычужанин. // Восточно-Европейский журнал передовых технологий. Інформаційні технології. Системи управління. — 2014. — Т. 1, № 2 (67). — С. 18–22.
2. ГОСТ 19176-85: Системы управления техническими средствами корабля. Термины и определения. — М.: ИПК Издательство стандартов, 2010. — 12 с.
3. Абеуов, Р.Б. Противоаварийное управление в энергосистемах / Р.Б. Абеуов. — Томск: Изд. ТПУ, 2007. — 39 с.
4. Коновалов, С.Н. Гибридные экспертные системы для противоаварийного управления сложными техническими системами / С.Н. Коновалов, В.В. Вычужанин. // Збірник наукових праць «Вісник Одеського національного морського університету». — 2017. — Т. 51, № 2. — С. 165–179.
5. Бубнов, Д.В. Разработка методов и алгоритмов принятия управленческих решений в автоматизированных машиностроительных системах: дисс. канд. техн. наук: 05.13.06 / Бубнов Д.В. — М, 2012. — 180 с.
6. Алгазин, Г.И. Информационные технологии комплексной оценки компетентности выпускника вуза / Г.И. Алгазин, О.В. Чудова. // Вестник НГУ. Серия: Информационные технологии. — 2009. — С. 70–78.
7. Козенко, З.Н. Поддержка принятия управленческих решений: инструментально-информационное обеспечение / З.Н. Козенко, А.Ф. Рогачёв, А.Л. Нахшунов, И.А. Карапузов. — Волгоград: ВГУ, 2001. — 124 с.
8. Жернаков, С.В. К вопросу о построении гибридных нейро-нечётких экспертных систем диагностики и контроля ГТД / С.В. Жернаков. // Управление в сложных системах. — 1999. — С. 119–126.
9. Шibaева, Н.О. Методы оценки и прогнозирования технического состояния судовых сложных систем: дисс. канд. техн. наук: 05.22.20 / Шibaева Н.О. — Одесса, 2016. — 225 с.
10. Бойко, В.Д. Методика оценки и обеспечения живучести функциональных комплексов судовых технических систем при их эксплуатации: дисс. канд. техн. наук: 05.22.20 / В.Д. Бойко. — Одесса, 2013. — 252 с.
11. Коновалов, С.Н. Информатизация дистанционного диагностирования состояния сложных технических систем / С.Н. Коновалов, В.В. Вычужанин. // Информатика и математические методы в моделировании. — 2016. — Т. 6, № 1. — С. 61–70.
12. Коновалов, С.Н. Разработка гибридной экспертной системы для противоаварийного управления сложными техническими системами / С.Н. Коновалов, В.В. Вычужанин. // Материалы XXIII Международной научно-технической конференции «Информационные системы и технологии» ИСТ-2017. — Нижний Новгород, 2017. — С. 835–840.
13. Виен, Л.Н. Дистанционная диагностическая система на основе гибридных моделей знаний: дисс. канд. техн. наук: 05.13.01 / Ле Нгуен Виен. — Волгоград, 2015. — 123 с.
14. Гаврилова, Т.А. Базы знаний интеллектуальных систем / Т.А. Гаврилова, В.Ф. Хорошевский. — СПб.: Питер, 2000. — 384 с.
15. Вычужанин, В.В. Повышение эффективности эксплуатации судовой системы комфортного кондиционирования воздуха при переменных нагрузках. Монография / В.В. Вычужанин. — Одесса: ОНМУ, 2009. — 206 с.
16. Вычужанин, В.В. Математические модели нестационарных режимов воздухообработки в центральной СКВ / В.В. Вычужанин. // Збірник наукових праць «Вісник Одеського національного морського університету». — 2007. — № 23. — С. 172–185.

ІНФОРМАТИЗАЦІЯ ПРОТИАВАРІЙНОГО КЕРУВАННЯ СКЛАДНИМИ ТЕХНІЧНИМИ СИСТЕМАМИ

С.М. Коновалов, В.В. Вичужанін

Одеський національний морський університет,
вул. Мечнікова, 34, Одеса, 65029, Україна; e-mail: konserge-serge@rambler.ru

У даній статті був розроблений метод протиаварійного керування складними технічними системами із застосуванням гібридної експертної системи.

Представлена гібридна експертна система, що заснована на нейронній мережі і нечіткій логіці, а також використовує автоматизовану базу знань. Розглянуто фактори та критерії працездатності технічних вузлів складних технічних систем, на основі діагностики і прогнозування яких будується остаточний висновок про працездатність системи, і знаходяться дії, застосовуючи які, можна усунути аварійні або позаштатні ситуації в разі їх виявлення. В якості прикладу для моделювання нейронної мережі, були взяті дані суднової системи кондиціонування повітря.

Ключові слова: протиаварійне керування, складна технічна система, гібридна експертна система, нейронна мережа, нечітка логіка, база знань, діагностика, прогнозування

INFORMATIZATION OF EMERGENCY CONTROL OF COMPLEX TECHNICAL SYSTEMS

S.N. Kononov, V.V. Vichuzhanin,

Odesa National Maritime University,
34, Mechnikova Str., Odesa, 65029, Ukraine; e-mail: konserge-serge@rambler.ru

In this article, a method was developed for antifault control of complex technical systems using a hybrid expert system. The presented hybrid expert system is based on a neural network and fuzzy logic, and also uses an automated knowledge base. Factors and criteria for the operability of technical units of complex technical systems are considered, on the basis of diagnostics and forecasting, a final conclusion is made about the operability of the system, and actions are in place to apply which can eliminate emergency or abnormal situations in the event of their detection. As an example for simulating a neural network, data from the ship's air conditioning system were taken.

Keywords: antifault control, complex technical system, hybrid expert system, neural network, fuzzy logic, knowledge base, diagnostics, forecasting

МЕТОД ВІДОКРЕМЛЕННЯ КЛОНУ ВІД ПРООБРАЗУ В ЦИФРОВОМУ ЗОБРАЖЕННІ В УМОВАХ ВІДСУТНОСТІ ВІДМІННОСТЕЙ ПРИ ЇХ ПОСТОБРОБЦІ

І.І. Бобок

Одеський національний політехнічний університет,
просп. Шевченка, 1, Одеса, 65044, Україна; e-mail: werter666@ukr.net

Одним з найпоширеніших інструментів, що використовується при несанкціонованих змінах цифрових зображень (ЦЗ), залишається клонування. Цей програмний інструмент реалізований у всіх сучасних графічних редакторах, роблячи задачу виявлення результатів клонування однією з самих актуальних при перевірці цілісності цифрових контентів. Дана робота присвячена підвищенню інформативності результатів виявлення порушень цілісності ЦЗ за рахунок розробки методу відокремлення області клону від прообразу в умовах відсутності відмінностей в їх обробці. Розроблений метод не використовує технологію цифрових водяних знаків, яка сама по собі приводить до порушення цілісності вихідного зображення, але є на сьогоднішній день дуже розповсюдженою при організації захисту ЦЗ від несанкціонованих змін. Ефективність алгоритмічної реалізації розробленого методу практично не залежить від сили збурної дії, що використовується в процесі постобробки клонованого зображення, та зростає зі зменшенням абсолютних розмірів клону, прообразу. Аналогів запропонованого методу з погляду розв'язуваної задачі (відокремлення клону від прообразу) і висунутих умов, у яких шукається її розв'язок, у відкритих джерелах не виявлено.

Ключові слова: цифрове зображення, клонування, клон, прообраз, відокремлення клону від прообразу, матриця мінімальних блокових відмінностей

Вступ

Забезпечення кібербезпеки, інформаційної безпеки на сьогоднішній день стає однією з самих актуальних проблем не тільки в інформаційному просторі, але й у життєвому просторі людини в цілому. Цілісність інформації є однією з основних вимог при організації процесу її захисту, порушення якої з різними цілями з урахуванням сучасного рівня розвитку інформаційних та комп'ютерних технологій стало задачею, яка легко реалізується.

Одним з основних інструментів, що використовуються при несанкціонованих змінах цифрових зображень (ЦЗ), є клонування [1-3], яке програмно реалізовано у всіх сучасних графічних редакторах. При клонуванні оригінальна область ЦЗ (прообраз) копіюється й переноситься в інше місце цього ж ЦЗ, створюючи клон. Цілі клонування можуть бути різними: усунення небажаного об'єкта з ЦЗ, збільшення кількості об'єктів на зображенні, зміна місця розташування об'єкта і т.д. Питанню виявлення результатів клонування в даний момент приділяється багато уваги [4-6], однак лише дуже незначна кількість робіт присвячена розробці методів і алгоритмів, що дозволяють визначати, яка з виявлених у ході експертизи цілісності зображення областей є оригінальною (прообразом), а яка є її клоном, тобто відокремлювати клон від прообразу [7,8]. Ця інформація часто виявляється надзвичайно важливою, підвищуючи інформативність результатів виявлення порушення цілісності зображення, наприклад, якщо ЦЗ використовується як речовинний доказ у судових розглядах, як доказ при слідчих заходах і ін., при цьому ключовим моментом є наявність/відсутність деяких деталей на

ньому, наприклад, наявність/відсутність на фото людини особливих прикмет (родимок, шрамів і т.д.) у чітко визначених місцях. Необхідно відзначити, що при розв'язку таких задач області клону, прообразу можуть бути дуже малими (як в абсолютному, так і у відносному сенсі), що створює додаткові труднощі, причому ще на етапі виявлення цих областей.

Існуючі методи для розв'язку задачі відокремлення клону від прообразу, інформація про які є доступною з відкритих джерел, ґрунтуються на двох принципово різних підходах до організації такого розподілу: виявлення відмінностей у результатах обробки клону й прообразу [7], використання цифрових водяних знаків (ЦВЗ), які вбудовуються у ЦЗ для організації його захисту від несанкціонованого порушення цілісності [8]. Перший підхід базується на тому, що після здійснення клонування область клону часто піддається якійсь обробці для її кращої «адаптації» до нової області розташування в межах ЦЗ, зокрема, розмиттю по контуру для зменшення ймовірності виникнення артефактів на зміненому ЦЗ. Саме виявлення результатів розмиття і є непрямым показником на клон в [7]. Однак така обробка є зовсім не обов'язковою для клону, більше того, з урахуванням того, що з'являються методи, які використовують наявність такої обробки для відокремлення клону від прообразу, для «автора» фальсифікації вона є просто небажаною, крім того, при малих розмірах клону/прообразу додаткова обробка клону з зазначеною вище метою просто не потрібна (рис.1 - у результаті клонування з лівої щочки хлопчика прибрана родимка, після чого ЦЗ без будь-якої постобробки збережено без втрат (Tif); артефакти на ЦЗ не виявлені). Другий підхід захисту ЦЗ від порушення цілісності, що пов'язаний з використанням ЦВЗ, хоча й розглядається автором даної статті як можливий [8], але не визначає остаточного рішення задачі в силу наступної основної причини: на формальному рівні процес вбудови ЦВЗ сам по собі порушує цілісність оригінального зображення.



а



б

Рис. 1. Ілюстрація можливості здійснення клонування в ЦЗ без використання будь-якої постобробки клону: а – оригінальне ЦЗ; б – ЦЗ, що отримано в результаті клонування

Таким чином, задача відокремлення клону від прообразу залишається до кінця невирішеною й актуальною, зокрема, для областей малих розмірів.

Мета статті й постановка досліджень

Метою роботи є підвищення інформативності результатів виявлення порушень цілісності ЦЗ шляхом розробки методу відокремлення виявлених клону від прообразу в умовах відсутності відмінностей у їхній обробці без використання технології ЦВЗ.

Для досягнення мети в роботі вирішуються наступні задачі:

1. Забезпечити незалежність ефективності розробленого методу від величини додаткової до клонування збурної дії на ЦЗ;
2. Забезпечити працездатність методу в умовах малих абсолютних розмірів областей клону й прообразу;
3. Розробити й оцінити кількісно ефективність алгоритмічної реалізації методу шляхом обчислювального експерименту.

Ефективність алгоритму буде оцінюватися кількістю помилок, що виникли при відокремленні клону від прообразу, тобто кількістю аналізованих зображень, що зазнали клонування, у яких області клону й прообразу були виявлені вірно (ці області мають непусте перетинання з реальними клоном і прообразом відповідно), але в результаті роботи розробленого алгоритму клон/прообраз трактується як прообраз/клон.

Основна частина

Оскільки передача інформації (ЦЗ, цифрові відео, аудіо) у даний момент по каналах зв'язку відбувається у форматах із втратами, то ЦЗ після того, як воно піддалося клонуванню, практично вірогідно буде збережено у форматі з втратами, що певним чином змінить зображення й може розглядатися як один зі способів його постобробки після клонування. Таким чином, можна вважати, що на практиці ЦЗ після клонування обов'язково буде піддаватися постобробці, але тільки ця постобробка не буде відрізнятися для областей клону й прообразу, що й припускається далі.

Наявність постобробки ЦЗ ускладнює задачу виявлення результатів клонування, оскільки області клону й прообразу з точно співпадаючих до неї стають такими, які відрізняються одна від одної, після неї. Найбільш складною ця задача стає у випадку значних додаткових збурних дій. В [9,10] нещодавно був запропонований метод *KL* для виявлення результатів порушення цілісності ЦЗ, що відбулося внаслідок клонування, розроблений для знаходження в ЦЗ областей клону й прообразу в умовах наявності/відсутності додаткових збурних дій (у тому числі, значних) на зображення, ефективність якого перевищує сучасні аналоги, при цьому області клону й прообразу, які визначаються цим методом, можуть мати дуже незначні розміри - до 0.06% ЦЗ, на відміну від аналогів, де, як правило, розглядається клон/прообраз, що становить не менш 1% ЦЗ. Однак у рамках розробленого методу не передбачене відокремлення клону від прообразу.

Висока ефективність методу *KL* змушує дуже уважно поставитися до використаного в ньому формальному математичному апарату з метою створення можливостей використання його для розв'язку задач даної роботи.

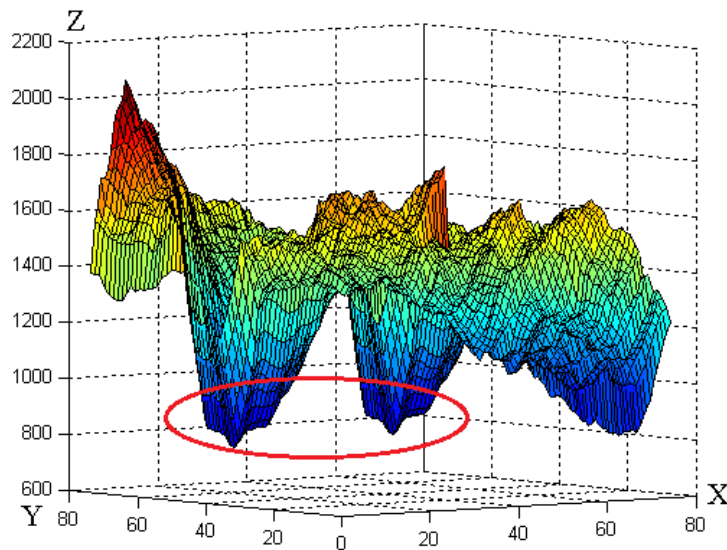
Основним об'єктом дослідження в блоково орієнтованому методі *KL* є матриця мінімальних блокових відмінностей (ММБВ), яка ставиться у відповідність аналізованому зображенню за наступним правилом [9,10]. Нехай F - $n \times m$ – матриця ЦЗ, для аналізу якого використовуються блоки розміру $l \times l$. Кожному елементу

f_{ij} , $i = \overline{1, n-l+1}$, $j = \overline{1, m-l+1}$, матриці F ставиться в співвідношення $l \times l$ - блок B_{ij} , який є підматрицею F , для якого на місці (1,1) знаходиться елемент f_{ij} :

$$B_{ij} = \begin{pmatrix} f_{ij} & f_{i,j+1} & \dots & f_{i,j+l-1} \\ f_{i+1,j} & f_{i+1,j+1} & \dots & f_{i+1,j+l-1} \\ \dots & \dots & \dots & \dots \\ f_{i+l-1,j} & f_{i+l-1,j+1} & \dots & f_{i+l-1,j+l-1} \end{pmatrix}$$

Елементи g_{ij} $(n-l+1) \times (m-l+1)$ - ММБВ G , яка ставиться в співвідношення ЦЗ, відображають величину найменшої відмінності $l \times l$ - блоку B_{ij} від будь-якого іншого $l \times l$ -блоку B_{kl} матриці F в сенсі величини $\sum_{t,p=1}^l r_{tp}$, де r_{tp} , $t, p = \overline{1, l}$, - елементи $l \times l$ -матриці R , $R = |B_{ij} - B_{kl}|$.

У даній роботі основна увага приділена відокремленню області клону від прообразу малих абсолютних розмірів. Відповідним блокам B_{ij} клону і B_{kl} прообразу в ММБВ відповідають однакові за значенням локальні (глобальні) мінімуми: $g_{ij} = g_{kl}$ (для частини ЦЗ, представленого на рис.1(б), що включає клон і прообраз, ММБВ і результат роботи алгоритмічної реалізації KL представлені на рис.2). ММБВ дає достатню інформацію для виявлення областей клону й прообразу в ЦЗ, що зазнало клонування, але вона ніяк не використовується в KL для відокремлення однієї області від іншої.



а



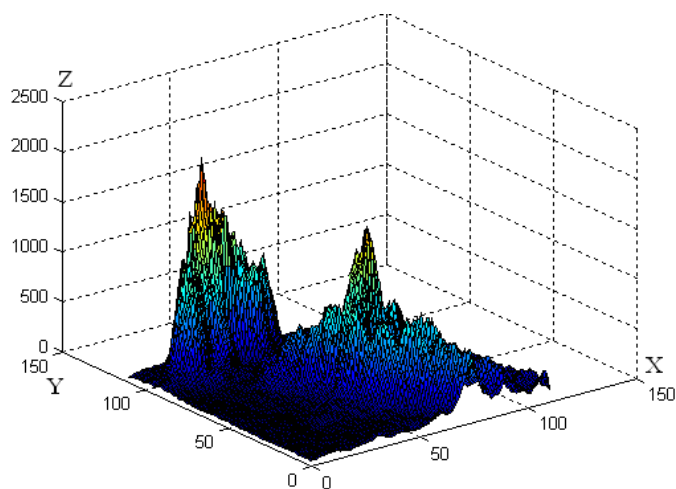
б

Рис. 2. Ілюстрація роботи алгоритмічної реалізації методу KL у випадку малих розмірів клону/прообразу для ЦЗ, представленого на рис.1(б): а – частина графіка функції, яка інтерполює елементи ММБВ, що відповідає частині ЦЗ (розміром 80×80 пікселів), яка містить області клону й прообразу; б – результат роботи KL (для наочності наведена частина аналізованого ЦЗ, що відповідає частині ММБВ, де області клону й прообразу пофарбовані в червоний колір)

Ідея запропонованого нижче методу відокремлення виявлених клону від прообразу з врахуванням формального математичного апарату KL у неформальному виді може бути сформульована наступним чином. Як відомо [11], для значень яскравості пікселів, що знаходяться поряд, в оригінальному ЦЗ існує кореляційний зв'язок: значення яскравості пікселів, що знаходяться на незначній відстані один від одного, відрізняються також незначно, не можуть змінюватися стрибкоподібно, що приводить до того, що сума значень яскравості пікселів, що належать $l \times l$ -блоку оригінального ЦЗ також змінюється незначно при зсуві блоку на кілька пікселів (униз, вгору, вправо, вліво). Ілюстрацією цьому може бути ММБВ (рис. 3), побудована для оригінального ЦЗ з використанням 8×8 -блоків, інтерполуюча функція для якої не має значних стрибків для близьких значень. Вбудова клону - «чужорідної» частини в деяку оригінальну підобласть ЦЗ приводить до порушення цієї кореляції в деякому околі клону, однак з врахуванням того, що кореляція має місце для близько розташованих пікселів, порушення цієї кореляції в сенсі формування ММБВ найбільш яскраво буде проявлятися для «чужорідної» частини малих розмірів (зв'язок між пікселями, що знаходяться поруч, пікселями, між якими 1, 2 інших, буде сильніше, ніж між тими, між якими знаходяться, наприклад, 10 інших пікселів. В останньому випадку зв'язку практично немає, тому він й «не розірветься», якщо між ними поставити 10 інших пікселів). Це імовірно приведе до того, що для «чужого» блоку клону мінімальна блокова відмінність від блоків у деякому його «новому» блоковому околі, що не включає прообраз, буде більше, ніж для оригінального блоку прообразу в його оригінальному околі. Висунута гіпотеза знайшла своє підтвердження на практиці в ході проведення обчислювального експерименту при використанні блоків ЦЗ малих розмірів і стала основою при розробці запропонованого нижче методу відокремлення клону від прообразу.



а



б

Рис. 3. Ілюстрація відсутності стрибкоподібних змін для близьких значень ММБВ оригінального ЦЗ: а – оригінальне ЦЗ; б – графік функції, що інтерполіє елементи частини ММБВ, яка відповідає області, виділеній на рис.3(а)

У ході експерименту було встановлено, що для значного числа ЦЗ, що зазнали клонування з малою областю клону, не тільки відповідні клону й прообразу елементи ММБВ, побудованих для їхніх околів, зв'язані зазначеним вище співвідношенням, але й графіки функцій, що інтерполіють елементи ММБВ, мають специфічний вигляд у випадку клону, визначаючи блок клону (або блок, що відповідає пікселю, який відстоїть від пікселя, що відповідає блоку клону, на 1,2,3 позиції) як локальний (глобальний) максимум ММБВ (рис.4(а)).

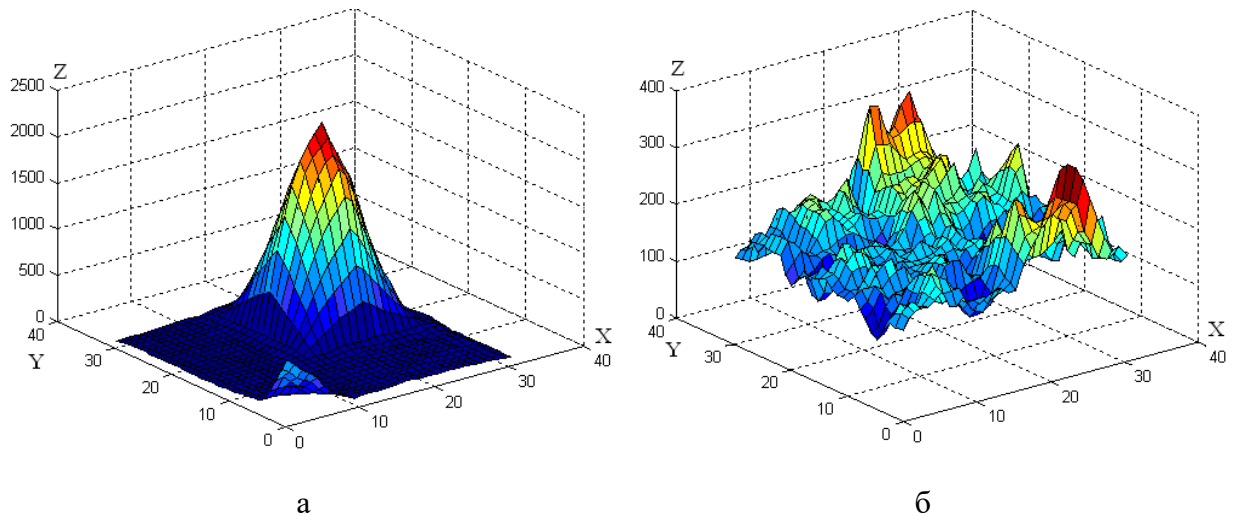


Рис. 4. ММБВ для околів клону й прообразу ЦЗ: а – ММБВ для околу клону; б – ММБВ для околу прообразу

З врахуванням всього вищесказаного, основні кроки запропонованого методу відокремлення клону від прообразу малих розмірів, що далі називається *КРМ*, наступні.

Крок 1. Нехай T_1 і T_2 - виявлені деяким алгоритмом області клону й прообразу, що мають малі абсолютні розміри. Визначити B_1 і B_2 - $l \times l$ -блоки матриці F аналізованого ЦЗ як можна меншого розміру такі, що

$$T_1 \subseteq B_1, T_2 \subseteq B_2, B_1 \cap B_2 = \emptyset.$$

Крок 2. Побудувати O_1 і O_2 - прямокутні $p \times p$ -околи B_1 і B_2 відповідно, розміри яких порівнянні з розмірами B_1 і B_2 , такі, що

$$\begin{aligned} B_1 &\subseteq O_1, B_2 \subseteq O_2, \\ O_1 \cap B_2 &= \emptyset, O_2 \cap B_1 = \emptyset. \end{aligned} \quad (1)$$

Крок 3. Для O_1 і O_2 побудувати ММБВ, використовуючи блоки розміру $l \times l$: M_1 і M_2 відповідно.

Крок 4. В M_1 і M_2 знайти значення елементів, що відповідають блокам B_1 і B_2 . Нехай ці елементи – це $m_{ij}^{(1)} \in M_1$ і $m_{kp}^{(2)} \in M_2$.

Крок 5 (відокремлення клону від прообразу).

$$\text{Якщо} \quad m_{ij}^{(1)} < m_{kp}^{(2)},$$

$$\text{то} \quad T_1 - \text{прообраз}, T_2 - \text{клон},$$

$$\text{інакше} \quad T_1 - \text{клон}, T_2 - \text{прообраз}.$$

Приклад виконання кроку 2 наведений на рис.5 для ЦЗ, представленого на рис.2. Зауважемо, що околи O_1 і O_2 можуть як перетинатися, так і бути непересічними. Основною умовою для них є (1).

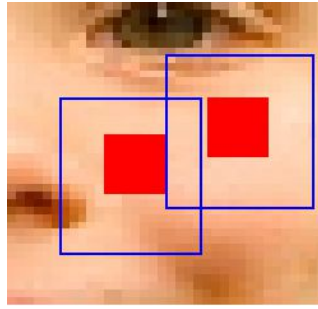


Рис. 5. Приклад побудови околів O_1 і O_2

Ефективність *KPM* при правильно виявлених попередньо областях клону й прообразу, очевидно, ніяк не повинна залежати від величини збурної дії, якій, можливо, піддалося клоноване ЦЗ, але явно буде залежати від розмірів B_1 і B_2 . Для підтвердження цього, а також для безпосередньої оцінки ефективності в середовищі *Matlab* був проведений обчислювальний експеримент. Експериментальна множина (ЕМ), на якій відбувалося тестування 2-ох алгоритмічних реалізацій розробленого методу, для яких використовувалися наступні значення переметрів: $l = 8$, $p = 48$; $l = 16$, $p = 56$, формувалося наступним чином. Кожне з обраних випадковим чином 200 ЦЗ з бази NRCS [12], піддавалося клонуванню, де як прообраз/клон брався один 8×8 – блок, після чого ЦЗ зберігалося у форматі Jpeg чотири рази з різними значеннями коефіцієнта якості $QF \in \{65, 75, 85, 95\}$, і ще двічі у форматі Tif після попереднього накладання на клоноване ЦЗ гауссівського шуму з нульовим математичним очікуванням і $D \in \{0.0001, 0.0005\}$. Все те ж саме було зроблено у випадку, коли в якості прообразу/клону брався один 16×16 – блок. Таким чином, потужність ЕМ – 2400 ЦЗ. Використання 8×8 у якості найменшого розміру блоку обумовлено тим, що клони/прообрази меншого розміру, як свідчать відкриті джерела, у даний момент існуючими методами систематично не визначаються, більше того, 8×8 -клони і прообрази визначаються лише деякими відповідними алгоритмами, зокрема [9,10].

Результати обчислювального експерименту представлені в табл.1,2.

Таблиця 1.

Помилки (%) при роботі алгоритму, що реалізує метод *KPM*, залежно від розмірів клону/прообразу й значення коефіцієнта якості QF , що використовувався для стиску клонованого ЦЗ з ЕМ

l	QF			
	65	75	85	95
8	19.5	20	20.5	19.5
16	25	24	25	26

Таблиця 2.

Помилки (%) при роботі алгоритму, що реалізує метод *KPM*, залежно від розмірів клону/прообразу й параметра гауссівського шуму, що накладався на клоноване ЦЗ

l	D	
	0.0001	0.0005
8	20.5	20
16	27.5	27.5

Таким чином, як і очікувалося, ефективність розробленого алгоритму є більшою у випадку менших розмірів областей клону/прообразу і не залежить від величини додаткової збурної дії. Для підвищення ефективності метод потребує вдосконалення, але необхідно відзначити, що до цього моменту способу розв'язку розглянутої задачі у відкритих джерелах в розглянутих умовах взагалі не пропонувалося.

Висновки

У роботі розроблений метод відокремлення виявлених попередньо клону від прообразу в умовах відсутності відмінностей у їхній обробці без використання технології ЦВЗ, що привело до підвищення інформативності результатів виявлення порушень цілісності ЦЗ, проведеного шляхом клонування.

Ефективність розробленого методу не залежить від величини збурної дії, яку зазнає клоноване ЦЗ, а залежить від розміру клону/прообразу, будучи більш високою для клону/прообразу менших розмірів. Кількість помилок при відокремленні клону від прообразу мінімально у випадку $l=8$, $p=48$ і становить $\approx 20\%$. Підвищення ефективності розробленого методу можливо за рахунок залучення аналізу додаткових формальних параметрів матриці клонованого ЦЗ, зокрема аналізу околів елементів, що відповідають клону й прообразу в ММБВ M_1 і M_2 , над чим зараз працює автор.

Список літератури

1. Kakar, P. Exposing postprocessed copy-paste forgeries through transform-invariant features / P. Kakar, N. Sudha. // IEEE Transactions on Information Forensics and Security. — 2012. — Vol. 7, No. 3. — Pp. 1018–1028.
2. Ali Qureshi, M. A review on copy move image forgery detection techniques / M. Ali Qureshi, M. Deriche. // IEEE 11th International Multi-Conference on Systems, Signals & Devices. — 2014. — Pp. 1–5.
3. Davarzani, R. Copy-move forgery detection using multiresolution local binary patterns / R. Davarzani, K. Yaghmaie, S. Mozaffari, M. Tapak. // Forensic science international. — 2013. — Vol. 231, No. 1. — Pp. 61–72.
4. Singh, Ratnam. Copy Move Tampering Detection Techniques: A Review / Ratnam Singh, Mandeep Kaur. // International Journal of Applied Engineering Research. — 2016. — Vol. 11, No. 5. — Pp. 3610–3615.
5. Mahdian, B. A bibliography on blind methods for identifying image forgery / B. Mahdian, S. Saic. // Signal Processing: Image Communication. — 2010. — Vol. 25, No. 6. — Pp. 389–399.
6. Mishra, M. Detection of Clones in Digital Images / M. Mishra, M.C. Adhikary. // International Journal of Computer Science and Business Informatics. — 2014. — Vol. 9, No. 1. — Pp. 91–102.
7. Лебедева, Е.Ю. Метод локализации и идентификации оригинальной и клонированной областей изображения / Е.Ю. Лебедева. // Информатика та математичні методи в моделюванні. — 2014. — Т. 4, № 1. — С. 76–84.
8. Кобозева, А.А. Выявление нарушений целостности цифрового изображения путем использования стеганографических алгоритмов / А.А. Кобозева, И.И. Бобок, Л.М. Дзюбинская. // Информатика та математичні методи в моделюванні. — 2015. — Т. 5, № 2. — С. 129–134.
9. Григоренко, С.М. Розвиток методу виявлення клонування в цифровому зображенні в умовах додаткових збурних дій / С.М. Григоренко. // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. — 2016. — Вип. 1 (31). — С. 85–98.
10. Кобозева, А.А. Метод выявления нарушения целостности цифрового изображения, проведенного путем клонирования, робастный к сжатию с потерями / А.А. Кобозева, И.И. Бобок, С.Н. Григоренко. // 17-я Международная научно-практическая конференция «Современные информационные и электронные технологии «СИЭТ-2016». — Одесса, 2016. — С. 127–128.
11. Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. — М.: Техносфера, 2006. — 1070 с.

12. NRCS Photo Gallery: [Электронный ресурс] // United States Department of Agriculture. Washington, USA. Режим доступа: <http://photogallery.nrcs.usda.gov> (Дата звернения: 26.07.2012).

МЕТОД ОТДЕЛЕНИЯ КЛОНА ОТ ПРООБРАЗА В ЦИФРОВОМ ИЗОБРАЖЕНИИ В УСЛОВИЯХ ОТСУТСТВИЯ РАЗЛИЧИЙ ПРИ ИХ ПОСТОБРАБОТКЕ

И.И. Бобок

Одесский национальный политехнический университет,
просп. Шевченко, 1, Одесса, 65044, Украина; e-mail: werter666@ukr.net

Одним из самых распространенных инструментов, используемый при несанкционированных изменениях цифровых изображений (ЦИ), остается клонирование. Этот программный инструмент реализован во всех современных графических редакторах, делая задачу выявления результатов клонирования одной из самых актуальных при проверке целостности цифровых контентов. Настоящая работа посвящена повышению информативности результатов выявления нарушений целостности ЦИ за счет разработки метода отделения области клона от прообраза в условиях отсутствия различий в их обработке. Разработанный метод не использует технологию цифровых водяных знаков, которая сама по себе приводит к нарушению целостности исходного изображения, но является на сегодняшний день очень распространенной при выявлении результатов несанкционированных изменений ЦИ. Эффективность алгоритмической реализации разработанного метода практически не зависит от силы возмущающего воздействия, использованного в процессе постобработки клонированного изображения, возрастает с уменьшением абсолютных размеров клона, прообраза. Аналогов предложенного метода с точки зрения решаемой задачи (отделение клона от прообраза) и выдвинутых условий, в которых ищется решение, в открытой печати не обнаружено.

Ключевые слова: цифровое изображение, клонирование, клон, прообраз, отделение клона от прообраза, матрица минимальных блоковых отличий

METHOD FOR SEPARATING OF CLONE AND PROTOTYPE IN A DIGITAL IMAGE IN THE CONDITIONS OF THE ABSENCE OF DIFFERENCES IN THEIR POST-PROCESSING

I.I. Bobok

Odesa National Polytechnic University,
1, Shevchenko Ave., Odesa, 65044, Ukraine; e-mail: werter666@ukr.net

The cloning is one of the most common and often used software tool for unauthorized changes of digital images that implemented in all modern graphics editors. Thus, the task of detection the results of cloning is one of the most urgent for verifying the integrity of digital content. This work is devoted to increasing the informative value of the results of the detection of integrity violations of the digital images due to the development of a new method for separating of clone and prototype in the absence of their post-processing or the absence of differences in this treatment. The developed method does not use the technology of digital watermarks which leads to a violation of the integrity of the original image. The effectiveness of the algorithmic implementation of the developed method is practically independent of the value of the disturbing effect used in the post-processing of the cloned image, and increases with decreasing absolute sizes of the clone and pre-image. Analogues of the proposed method from the point of view of the problem being solved (separating of clone and prototype) and the proposed conditions have not been found in the open sources.

Keywords: digital image, cloning, clone, prototype, separation of clone and prototype, matrix of minimal block differences

СПРОЩЕННЯ МАТЕМАТИЧНИХ МОДЕЛЕЙ ОБ'ЄКТІВ З РОЗПОДІЛЕНИМИ
ПАРАМЕТРАМИ НА ОСНОВІ МЕТОДУ РОЗЩЕПЛЕННЯА.А. Верлань¹, В.А. Іванюк²¹ Норвезький університет науки і технологій,
NTNU Gjøvik, Teknologiveien 22, 2815 Gjøvik, Norway; e-mail: andriy.verlan@ntnu.no² Кам'янець-Подільський національний університет імені Івана Огієнка,
вул. Уральська, 1, Кам'янець-Подільський, 32300, Україна; e-mail: wivanyuk@gmail.com

У статті розглянуто методи побудови спрощених моделей багатовимірних задач теплопровідності у вигляді добутку розв'язків одновимірних задач. Ефективність запропонованого підходу показано на основі аналізу складності чисельної реалізації даних моделей за допомогою різницевих методів.

Ключові слова: диференціальні рівняння в частинних похідних, задача теплопровідності, еквівалентні перетворення

Вступ

Інтенсивний розвиток комп'ютерно-інтегрованих систем характеризується постійно зростаючою складністю їх режимів, підвищення вимог до якості функціонування (швидкодія, точність, надійність, економічність тощо). Ці чинники породжують нові вимоги до методів і засобів математичного моделювання динамічних процесів у вказаних системах, в тому числі необхідність суттєвого врахування розподіленості параметрів елементів, вузлів, блоків та систем в цілому. Існуючі методи і засоби, що використовуються для розв'язування задач моделювання об'єктів з розподіленими параметрами, в переважній більшості ґрунтуються на використанні моделей у вигляді диференціальних рівнянь з частинними похідними [1, 2, 3, 4]. Серед чисельних методів розв'язання даних задач з двома або трьома просторовими змінними найбільш відомими є метод сіток та метод скінченних елементів, що породжує системи лінійних алгебраїчних рівнянь великої розмірності. Для підвищення точності часто використовується методика згущення сіток навколо точок з особливостями [3, 5]. Це приводить до збільшення часу розрахунків при розв'язуванні поставлених задач. Тому розробка методів еквівалентного перетворення базових моделей з метою забезпечення специфічних часових і ресурсних вимог є актуальною задачею [6].

Метою роботи є аналіз складності еквівалентних моделей багатовимірних задач теплопровідності.

Основна частина

Еквівалентні перетворення базових моделей. Розглянемо рівняння теплопровідності [1]:

$$\frac{\partial^2 v}{\partial x_1^2} + \frac{\partial^2 v}{\partial x_2^2} + \frac{\partial^2 v}{\partial x_3^2} = \frac{1}{\chi} \frac{\partial v}{\partial t}, \quad t > 0, \quad (1)$$

для прямокутного паралелепіпеда:

$$a_1 < x_1 < b_1, \quad a_2 < x_2 < b_2, \quad a_3 < x_3 < b_3, \quad (2)$$

де $v = v(x_1, x_2, x_3, t)$ – шукана функція температури; x_1, x_2, x_3, t – незалежні змінні, які визначають просторові координати та час; $a_1, a_2, a_3, b_1, b_2, b_3$ – сталі; χ – коефіцієнт теплопровідності.

Для деяких важливих типів початкових і граничних умов його розв'язком є добуток розв'язків трьох задач з однією змінною [1]; таким чином, якщо останні відомі, можна відразу ж написати і розв'язок поставленої задачі.

Припустимо, що функції $v_r = v_r(x_r, t)$, де $r=1, 2, 3$, є розв'язками одновимірних рівнянь теплопровідності:

$$\frac{\partial^2 v_r}{\partial x_r^2} = \frac{1}{\chi} \frac{\partial v_r}{\partial t}, \quad a_r < x_r < b_r, \quad t > 0, \quad (3)$$

з граничними умовами:

$$\alpha_r \frac{\partial v_r}{\partial x_r} - \beta_r v_r = 0, \quad x_r = a_r, \quad t > 0;$$

$$\alpha'_r \frac{\partial v_r}{\partial x_r} + \beta'_r v_r = 0, \quad x_r = b_r, \quad t > 0,$$

де $\alpha_r, \beta_r, \alpha'_r, \beta'_r$ – сталі, кожна з яких може бути рівною нулю (таким чином, включені випадки нульової початкової температури і відсутність теплового потоку на поверхні). Нехай початкові умови задаються функціями $V_r(x_r)$:

$$v_r(x_r, t) = V_r(x_r), \quad t = 0, \quad a_r < x_r < b_r, \quad r=1, 2, 3.$$

Тоді розв'язком рівняння (1) в області (2) при початковій умові:

$$v = V_1(x_1)V_2(x_2)V_3(x_3), \quad t = 0, \quad (4)$$

і при граничних умовах:

$$\alpha_r \frac{\partial v}{\partial x_r} - \beta_r v = 0, \quad x_r = a_r, \quad t > 0, \quad r=1, 2, 3: \quad (5)$$

$$\alpha'_r \frac{\partial v}{\partial x_r} + \beta'_r v = 0, \quad x_r = b_r, \quad t > 0, \quad r=1, 2, 3, \quad (6)$$

є

$$v = v_1(x_1, t)v_2(x_2, t)v_3(x_3, t). \quad (7)$$

При підстановці виразу (7) в рівняння (1) і при використанні (3) отримаємо:

$$v_2 v_3 \frac{\partial^2 v_1}{\partial x_1^2} + v_1 v_3 \frac{\partial^2 v_2}{\partial x_2^2} + v_1 v_2 \frac{\partial^2 v_3}{\partial x_3^2} = \frac{1}{\chi} \left(v_2 v_3 \frac{\partial v_1}{\partial t} + v_1 v_3 \frac{\partial v_2}{\partial t} + v_1 v_2 \frac{\partial v_3}{\partial t} \right).$$

При цьому очевидно, що задовольняються початкові та граничні умови (4), (5) і (6).

Розглянемо також диференціальне рівняння в частинних похідних, задане в циліндричних системах координат [1]:

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial v}{\partial r} \right) + \frac{\partial^2 v}{\partial z^2} = \frac{1}{\chi} \frac{\partial v}{\partial t}, \quad (8)$$

де $v = v(r, z, t)$ – шукана функція температури; r, z, t – незалежні змінні, які визначають просторові координати та час.

Припустимо, що його необхідно розв'язати в області

$$a < r < b, \quad z_1 < z < z_2, \quad (9)$$

де a, b, z_1, z_2 – сталі.

Нехай функція $v_1 = v_1(r, t)$ є розв'язком рівняння

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial v_1}{\partial r} \right) = \frac{1}{\chi} \frac{\partial v_1}{\partial t}, \quad t > 0, \quad a < r < b, \quad (10)$$

при граничних та початкових умовах:

$$\alpha_1 \frac{\partial v_1}{\partial r} - \beta_1 v_1 = 0, \quad r = a, \quad t > 0;$$

$$\alpha_1' \frac{\partial v_1}{\partial r} + \beta_1' v_1 = 0, \quad r = b, \quad t > 0;$$

$$v_1 = V_1(r) \quad \text{при } t=0,$$

де $\alpha_1, \beta_1, \alpha_1', \beta_1'$ – сталі, кожна з яких може бути рівною нулю, $V_1(r)$ – задана функція.

Нехай також функція $v_2 = v_2(z, t)$ є розв'язком рівняння

$$\frac{\partial^2 v_2}{\partial z^2} = \frac{1}{\chi} \frac{\partial v_2}{\partial t}, \quad t > 0, \quad z_1 < z < z_2,$$

при граничних та початкових умовах:

$$\alpha_2 \frac{\partial v_2}{\partial z} - \beta_2 v_2 = 0, \quad z = z_1, \quad t > 0;$$

$$\alpha_2' \frac{\partial v_2}{\partial z} + \beta_2' v_2 = 0, \quad z = z_2, \quad t > 0;$$

$$v_2 = V_2(z) \text{ при } t=0,$$

де $\alpha_2, \beta_2, \alpha'_2, \beta'_2$ – сталі, кожна з яких може бути рівною нулю, $V_2(z)$ – задана функція.

Тоді розв'язком рівняння (8) в області (9) є:

$$v = v_1(r, t) v_2(z, t), \quad (11)$$

при граничних умовах:

$$\alpha_1 \frac{\partial v}{\partial r} - \beta_1 v = 0, \quad r = a, \quad z_1 < z < z_2, \quad t > 0,$$

$$\alpha'_1 \frac{\partial v}{\partial r} + \beta'_1 v = 0, \quad r = b, \quad z_1 < z < z_2, \quad t > 0,$$

$$\alpha_2 \frac{\partial v}{\partial z} - \beta_2 v = 0, \quad z = z_1, \quad a < r < b, \quad t > 0,$$

$$\alpha'_2 \frac{\partial v}{\partial z} + \beta'_2 v = 0, \quad z = z_2, \quad a < r < b, \quad t > 0$$

і при початковій умові

$$v = V_1(r) V_2(z) \text{ при } t = 0.$$

Такий же прийом може бути використаний і для інших областей, наприклад у випадку необмеженого прямого двогранного кута, напівобмеженого циліндра та ін.

Оцінка моделей. При пошуку розв'язків вказаних задач за допомогою чисельних методів, що побудовані на основі різницьових схем, отримуються системи лінійних алгебраїчних рівнянь, розмірність яких залежить від розбиття за просторовими та часовою змінними. Складність чисельної реалізації можна оцінити на основі кількості операцій, що необхідні для розв'язання систем алгебраїчних рівнянь [3, 5].

Так, у випадку двовимірної моделі (8), вважатимемо, що кількість точок розбиття по просторовій змінній r буде рівною M_1 , по змінній z – M_2 , по часовій змінній t – N . В результаті, при застосуванні різницьових методів, буде отримано систему лінійних алгебраїчних рівнянь розмірності $M_1 M_2 N$ для моделі (8) та $M_1 N + M_2 N = (M_1 + M_2) N$ для моделі (11). Без зменшення загальності, для спрощення наступної оцінки, припустимо, що розбиття по просторових змінних однакове: $M = M_1 = M_2$. Тобто, матимемо системи алгебраїчних рівнянь розмірності $M^2 N$ та $2 M N$ відповідно для двовимірної моделі та двох одновимірних.

Для розв'язування систем лінійних алгебраїчних рівнянь часто застосовуються такі методи, як метод Гауса, LU-розклад, метод Холецкого, метод ітерацій тощо. Кількість операцій, необхідних для використання кожного з методів при розв'язуванні n -вимірної системи лінійних рівнянь рівна: $\frac{2}{3} n^3$ – метод Гауса, $\frac{2}{3} n^3$ – LU-розклад,

$\frac{1}{3} n^3$ – метод Холецкого, $2 n^2 L$ – метод ітерацій, де L – кількість ітерацій. Оцінка складності розв'язання двовимірної задачі (8) та сукупності одновимірних задач (10) представлена в табл. 1.

Аналогічні оцінки зроблені і для випадку тривимірної моделі (1). Вважатимемо, що кількість точок розбиття по просторових змінних x_1, x_2, x_3 буде рівною M , за часовою змінною $t - N$. В результаті застосування різницьових схем отримано системи алгебраїчних рівнянь розмірності M^3N та $3MN$ відповідно для тривимірної моделі та трьох одновимірних. Оцінка складності розв'язання тривимірної задачі (1) та сукупності одновимірних задач (3) представлена в таблиці 2.

Для оцінки складності використовується відносний показник складності δ_T , який визначається, як відношення кількості операцій, що необхідні для розв'язання базової моделі та спрощених моделей.

Таблиця 1.

Оцінка складності розв'язання двовимірної задачі

Тип задачі	Метод Гауса	LU-розклад	Метод Холецкого	Метод ітерації
	Кількість операцій			
Двовимірна задача	$\frac{2}{3}M^6N^3$	$\frac{2}{3}M^6N^3$	$\frac{1}{3}M^6N^3$	$2M^4N^2L$
Дві одновимірні задачі	$\frac{2}{3}2^3M^3N^3$	$\frac{2}{3}2^3M^3N^3$	$\frac{1}{3}2^3M^3N^3$	$4M^2N^2L$
	Відносний показник складності δ_T			
	$\frac{1}{2^3}M^3$	$\frac{1}{2^3}M^3$	$\frac{1}{2^3}M^3$	$\frac{1}{2}M^2$

Таблиця 2.

Оцінка складності розв'язання тривимірної задачі

Тип задачі	Метод Гауса	LU-розклад	Метод Холецкого	Метод ітерації
	Кількість операцій			
Тривимірна задача	$\frac{2}{3}M^9N^3$	$\frac{2}{3}M^9N^3$	$\frac{1}{3}M^9N^3$	$2M^6N^2L$
Три одновимірні задачі	$\frac{2}{3}3^3M^3N^3$	$\frac{2}{3}3^3M^3N^3$	$\frac{1}{3}3^3M^3N^3$	$6M^2N^2L$
	Відносний показник складності δ_T			
	$\frac{1}{3^3}M^6$	$\frac{1}{3^3}M^6$	$\frac{1}{2^3}M^6$	$\frac{1}{3}M^4$

При реалізації двовимірної та тривимірної моделей також витрачається час для знаходження загального розв'язку на основі розв'язків одновимірних задач. Кількість операцій при цьому дорівнює, відповідно, M^2N та M^3N , але вона не впливає на порядок відносного показника складності.

На основі відносного показника порівняння складності, можна зробити висновок, що спрощення вихідних двовимірних та тривимірних задач дозволяє скоротити кількість обчислювальних операцій при знаходженні чисельних розв'язків при застосуванні прямих методів розв'язування систем лінійних алгебраїчних рівнянь в $\square M^3$ разів та в $\square M^6$ разів відповідно; у випадку застосування методу ітерацій кількість операцій скорочується в $\square M^2$ разів та в $\square M^4$ разів відповідно.

Висновки

Отримані результати мають суттєве практичне значення і можуть застосовуватись при побудові комп'ютерно-інтегрованих систем, що мають у своєму складі моделі об'єктів з розподіленими параметрами, зокрема, вони дозволяють достатньо економно отримувати чисельні значення температур у твердих тілах, якщо їх початкова температура постійна, а з поверхні відбувається теплообмін в середовище з постійною температурою.

Список літератури

1. Carslaw, H.S. Conduction of Heat in Solids / H.S. Carslaw, J.C. Jaeger. — 2 Edition. — Oxford: Oxford University Press, 1986. — 520 p.
2. Иванюк, В.А. Способы формирования передаточных функций приемника теплового потока / В.А. Иванюк, Н.Л. Костян, А.И. Махович. // Математичне та комп'ютерне моделювання. Серія: Технічні науки: зб. наук. праць / Інститут кібернетики імені В.М. Глушкова Національної академії наук України, Кам'янець-Подільський національний університет імені Івана Огієнка. — 2013. — Вип. 8. — С. 61–69.
3. Самарский, А.А. Вычислительная теплопередача / А.А. Самарский, П.Н. Вабищевич. — М.: Едиториал УРСС, 2003. — 784 с.
4. Скопецкий, В.В. Математичне моделювання прямих та обернених задач динаміки систем з розподіленими параметрами / В.В. Скопецкий, В.А. Стоян, Ю.Г. Кривонос. — К.: Наукова думка, 2002. — 361 с.
5. Самарский, А.А. Аддитивные схемы для задач математической физики / А.А. Самарский, П.Н. Вабищевич. — М.: Наука, 2001. — 319 с.
6. Verlan, A.A. An Approach to the Precision Parametric Reduction of Mathematical Models / A.A. Verlan. // Математичне та комп'ютерне моделювання. Серія: Технічні науки: зб. наук. праць / Інститут кібернетики імені В. М. Глушкова Національної академії наук України, Кам'янець-Подільський національний університет імені Івана Огієнка. — 2016. — Вип. 14. — С. 26–35.

УПРОЩЕНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ОБЪЕКТОВ С РАСПРЕДЕЛЕННЫМИ ПАРАМЕТРАМИ НА ОСНОВЕ МЕТОДА РАСЩЕПЛЕНИЯ

А.А. Верлань¹, В.А. Иванюк²

¹ Норвежский университет науки и технологий,
NTNU Gjøvik, Teknologiveien 22, 2815 Gjøvik, Norway; e-mail: andriy.verlan@ntnu.no

² Каменец-Подольский национальный университет имени Ивана Огиенка,
ул. Уральская, 1, Каменец-Подольский, 32300, Украина; e-mail: wivanyuk@gmail.com

В статье рассмотрены методы построения упрощенных моделей многомерных задач теплопроводности в виде произведения решений одномерных задач. Эффективность предложенного подхода представлено на основе анализа сложности численной реализации данных моделей с помощью разностных методов.

Ключевые слова: дифференциальные уравнения в частных производных, задача теплопроводности, эквивалентные преобразования

SIMPLIFICATION OF MATHEMATICAL MODELS OF OBJECTS WITH DISTRIBUTED PARAMETERS BY THE METHOD OF SPLITTING

A.A. Verlan¹, V.A. Ivanyuk²

¹ Norwegian University of Science and Technology,
NTNU Gjøvik, Teknologiveien 22, 2815 Gjøvik, Norway; e-mail: andriy.verlan@ntnu.no

² Kamyanyets-Podilskyi Ivan Ohienko National University,
1, Uralska Str., Kamyanyets-Podilskyi, 32300, Ukraine; e-mail: wivanyuk@gmail.com

The article deals with the methods of constructing simplified models of multidimensional heat conduction problems in the form of products of solutions of one-dimensional problems. The effectiveness of the proposed approach is shown based on the analysis of the complexity of numerical implementation of these models.

Keywords: differential equations in partial derivatives, heat conduction problem, equivalent transformations

ПАРАМЕТРИЧЕСКИЙ СИНТЕЗ ДИСКРЕТНО-АДАПТИВНОЙ ОДНОКАНАЛЬНОЙ ПО ДОПЛЕРОВСКОЙ ЧАСТОТЕ СИСТЕМЫ ОБРАБОТКИ СИГНАЛА В УСЛОВИЯХ ГАУССОВЫХ ПОМЕХ

И.В. Цевух, А.В. Соколов, А.А. Сакович

Одесский национальный политехнический университет,
просп. Шевченко, 1, Одесса, 65044, Украина; e-mail: itsevukh@opu.ua

В работе предложена методика минимизации числа режимов и оптимизации параметров одноканальной по доплеровской частоте дискретно-адаптивной системы (ДАС) обработки сигналов на фоне гауссовых помех при дополнительном ограничении на величину потерь в эффективности по сравнению с системами с непрерывной перестройкой параметров. Разработана с использованием теоретико-графовой модели процедура параметрического синтеза ДАС. Приведены результаты математического моделирования одноканальной по доплеровской частоте ДАС обработки сигнала в условиях аддитивной смеси коррелированной и некоррелированной гауссовых помех.

Ключевые слова: отношение правдоподобия, достаточная статистика, ковариационная матрица

Введение

Одним из основных требований, предъявляемых к существующим и проектируемым многофункциональным системам передачи и извлечения информации (СПИИ), является высокая помехозащищенность в условиях сложной, изменчивой и априорно неизвестной помеховой обстановки. Известно [1], что эффективные системы обработки полезного сигнала в условиях помех должны синтезироваться с использованием теста отношения правдоподобия. Полученные таким образом алгоритмы минимизируют, в определенном смысле, риск, связанный с применением неправильных решений, что немаловажно при построении систем, на выходе которых принимается решение о наличии или отсутствии в исходном процессе полезного сигнала. Применение оптимальных алгоритмов для решения конкретных радиотехнических задач осложняется в силу специфики, заключающейся в невозможности достаточно обоснованного задания вероятностных или даже спектрально-корреляционных характеристик сигнала и помехи, а также требований к реализации алгоритмов в реальном масштабе времени при сохранении их устойчивости к асимптотически оптимальным решениям.

Сложность практической реализации адаптивных оптимальных обнаружителей полезного сигнала с непрерывной перестройкой параметров в зависимости от помеховой обстановки делает актуальным поиск упрощенных вариантов построения и расчет параметров таких систем.

Цель статьи и постановка задач исследования

Целью статьи является проведение параметрического синтеза упрощенной, с точки зрения практической реализации, одноканальной по доплеровской частоте дискретно-адаптивной системы обработки сигналов при дополнительном ограничении на величину потерь в эффективности по сравнению с системами с непрерывной перестройкой параметров.

Для достижения цели необходимо решить следующие задачи:

1. Разработать упрощенный квазиадаптивный алгоритм обработки сигналов в условиях априорно неизвестной помеховой обстановки;
2. Предложить методику минимизации числа режимов и оптимизации параметров одноканальной по доплеровской частоте дискретно-адаптивной системы обработки сигналов при дополнительном ограничении на величину потерь в эффективности по сравнению с системами с непрерывной перестройкой параметров;
3. Провести параметрический синтез одноканальной по доплеровской частоте дискретно-адаптивной системы обработки сигналов в условиях аддитивной смеси коррелированной и некоррелированной гауссовых помех.

Основная часть

Разработка упрощенного варианта построения системы обработки сигнала в условиях сложной и априорно неизвестной помеховой обстановки. Для упрощения построения системы обработки предположим, что априорная неопределенность спектрально-корреляционных характеристик процесса на входе СПИИ носит характер параметрической, а значит одним из путей ее преодоления может быть использование адаптивного Байесова подхода. В этом случае, при допустимости гауссовой аппроксимации плотности распределения вероятности помехи, приемлемыми с точки зрения скорости сходимости оказываются пути, основанные на идее замены в полученных при априорной определенности алгоритмах, корреляционной матрицы помехи ее состоятельной оценкой [2]. С учетом этого в рамках адаптивного Байесового подхода заменим фиксированные параметры ковариационной матрицы помехи B_{Π} их оценками максимального правдоподобия, полученными из обучающей выборки в отсутствии полезного сигнала.

На втором этапе поиска упрощенного варианта построения используем тот факт, что для ряда радиотехнических приложений допустимым оказывается применение вместо оптимальной многоканальной системы обработки - упрощенной системы одноканальной по доплеровской частоте сигнала. Для построения таких одноканальных систем, решающих задачу обнаружения импульсного сигнала на фоне гауссовых помех с неизвестными корреляционными свойствами, в [3] и [4] предложено использовать, соответственно, достаточные статистики

$$T^2 = X^* \hat{B}_{\Pi}^{-1} X ; \quad (1)$$

$$C^2 = X^* \hat{B}_{\Pi}^2 X , \quad (2)$$

где X - N -мерный вектор выборочных отсчетов входного процесса, $\hat{B}_{\Pi}$ - оценка максимального правдоподобия ковариационной матрицы помехи, $*$ - знак эрмитова сопряжения и транспонирования. В [4] показано, что эффективность алгоритма, использующего C^2 статистику, по вероятностным показателям для гауссовых моделей сигнала и помех оказывается несколько выше алгоритма, использующего T^2

статистику Хотеллинга. Тем не менее, построение систем обработки на базе (1) и (2) затруднено из-за необходимости измерения и преобразования ковариационной матрицы помехи, что требует большого числа вычислений, а для (2) еще и дополнительного, по сравнению с (1), умножения на оценку обратной ковариационной матрицы помехи.

И, наконец, на третьем этапе для уменьшения вычислительной сложности перейдем от адаптивной системы с непрерывной перестройкой параметров (СНП), к дискретно-адаптивной системе [5], состоящей из совокупности неадаптивных подсистем (режимов), автоматически переключающихся в зависимости от результатов анализа помеховой обстановки. Обеспечим при этом контролируемый уровень потерь в эффективности такой ДАС по сравнению с СНП.

Разработка методики параметрического синтеза ДАС. При гауссовой аппроксимации плотности распределения вероятностей аддитивной смеси действующих на входе приемника помех, любую из возможных помеховых ситуаций, принадлежащих некоторой области θ , можно полностью характеризовать значением элементов ковариационной матрицы $B_{\Pi} \in B(\theta)$. Основная идея перехода от систем обработки с непрерывной к системам с дискретной перестройкой параметров заключается в разбиении непрерывной области θ на минимально возможное число M неперекрывающихся подобластей, в каждой из которых потеря в эффективности ДАС относительно СНП не превышает заданной величины C_0 (1-й этап):

$$M \rightarrow \min_{B_{\Pi} \in B(\theta)} = \dot{M};$$

при ограничении

$$C(B_{\Pi}, W_{gi}) \leq C_0;$$

$$i = k, \text{ если } B_{\Pi} \in B(\theta_k),$$

где $C(B_{\Pi}, W_{gi})$ – функционал потерь, вид которого определяется выбранным критерием оптимальности системы обработки G , $B(\theta_k)$ – совокупность всех возможных ковариационных матриц помехи подобласти θ_k , W_{gi} – вектор параметров систем обработки, определяемый ковариационной матрицей B_{gi} .

При этом координаты «центров» (определяющие параметры режимов ДАС) и границы подобластей θ_i , $i = \overline{1, M}$, должны выбираться так, чтобы доставлять минимум средним по области θ потерям (2-й этап):

$$\int_{B(\theta)} C'(B_{\Pi}, W_{gi}) \rightarrow \min_{B_{\Pi} \in B(\theta)},$$

при условии, что

$$C'(B_{\Pi}, W_{gi}) \rightarrow \min_{j=1, M} C(B_{\Pi}, W_{gi}).$$

С учетом этого, правило выбора вектора параметров ДАС может быть записано в следующем виде:

$$W_{gi} = Z_g H_i,$$

где

$$H_i = [h_1 = 0, \dots, h_{i-1} = 0, h_i = 1, h_{i+1} = 0, h_M = 0],$$

если $B_{\Pi} \in B(\theta_i)$; Z_g – матрица, состоящая из $\dot{M}$ векторов параметров ДАС.

Отметим, что пространство θ образует в общем случае бесконечное множество помеховых ситуаций. При численном решении задачи параметрического синтеза необходимо будет перейти к конечным множествам, разбивая пространство входных переменных сеткой из w опорных точек (метод сеток).

Конкретизируем вид критериев оптимизации для каждого из этапов предложенной процедуры параметрического синтеза в соответствии с выбранной моделью дискретно-адаптивной системы обработки сигналов. С учетом [4,6], формальное представление задач поиска минимального числа режимов помехозащиты и оптимизации их параметров имеет вид:

1-й этап:

$$M \rightarrow \min_{B_{\Pi} \in B(\theta)} = \dot{M};$$

при ограничении

$$C(B_{\Pi_j}, B_{\Pi_k}^{-2}) = \frac{\sqrt{\text{Tr} B_{\Pi_j}^{-2} \text{Tr} B_{\Pi_k}^{-2}}}{\sqrt{\text{Tr}(B_{\Pi_k}^{-2} B_{\Pi_j})^2}} \leq C_0;$$

$$j = \overline{1, w}; \quad k = \overline{1, \dot{M}},$$

где $B_{\Pi_k}^{-2} = ZH_k$, Z – блочная матрица, составленная из $\dot{M}$ матриц обработки ДАС, при синтезе которой использовалась достаточная статистика C^2 ; $\text{Tr}[\cdot]$ – след матрицы.

2-й этап:

$$\dot{V} = \min_{V \in V(\theta)} Q(V);$$

$$Q(V) = \frac{1}{w} \sum_{i=1}^w C'(B_{\Pi_j}, B_{\Pi_k}^{-2});$$

$$C'(B_{\Pi_i}, B_{\Pi_k}^{-2}) = \min_{\forall n \neq k} C(B_{\Pi_i}, B_{\Pi_n}^{-2}).$$

В столь общей постановке оптимизационная задача первого этапа является NP -полной, что говорит о невозможности, в общем случае, получения точных оптимальных решений за полиномиальное время без помощи недетерминированных алгоритмов, то есть о необходимости для любого детерминированного алгоритма, корректно решающего данную NP -полную задачу, в худшем случае, экспоненциального количества времени и, следовательно, о его неприменимости на практике ни к каким, за исключением очень малой размерности, индивидуальным задачам. Понимание сложности задачи имеет ряд преимуществ: с одной стороны – показывает безнадёжность попыток алгоритмически решить задачу, а с другой – открывает пути для достижения менее амбициозных целей, чем построение алгоритма,

который всегда находит точное решение и время работы которого никогда не превышает полиномиальной оценки.

С учетом этого, и пользуясь рекомендациями [7] для практического решения комбинаторных задач подобного типа, используем приближенный метод. Одним из самых удобных и широко используемых при проектировании сложных систем, в которых естественным образом можно выделить связи между элементами, является теоретико-графовая модель представления данных.

В рамках указанной модели ограниченный ансамбль комбинаций, «помеховая ситуация – режим помехозащиты» может быть представлен с помощью двудольного неориентированного графа $\Gamma(U, V, E)$, в котором множество вершин $U = \{u_1, \dots, u_L\}$ соответствуют L ($L \gg M$) помеховым ситуациям с ковариационными матрицами B_{u_j} , $j = \overline{1, L}$, а координаты вершин множества $V = \{v_1, \dots, v_\mu\}$ однозначно определяют параметры $\mu \gg M$ режимов работы ДАС.

Пусть H – некоторая нумерация на множестве вершин графа $\Gamma(U, V, E)$. Номер вершины u_j при этом обозначается $H(u_j)$, а номер вершины v_i – $H(v_i)$.

Пусть также вершина u_j достигается из v_i , если выполняется условие

$$C(B_{u_j}, W_{v_i}) \leq C_0.$$

Тогда решение задачи сводится к построению подграфа $\Gamma_\Pi(U, \dot{S}, \dot{E})$ с минимально возможным множеством вершин $\dot{S}$, таким, что

$$\dot{S} \in V, |\dot{S}| = M,$$

а для любой вершины $u_j \in U$ существует единственное ребро $(u_j, s_i) \in \dot{E}$, вес которого

$$C(B_{u_j}, W_{v_i}) = \min_k C(B_{u_j}, W_{v_k}).$$

Предлагается следующий алгоритм построения подграфа $\Gamma_\Pi(U, \dot{S}, \dot{E})$.

Вход: граф $\Gamma(U, V, E)$. **Выход:** подграф $\Gamma_\Pi(U, \dot{S}, \dot{E})$.

Начало.

Шаг 1. $M := 0$, $S := \emptyset$.

Шаг 2. Выбор в V вершины

$$V_m = \arg \max_{j=H_v, j \neq m} \deg(V_j),$$

где $\deg(V_j)$ – степень вершины. Неопределенность разрешается произвольно.

Шаг 3. Удалить из $\Gamma(U, V, E)$ найденную вершину V_m и все смежные с ней вершины из U .

Шаг 4. Добавить V_m в S и увеличить $M := M + 1$.

Шаг 5. Если $U \neq \emptyset$, то переход к шагу 2.

Шаг 6. $j := 0$, $N := M$.

Шаг 7. $j := j + 1$, если $j > M$ то переход к шагу 11.

Шаг 8. Удалить из S вершину $V_{H_{S(j)}}$.

Шаг 9. При отсутствии в $\Gamma(U, V, E)$ на данной итерации изолированных вершин уменьшить $N := N + 1$ и перейти к шагу 7.

Шаг 10. Включить $V_{H_{S(j)}}$ в S и перейти к шагу 7.

Шаг 11. $\dot{M} := N$, $\dot{S} := S$.

Шаг 12. Вершина $u_j \in U$ смежна вершине $s_i \in \dot{S}$, если выполняется условие

$$C(B_{u_j}, W_{v_i}) = \min_k C(B_{u_j}, W_{v_k}).$$

Конец .

При практической реализации алгоритма для больших мощностей множеств V и U существенный вопрос заключается в том, каким образом хранить данные в памяти ПК, потому что способ хранения описаний ограниченного ансамбля комбинаций «помеховая ситуация – режим помехозащиты» (вершин двудольного графа) и отношений между ними (ребер графа) влияет на возможность и скорость получения решения поставленной задачи. Указанные отношения могут задаваться с помощью матриц смежности, инцидентности, достижимости, списков смежности, списков ребер. Проведенный анализ показал целесообразность представления графа $\Gamma(U, V, E)$ в памяти ПК с помощью бинарной матрицы инцидентности, так как в этом случае учет априорной информации о двудольном характере данного графа позволяет существенно сократить объем используемой памяти по сравнению с необходимым для представления графа произвольной конфигурации.

Элементы d_{ij}^u матрицы инцидентности $D_u[\Gamma(U, V, E)]$ определяются следующим образом:

$$d_{ij}^u = \begin{cases} 1, & \text{если существует ребро } (u_j, v_i); \\ 0, & \text{в противном случае.} \end{cases}$$

Временная сложность разработанного алгоритма и его эффективность в значительной степени зависят от конфигурации вводимого графа $\Gamma(U, V, E)$, а его емкостная сложность определяется размерами матрицы инцидентности $D_u[\Gamma(U, V, E)]$ и, несмотря на экономию, достигаемую за счет учета двудольного характера графа $\Gamma(U, V, E)$, ограничивает применение этих алгоритмов объемом памяти ПК, имеющегося у разработчика.

Результаты параметрического синтеза одноканальной по доплеровской частоте ДАС сигнала в условиях аддитивной смеси некоррелированной (НП) и унимодальной коррелированной гауссовых помех. С помощью программной реализации предложенной методики проведем параметрический синтез одноканальной по доплеровской частоте ДАС сигнала в условиях характерной для некоторых типов СПИИ помеховой обстановки, когда отношение некоррелированная помеха(НП)/коррелированная помеха(КП) изменяется в пределах от 0 до -40дБ, форма спектра флуктуаций КП – от резонансной до гауссовой, а его относительная полуширина от 0,05 до 0,2. Предположим, что адаптация к аргументу корреляционной функции помехи носит непрерывный характер. Такое уточнение области помеховых ситуаций позволяет существенно снизить время расчетов за счет устранения в процессе формирования множеств U и V графа $\Gamma(U, V, E)$ операции проверки генерируемых матриц B_{u_j} , $j = \overline{1, L}$, на принадлежность к классу ковариационных.

Результаты расчетов числа подобластей (режимов помехозащиты $\dot{M}$), требуемого для обеспечения заданного уровня максимальных потерь (C_0), и значений средних, относительно СНП, потерь в эффективности ($\overline{\Delta C}$) в зависимости от размерности (N)

вектора входного процесса ДАС, синтезированной с использованием решающей статистики (2), приведены в таблице 1.

Таблица 1.
Результаты параметрического синтеза одноканальной по доплеровской частоте ДАС

N	C_0, dB	$\dot{M}$	$\overline{\Delta C}, dB$
3	1	12	0,093
	2	9	0,148
	3	7	0,210
	4	5	0,281
	5	3	0,409
4	1	11	0,430
	2	8	0,453
	3	6	0,592
	4	4	0,705
	5	3	0,767
5	1	9	0,89
	2	7	0,913
	3	5	0,937
	4	3	1,179
	5	3	1,179

Из таблицы 1 следует, что с увеличением N сокращается число режимов ДАС: при $C_0 = 3$ дБ $\dot{M}$ уменьшается с семи для $N=3$ до пяти для $N=5$. Это можно объяснить тем, что с увеличением N разность между значениями условных математических ожиданий статистики $C^2 \{M[C^2 / \text{сигнал присутствует}] - M[C^2 / \text{сигнал отсутствует}]\}$ растет быстрее примерно в $\sqrt{N}$ раз, чем условное среднеквадратическое отклонение $D^{1/2}[C^2 / \text{сигнал отсутствует}]$, а следовательно, снижается влияние “хвостов” распределения $p[C^2 / \text{сигнал отсутствует}]$ на величину показателя качества обнаружения.

Для наглядности в таблице 2 приведены результаты расчетов значений элементов нижней треугольной матрицы обработки Q (разложение Холецкого $B_{\Pi_k}^{-2} = QQ^T$) одноканальной по доплеровской частоте ДАС обработки сигнала, синтезированной с использованием достаточной статистики C^2 для $N = 3$ и $C_0 = 3$ дБ.

Таблица 2.
Значения элементов нижней треугольной матрицы обработки Q одноканальной ДАС

C_0, dB	Номер режима	Q			$\overline{\Delta C}, dB$
3	1	1024	-2026	1024	0,21
		0	2865	-1448	
		0	0	2047	
	2	458,2	-902,9	457,3	
		0	25,24	-25,57	
		0	0	0,701	

Продолжение таблицы 2.

C_0, dB	Номер режима	Q			$\overline{\Delta C}, dB$
3	3	273,9	-537,4	273,1	0,21
		0	19,82	-20,17	
		0	0	0,5	
	4	724,1	-1417	723,9	
		0	16,03	-16,39	
		0	0	0,612	
	5	63,81	-123,0	62,59	
		0	12,11	-12,42	
		0	0	0,605	
	6	149,4	-286,2	149,2	
		0	8,231	-8,540	
		0	0	0,625	
	7	8,191	-13,94	6,709	
		0	4,440	-4,652	
		0	0	0,661	

Выводы

В работе предложена методика и проведен параметрический синтез одноканальной по доплеровской частоте сигнала ДАС, в которой, для сокращения числа режимов был сохранен непрерывный характер адаптации к аргументу корреляционной функции помехи. Показано, что синтезированная, в соответствии с разработанной методикой ДАС для заданных ограничений на максимальные потери позволяет, при сравнительно небольшом числе режимов помехозащиты (не более семи для $C_0 \geq 3\text{дБ}$, $N \geq 3$), обеспечить, в предположении безошибочной классификации помеховой обстановки, незначительные (менее 1дБ для $C_0 \geq 3\text{дБ}$, $N \geq 3$) средние относительно СНП потери в эффективности обработки сигналов на фоне аддитивной смеси некоррелированной и унимодальной коррелированной помех.

Список литературы

1. Merrill, I. Skolnik Radar Handbook. Vol. 1. – 3 Edition. — The McGraw-Hill Companies, 2008. — 273 p. — ISBN: 9780071485470.
2. Levy, B.C. Principles of Signal Detection and Parameter Estimation N. — Y.: Springer, 2008. — 642 p.
3. Бартенев, В.Г. О построении адаптивного обнаружителя импульсных сигналов на фоне нормальных помех с неизвестными корреляционными свойствами / В.Г. Бартенев, А.М. Шлома. // Радиоэлектроника. — 1988. — № 2. — С. 3–8.
4. Цевух, И.В. Алгоритм обработки гауссовых сигналов условиях гауссовых помех / И.В. Цевух. // Радиоэлектроника. — 1988. — № 12. — С. 53–54.
5. Maurice, G. Bellanger Adaptive digital filters. — Marcel Dekker Inc, 2001. — 464 p.
6. Бакулев, П.А. Радиолокационные системы. — М: Радиотехника, 2004. — 320 с.
7. Avis, D. Graph theory and combinatorial optimization / D. Avis, A. Hertz, O. Marcotte. — Y.: Springer, 2005. — 272 p.

**ПАРАМЕТРИЧНИЙ СИНТЕЗ ДИСКРЕТНО-АДАПТИВНОЇ ОДНОКАНАЛЬНОЇ
З ДОПЛЕРОВСЬКОЇ ЧАСТОТИ СИСТЕМИ ОБРОБКИ СИГНАЛУ В УМОВАХ
ГАУССОВИХ ЗАВАД**

І.В. Цевух, А.В. Соколов, А.А. Сакович

Одеський національний політехнічний університет,
просп. Шевченка, 1, Одеса, 65044, Україна; e-mail: itsevukh@opu.ua

В роботі запропонована методика мінімізації числа режимів і оптимізації параметрів одноканальної з доплеровської частоти дискретно-адаптивної системи (ДАС) обробки сигналів на тлі гаусових завад при додатковому обмеженні на величину втрат в ефективності в порівнянні системами з безперервною зміною параметрів. Розроблено, з використанням теоретико-графової моделі, процедуру параметричного синтезу ДАС. Наведено результати математичного моделювання одноканальної з доплеровської частоти ДАС обробки сигналу в умовах адитивної суміші корельованої та некорельованої гаусових завад.

Ключові слова: відношення правдоподібності, достатня статистика, коваріаційна матриця

**PARAMETRIC SYNTHESIS OF ONE-CHANNEL ON DOPPLER FREQUENCY
OF DISCRETE-ADAPTIVE PROCESSING SYSTEM FOR SIGNAL IN GAUSSIAN NOISES**

I.V. Tsevukh, A.V. Sokolov, A.A. Sakovich

Odesa National Polytechnic University,
1, Shevchenko Ave., Odesa, 65044, Ukraine; e-mail: itsevukh@opu.ua

In this paper a technique for minimizing the number of regimes and optimizing the parameters of a discrete-adaptive signal processing system (DAS) in conditions of gaussian noises is proposed, with an additional restriction on the magnitude of efficiency losses compared to systems with continuous parameters variation. A procedure for parametric DAS synthesis was developed using the graph-theoretic model. Results of the mathematical modeling of a single-channel on doppler frequency DAS in the conditions of an additive mixture of correlated and uncorrelated gaussian noises are presented.

Keywords: likelihood ratio, sufficient statistics, covariance matrix

МОДЕЛЮВАННЯ ТА ОПТИМІЗАЦІЯ ДОСТУПУ ДО СТОРІНОК WEB-САЙТУ ДЛЯ РІЗНИХ ЗАКОНІВ РОЗПОДІЛУ ЙМОВІРНОСТЕЙ ЗВЕРТАННЯ ДО СТОРІНОК

М.І. Юськів, Г.Г. Цегелик

Львівський національний університет ім. Івана Франка,
вул. Університетська, 1, Львів, 79000, Україна; e-mail: markiyau_yuskiv@live.com

У статті розглянуто підхід до математичного моделювання оптимального доступу користувачів до послідовно організованих сторінок Web-сайту. Знайдено вираз для математичного очікування загального часу, необхідного для пошуку сторінки, для різних законів розподілу ймовірності звертання до сторінок. Виведено співвідношення для знаходження параметрів, за яких математичне очікування досягає мінімуму. Запропоновано підходи для реалізації даного алгоритму у прикладних додатках. На основі побудованої математичної моделі оптимального доступу до сторінок інформаційного Web-сайту зі сторони користувача та отриманого співвідношення для знаходження параметрів, за яких математичне очікування досягає мінімуму, можна зробити висновок, що дане дослідження має потенціал для подальшого застосування його результатів на виробництві. Зокрема для побудови високопродуктивних модулів для реалізації пошуку у додатках з хмарною інфраструктурою.

Ключові слова: моделювання доступу до сторінок, Web-сайт, закони розподілу ймовірностей, мікросервісна архітектура, оптимізація доступу

Вступ

Мережа Інтернет є унікальним інформаційним ресурсом, через який кожен день проходять великі об'єми інформації. Основними складовими цього ресурсу є сервери та Web-сайти. Кожен сайт відповідно складається з набору сторінок. Сторінки сайту – атомарні одиниці, які між собою пов'язані логічними зв'язками. Відповідно, об'єктом дослідження є процес пошуку інформації на Web-сайті, де сторінки розбиті на блоки з використанням методу послідовного читання блоків користувачем та їхнього послідовного перегляду. Предметом дослідження виступає оптимальність вибраного методу та можливість його реалізації у промислових проектах.

Метою дослідження є побудова математичної моделі для заданого способу пошуку інформації на сайті та визначення ефективності пошуку інформації з використанням методу послідовного читання та перегляду блоків та сторінок у блоці для подальшої реалізації даного методу в рамках пошукової компоненти у додатку з мікросервісною архітектурою; визначення можливості подальшого використання результатів даного дослідження на виробництві.

Розглянуто організацію і перегляд сторінок сайту при заданих ймовірностях звертання до сторінок. За критерій ефективності пошуку сторінки приймається математичне очікування загального часу, необхідного для пошуку потрібної сторінки користувачем. Серед законів розподілу ймовірностей звертання до сторінок розглядаються рівномірний, бінарний, узагальнений та закон Зіпфа [1,2].

Основна частина

Припустимо, що інформація, яка міститься на Web-сайті, розміщена на N сторінках, розбитих на n блоків по m сторінок в кожному ($N = n \times m$). Пошук потрібної сторінки відбувається шляхом послідовного читання блоків користувачем і їх послідовного перегляду. Тоді математичне очікування загального часу, необхідного для пошуку сторінки на Web-сайті E виразиться наступною формулою:

$$E = \sum_{i=1}^n \sum_{j=1}^m (ia + ((i-1)m + j)t) p_{(i-1)m+j},$$

де $a = b + dm$ – час читання блоку сторінок користувачем; b, d – деякі сталі; p_i – ймовірність звертання до i -ої сторінки; t – середній час перегляду однієї сторінки користувачем.

Знайдемо явний вираз для E як для рівномірного закону розподілу ймовірностей звертання до сторінок, так і для таких законів нерівномірного розподілу ймовірностей як [1,3]:

– “бінарного” закону: $p_i = \frac{1}{2^i}$, $i = \overline{1, N-1}$, $p_N = \frac{1}{2^{N-1}}$;

– закону Зіпфа: $p_i = \frac{1}{iH_N}$, $i = \overline{1, N}$, $H_N = \sum_{k=1}^N \frac{1}{k}$;

– узагальненого закону розподілу: $p_i = \frac{1}{i^c H_N^{(c)}}$, $i = \overline{1, N}$, $H_N = \sum_{k=1}^N \frac{1}{k^{(c)}}$, де c ($0 < c < 1$).

Серед найпоширеніших методів оптимізації пошуку інформації на сайті, які використовуються у світовій практиці та згадуються у науковій періодиці, виділяють наступні види:

- пошук інформації з використанням теорії графів, що знайшло своє відображення у пошуку, реалізованому в мережі Facebook, та згадується у джерелі [1];
- розбиття сайту та сторінок на атомарні одиниці, що знаходять відображення у джерелах [4,5];
- використання експертних систем для повнотекстового пошуку з використанням доменів та попередніх результатів пошуку користувачем інформації на сайті [7,8];
- оптимізація пошукових запитів [8];
- кешування найпопулярніших запитів.

Продуктивність обчислювальних систем визначається ефективністю методів пошуку інформації на сайті. Оскільки в більшості систем опрацювання інформації типовими є випадки нерівномірного розподілу ймовірностей звертання до сторінки, то підхід використаний в даній статті, дає змогу не лише дослідити ефективність методу пошуку інформації для конкретного закону розподілу, а й мати залежності ефективності методу від зміни закону розподілу ймовірностей звертання до сторінок. Деякі часткові результати дослідження, одержані зарубіжними авторами, відображенні в джерелі [1]. Проте, як показує практика реальних проектів, не існує єдиного методу для покращення швидкості пошуку інформації на сторінці. Навпроти, поєднання покращення швидкодії алгоритмів, використання розподілених хмарних систем, оптимізація пошукових запитів та, при потребі, їх кешування на сьогоднішній день становлять основу підходів, які спільно використовуються для підвищення швидкодії пошуку інформації на сторінках Web-сайту.

Розглянемо докладно формальне визначення й властивості для різних законів розподілу ймовірностей звертання до сторінок Web-сайту.

У випадку рівномірного закону розподілу ймовірностей звертання до сторінок отримаємо наступний вираз для E :

$$E = \frac{1}{2}((n+1)a + (N+1)t),$$

або

$$E = \frac{1}{2} \left(\left(\frac{N}{m} + 1 \right) (b + dm) + (N+1)t \right).$$

Функція E досягає мінімуму при $m = \left(\frac{Nb}{d} \right)^{1/2}$. Тоді $n = \left(\frac{Nd}{b} \right)^{1/2}$.

У випадку, коли ймовірності звертання до сторінок задовольняють “бінарний” закон, отримаємо наступну формулу:

$$E = \left(\frac{2^m}{2^m - 1} a + 2t \right) (1 - 2^{-N}).$$

Якщо знехтувати величиною 2^{-N} , то з достатньо високою точністю можемо прийняти:

$$E = \left(\frac{2^m}{2^m - 1} a + 2t \right).$$

Для знаходження значення параметра m , при якому функція E досягає мінімуму, отримуємо рівняння: $2^m = 1 + \left(\frac{b}{d} + m \right) \ln 2$.

Якщо ймовірності звертання до сторінок розподілені за законом Зіпфа, то

$$E = \frac{1}{H_N} \left(((n+1)H_N - S_m(n))(a + mt) + \left(\left(\frac{1}{n} S_m(n) - H_N + 1 \right) N - mH_N \right) t \right),$$

де

$$S_m(n) = \sum_{k=1}^n H_{km}.$$

Використовуючи апроксимацію $S_m(n)$ функцією $\bar{S}_m(n)$,

$$\bar{S}_m(n) = n(H_N - 1) + \frac{1}{2} \ln n + C_1,$$

де

$$C_1 = \frac{1}{2} \ln 2\pi,$$

отримуємо:

$$E = \frac{1}{H_N} \left(\left(H_N + n - \frac{1}{2} \ln n - \frac{1}{2} \ln 2\pi \right) a + Nt \right).$$

Для знаходження значення параметра n , при якому функція E досягає мінімуму, отримуємо наступне рівняння: $(2n-1)n = \frac{Nd}{b} (2H_N - 2C_1 + 1 - \ln n)$.

У випадку, якщо розподіл ймовірностей звертання до сторінок задовольняє узагальнений закон розподілу, отримуємо:

$$E = \frac{1}{H_N^{(c)}} \left(\left((n+1)H_N^{(c)} - S_m^{(c)}(n) \right) (a + mt) + \left(H_N^{(c-1)} + mS_m^{(c)}(n) - NH_N^{(c)} - mH_N^{(c)} \right) t \right),$$

де

$$S_m^{(c)}(n) = \sum_{k=1}^n H_{km}^{(c)}.$$

Використовуючи апроксимацію $S_m^{(c)}(n)$ функцією $\bar{S}_m^{(c)}(n)$, де

$$\bar{S}_m^{(c)}(n) = nH_N^{(c)} + \frac{N^{1-c}}{1-c} \left(\frac{c-1}{2-c} n + \frac{a^{(c)}(n)}{n^{1-c}} \right),$$

з достатньо високою точністю отримуємо

$$E = \frac{1}{H_N^{(c)}} \left(\left(H_N^{(c)} - \frac{N^{1-c}}{1-c} \right) \left(\frac{c-1}{2-c} n + \frac{a^{(c)}(n)}{n^{1-c}} \right) \right) \left(b + \frac{dN}{n} \right) + H_N^{(c-1)} t.$$

Якщо похідну від функції $a^{(c)}(n)$ за змінною n замінити скінченною різницею $a^{(c)}(n+1) - a^{(c)}(n)$, то для наближеного знаходження параметра n , при якому функція E досягає мінімуму, отримаємо рівняння:

$$\begin{aligned} n^{3-c} + (2-c) \left(n + \frac{2-c}{1-c} \frac{Nd}{b} \right) a^{(c)}(n) = \\ = (2-c) \frac{d}{b} N^c n^{1-c} + H_N^{(c)} + \frac{2-c}{1-c} n \left(n + \frac{Nd}{b} \right) (a^{(c)}(n+1) - a^{(c)}(n)). \end{aligned}$$

У таблиці 1 наведено значення параметрів n , за яких математичне сподівання досягає мінімуму для рівномірного, бінарного, закону Зіпфа та узагальненого (в якому $0 < c < 1$), законів розподілу ймовірностей звертання до записів, деяких значень b/d , $t/d=0.1$ та $N=10^6$.

Використання результатів отриманого дослідження дає можливість на основі математичної моделі розрахувати математичне очікування для методу послідовного читання блоків користувачем і їх послідовного перегляду для різних законів розподілу ймовірностей звертання до сторінок. На основі вхідних даних ми можемо побачити, наскільки метод ефективний, та розробити подальші кроки для удосконалення алгоритмів пошуку інформації на Web-сайті.

Таблиця 1.

Значення параметрів n

b/d	Рівномірний	$c = 0,2$	$c = 0,4$	$c = 0,6$	$c = 0,8$	Закон Зіпфа	Бінарний
10	317,24	335,3	367,85	429,53	580,47	1025,5	297716,10
100	100,00	106,5	117,3	138,3	188,4	333,3	160730,1
1000	31,7	32,25	37,6	44,5	61,5	108,5	105792,4

До сильних сторін даного дослідження належить те, що без введення додаткових доробок виведений метод можна застосувати у реальних проектах. Для цього найкращим підходом буде розробка REST-сервіса як частини Web-додатку з мікросервісною архітектурою. Завданням цього сервіса є організація пошуку інформації на сайті. Перевагою такого підходу є можливість окремого тестування функціоналу пошуку, що, в свою чергу, не буде впливати на функціонал основної частини Web-сайту та її тестування. Іншою перевагою саме такого підходу буде можливість використовувати розпаралелювання роботи та збільшення пропускну та обчислювальної спроможності через використання хмарних платформ, таких як Microsoft Azure, які дають багато можливостей для хостингу, тестування, діагностики та побудови метрик для REST-сервісів.

Висновки

Розглянуто актуальну проблему оптимізації пошуку інформації на Web-сайті з використанням методу послідовного перегляду.

Наукова новизна полягає у тому, що побудовано математичну модель оптимального доступу до сторінок інформаційного Web-сайту зі сторони користувача. За критерій оптимальності прийнято математичне очікування загального часу, необхідного для послідовного пошуку сторінки на Web-сайті. Математична модель враховує ймовірності звертання до сторінок, час читання блоку сторінок та час перегляду сторінок користувачем. Знайдено вирази для математичного очікування, залежні від різних законів розподілу ймовірності звертання до сторінок. Також виведено співвідношення для знаходження параметрів, за яких математичне очікування досягає мінімуму.

Практична цінність даного дослідження у тому, що на основі отриманих даних можна побудувати програмний мікросервіс, який буде частиною додатку для оптимізації пошуку інформації на сайті.

Список літератури

1. Кнут, Д. Искусство программирования для ЭВМ. Т.3: Сортировка и поиск. – 2 изд. — М.: Издательский дом “Вильямс”, 2013. — 824 с.
2. Цегелик, Г.Г. Ефективність методу r -рівневого блочного пошуку для різних законів розподілу ймовірностей звертання до записів / Г.Г. Цегелик, А.В. Мельничанин. // Вісн. НУ “Львівська Політехніка”. Сер. інформ. Системи та мережі. — 2005. — № 549. — С. 184-192.
3. Цегелик, Г.Г. Математичне моделювання та оптимізація до інформації індексно-послідовних файлів баз даних / Г.Г. Цегелик, А.В. Мельничанин. // Волин. матем. вісн. Сер. прикл. матем. — 2009. — Вип. 6 (15) . — С. 179–196.

4. Юськів М.І., Цегелик Г.Г. Моделювання та ефективність доступу до послідовно організованих сторінок Web-сайту для різних законів розподілу ймовірностей звертання до сторінок – Інформатика та системні науки ІСН-2016: матеріали VII Всеукраїнської науково-практичної конференції за міжнародною участю, м. Полтава, 10–12 берез. 2016 р. — Полтава: ПУЕТ, 2016.
5. Юськів М.І., Цегелик Г.Г. Ефективність методу послідовного перегляду сторінок на Web-сайті для різних законів розподілу ймовірностей звертання до сторінок, м. Львів, 28-30 вересня 2016 р. — Львів: ФМІ, 2016.
6. Baeza-Yates R., Castilio C. Relating Web Structure and User Search Behavior. – *Advances in Web mining and Web usage Analysis. Materials of 8th International Workshop.* University of Philadelphia. — 2006. — Pp. 226–232.
7. Jansen B., Spink A. How are we searching the world wide web? A comparison of nine search engine transaction logs. *Information Processing and Management: an International Journal - Special issue: Formal methods for information retrieval.* — 2006. — Vol. 42, № 1. — Pp. 248–263.
8. Maharrey, B.K. Cloud Computing on Amazon's EC2 / Department of Computer Science. Auburn University. — 2010. — 6 p.
9. Ingwersen P., Jarvelin K. The Turn. Integration of information seeking and Retrieval in Context. — Dordrecht: Springer, 2005. — 417 p.
10. How the complexity of Google's search ranking algorithms changes over time. [Електронний ресурс]. Режим доступу: <https://medium.com/@nikhilbd/how-the-complexity-of-googles-search-ranking-algorithms-changes-over-time-6b0589a3c90f> (Дата звернення 10.09.2017).

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ ДОСТУПА К СТРАНИЦАМ WEB-САЙТА ДЛЯ РАЗЛИЧНЫХ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ ОБРАЩЕНИЯ К СТРАНИЦЕ

М.І. Юськів, Г.Г. Цегелик

Львовский национальный университет им. Ивана Франка,
ул. Университетская, 1, Львов, 79000, Украина; e-mail: markiyan_yuskiv@live.com

В статье рассмотрен подход к математическому моделированию оптимального доступа пользователей к последовательно организованных страниц Web-сайта. Найдено выражение для математического ожидания общего времени, необходимого для поиска страницы, для различных законов распределения вероятности обращения к страницам. Также выведено соотношение для нахождения параметров, при которых математическое ожидание достигает минимума. Предложены подходы для реализации данного алгоритма в прикладных приложениях. На основе построенной математической модели оптимального доступа к страницам информационного Web-сайта со стороны пользователя и полученного соотношения для нахождения параметров, при которых математическое ожидание достигает минимума можно сделать вывод, что данное исследование имеет потенциал для дальнейшего применения его результатов на производстве. В частности для построения высокопроизводительных модулей для реализации поиска в приложениях с облачной инфраструктурой.

Ключевые слова: моделирование доступа к страницам, Web-сайт, законы распределения вероятностей, микросервисная архитектура, оптимизация доступа

MODELING AND OPTIMIZATION OF ACCESS TO THE PAGES OF WEB-SITE FOR DIFFERENT DISTRIBUTION LAWS OF PROBABILITY OF ACCESSING TO THE PAGE

M.I. Yuskiv, G.G. Tsegelik

Lviv National University named after Ivan Franko,
1, Universitetskaya Str., Lviv, 79000, Ukraine; e-mail: markiyan_yuskiv@live.com

The article deals with mathematical modeling approach to optimal user access to pages organized sequentially Web-site. An expression is found for the expectation of the total time required to search the page for the various laws of probability distribution pages. Also, displayed for the value of the parameters under which expectation reaches a minimum. The approaches to implement the algorithm in applied applications. Based on the mathematical model of optimal access to information pages from the Web-site user side and got value for finding the parameters under which expectation reaches a minimum, we can conclude that the data the research has the potential for further application of the results at work. In particular for building high-performance modules to implement a search applications with using cloud.

Keywords: modelling of access, adherently organized pages, web-site, distribution laws probability, microservice arhitecture, optimization of access

АНАЛИЗ МЕДИАННЫХ МЕТОДОВ КОНSENSУСНОГО АГРЕГИРОВАНИЯ РАНГОВЫХ ПРЕДПОЧТЕНИЙ

В.А. Болтенков, В.И. Куваева, А.В. Позняк

Одесский национальный политехнический университет,
просп. Шевченко, 1, Одесса, 65044, Украина; e-mail: vaboltenkov@gmail.com

В работе исследованы медианные методы консенсусного агрегирования многоагентных индивидуальных предпочтений, выполненных в ранговых шкалах. Рассмотрены три медианных ранжирования: медиана Кука-Сейфорда в пространстве позиций альтернатив, медиана Литвака в пространстве векторов предпочтений, медиана Кемени в пространстве парных сравнений. Поскольку все медианные ранжирования являются NP-полными задачами, проведена оценка вычислительной сложности различных вариантов приближенных алгоритмов медианного агрегирования. Введено понятие задач агрегирования большой размерности с числом альтернатив более пяти. Установлено, что для задач рангового агрегирования большой размерности точное решение при минимальных вычислительных затратах обеспечивает эвристический алгоритм построения медианы Кемени, предложенный Б.Г. Литваком. Эвристический алгоритм Литвака реализован в электронных таблицах Microsoft Excel для практической задачи построения системы поддержки принятия решений при выборе абитуриентом ИТ-специальности для обучения в университете. Полученные результаты позволяют рекомендовать эвристический алгоритм Литвака для вычисления медианного консенсуса по Кемени для практического применения в задачах агрегирования рангов большой размерности.

Ключевые слова: агрегирование ранговых предпочтений, ранговые шкалы, консенсусное ранжирование, медианные алгоритмы, системы поддержки принятия решений

Введение

Задача агрегирования ранговых предпочтений берет свое начало в теории социального выбора, первые формулировки которой относятся к концу XVIII века [1]. В последние 30 лет она привлекает внимание исследователей как полезный инструмент для коллективного экспертного оценивания, метапоиска, выбора подобных объектов и других прикладных информационных технологий. Задача агрегирования рангов заключается в объединении индивидуальных предпочтений, построенных в ранговой шкале для некоторого множества альтернатив несколькими независимыми агентами (экспертами или другими информационными объектами, например, поисковыми машинами), в единое коллективное ранжирование в той же шкале, которое принято называть «консенсусом» или «консенсусным ранжированием» [1-2]. Имея консенсусное ранжирование, далее можно отсортировать индивидуальные ранжирования по мере их близости к консенсусному. Особое место в консенсусных ранжированиях занимают ранжирования, основанные на медианных принципах. Эти ранжирования являются аксиоматически обоснованными и математически строгими. Однако в литературе отсутствует систематическое, формализованное и содержательное изложение медианных принципов ранжирования и оценка границ их применимости в задачах с большим числом альтернатив.

Целью работы является систематизация и анализ методов консенсусного агрегирования ранговых предпочтений, основанных на медианных принципах, оценка их вычислительной сложности и применимости в задачах с большой размерностью множества альтернатив.

Основная часть

Все современные методы поиска рангового консенсуса берут свое начало в теории «справедливого» социального выбора, относящегося к работам Борда и Кондорсе [3]. Все методы агрегирования рангов можно разделить на две категории. Первая категория – так называемые позиционные методы, основанные на работе Борда [4]. Консенсусное ранжирование в них определяется с учетом позиций, которые занимают альтернативы в индивидуальных ранжированиях агентов. Простейшим из них является «счет Борда», согласно которому альтернативы в групповом ранжировании располагаются в соответствии с суммой их мест, набранных в индивидуальных ранжированиях. Вторая категория – методы, основанные на парных сравнениях. Эти методы основаны на интуитивной концепции Кондорсе [5], согласно которой альтернатива выигрывает (победитель по Кондорсе), если большинство экспертов считают ее предпочтительной в парных сравнениях со всеми остальными. Обе категории методов консенсусного агрегирования имеют свои преимущества и недостатки и достаточно подробно изучены в работе [6].

Формализуем задачу рангового агрегирования согласно [3]. Пусть существует множество альтернатив $A = \{A_1, A_2, \dots, A_n\}$, подлежащих упорядочению группой из K агентов по какому-либо критерию (или ряду критериев). Каждый из агентов k , $k = \overline{1, K}$ упорядочивает альтернативы и представляет индивидуальное ранжирование

$$P^k = \{A_{k_1}, A_{k_2}, \dots, A_{k_n}\}. \quad (1)$$

Предполагается, что каждый из агентов может установить, как строгий порядок на множестве альтернатив, так и слабый порядок, т.е. ввести одинаковые ранги для «неразличимых» альтернатив. Например, для множества $A = \{a_1, a_2, a_3, a_4, a_5\}$ может быть введено ранжирование $\langle a_1 \succ a_2 \succ a_3 \sim a_4 \succ a_5 \rangle$, где знак $\succ$ означает «предпочтительнее, чем ...», знак $\sim$ – «равнозначно с ...».

Альтернативы с одинаковыми рангами в русскоязычной литературе называют альтернативами со «связанными рангами», в англоязычной – ties (т.е. ничьими). Каждое индивидуальное ранжирование (1) может быть представлено в виде $P^k = \{q_1^k, q_2^k, \dots, q_n^k\}$, где q_i^k – позиция, занимаемая i -й альтернативой в ранжировании k -го агента.

Групповое консенсусное ранжирование, основанное на концепции расстояния, заключается в определении ранжирования P^k , ближайшего по некоторой введенной мере ко всем индивидуальным ранжированиям, т.е.

$$\arg \min_P \sum_{k=1}^K d(P^k, P) \rightarrow \hat{P}. \quad (2)$$

Решение задачи (2) называется медианным консенсусным ранжированием или в соответствующем контексте просто медианой. Задача (2) может быть решена как дискретная целочисленная задача оптимизации полным перебором, однако показано [3], что объем перебираемых вариантов стремительно растет с ростом мощности

множества альтернатив. Так, для множества из 5 альтернатив со связанными рангами требуется перебор 541 варианта, а для множества из 10 альтернатив – уже 102247563 вариантов.

Рассмотрим детально существующие медианные консенсусные ранжирования.

Медиана Кука-Сейфорда. Расстояние между двумя ранжированиями по Куку-Сейфорду определяется в пространстве позиций, занимаемых альтернативами:

$$d(P^{k_1}, P^{k_2}) = \sum_{i=1}^n |q_i^{k_1} - q_i^{k_2}|, \quad (3)$$

где q_i^k – позиция, занимаемая i -й альтернативой в ранжировании k -го агента.

В [7] показано, что расстояние (3) удовлетворяет системе аксиом меры близости, введенных в этой работе. Расстояние ранжирующего вектора P от всего множества ранжирований агентов $\{P^k\}$ равно $d(P, P^{(k)}) = \sum_{k=1}^K d(P, P^k) = \sum_{k=1}^K \sum_{i=1}^n |q_i^k - q_i|$, где q_i – позиция альтернативы A_i в ранжировании P .

Задача оптимизации по Куку-Сейфорду формулируется в виде: найти такое ранжирование $\hat{P}$, для которого

$$d(\hat{P}, P^{(k)}) = \arg \min_P \sum_{k=1}^K d(P^k, P) = \arg \min_P \sum_{k=1}^K \sum_{i=1}^n |q_i^k - q_i|. \quad (4)$$

Если предположить, что альтернатива A_i может занимать j -ю позицию в ранжировании P ($j = 1, 2, \dots, n$), расстояние (4) можно записать в виде

$$d(\hat{P}, P^{(k)}) = \sum_{i=1}^n \sum_{j=1}^n d_{ij} y_{ij}, \quad (5)$$

где $d_{ij} = \sum_{k=1}^K |d_i^k - j|$, $y_{ij} = \begin{cases} 1, & \text{если } A_i \text{ занимает позицию } j \text{ в ранжировании } P, \\ 0, & \text{в противном случае.} \end{cases}$

Задача минимизации (5) может быть представлена как хорошо известная в дискретном программировании задача о назначениях [8]:

$$\arg \min_{y_{ij}} \sum_{i=1}^n \sum_{j=1}^n d_{ij} y_{ij},$$

с ограничениями

$$\forall_{i=1, \dots, n} \sum_{j=1}^n y_{ij} = 1, \quad (6)$$

(одна альтернатива может занимать только одну позицию),

$$\forall_{j=1, \dots, n} \sum_{i=1}^n y_{ij} = 1, \quad (7)$$

(в одной позиции может находиться только одна альтернатива).

Медиана Литвака. Б.Г. Литвак предложил еще один способ построения консенсусного агрегированного ранжирования [9], который в англоязычной литературе получил наименование медианы Литвака [10]. Подход Литвака основан на так называемом векторе предпочтений. Для ранжирования k -го агента вводится вектор предпочтений $\pi^k = [\pi_1^k, \pi_2^k, \dots, \pi_n^k]$, где π_i^k – число альтернатив, предшествующих i -й альтернативе в ранжировании. По определению Литвака для двух векторов предпочтений π^{k_1} и π^{k_2} , сформированных ранжирований P^{k_1} и P^{k_2} соответственно, расстояние между ними определяется как $d(P^{k_1}, P^{k_2}) = \sum_{i=1}^n |\pi_i^{k_1} - \pi_i^{k_2}|$. В [11] доказано, что введенное расстояние удовлетворяет введенным аксиомам расстояния близости.

По определению Литвака расстояние от ранжирования P на полном множестве индивидуальных ранжирований агентов $\{P^k\}$ определяется следующим образом:

$$d(P, P^{(k)}) = \sum_{k=1}^K \sum_{i=1}^n |\pi_i^P - \pi_i^k|.$$

Введем $h_i^{k(j)} = |\pi_i^{P(j)} - \pi_i^k|$, где $i = \overline{1, n}$; $j = \overline{1, n}$; $k = \overline{1, K}$. При таких обозначениях медиана Литвака есть результат задачи оптимизации:

$$\arg \min_{y_{ij}} \sum_{i=1}^n \sum_{j=1}^n h_i^{(j)} y_{ij},$$

с ограничениями (6)-(7).

Медиана Кемени. Расстояние между ранжированиями для медианы Кемени [11] определено с применением матрицы парных сравнений. Для заданного ранжирования n альтернатив k -м агентом ($k = \overline{1, K}$) $P^k : A_{k_1}, A_{k_2}, \dots, A_{k_n}$ может быть следующим образом построена матрица парных сравнений:

$$C^k = \begin{bmatrix} c_{11}^k, \dots, c_{1n}^k \\ \dots \dots \dots \\ c_{n1}^k, \dots, c_{nn}^k \end{bmatrix}, \quad (8)$$

$$c_{il}^k = \begin{cases} 1, & \text{для } A_i \succ A_l \\ 0, & \text{для } A_i \sim A_l \\ -1, & \text{для } A_i \prec A_l \end{cases},$$

где $c_{il} = 0$, $i = \overline{1, n}$; $l = \overline{1, n}$.

Для двух заданных ранжирований P^{k_1} и P^{k_2} расстояние между ними равно

$$d(P^{k_1}, P^{k_2}) = \frac{1}{2} \sum_{i=1}^n \sum_{l=1}^n |c_{il}^{k_1} - c_{il}^{k_2}|.$$

Расстояние между заданным ранжированием P и набором ранжирований, сформированных агентами, определяется как

$$d = d(P, P^{(k)}) = \frac{1}{2} \sum_{i=1}^n \sum_{l=1}^n \sum_{k=1}^K |c_{il}^P - c_{il}^k|.$$

Предположим, что в ранжировании P $A_i \succ A_l$, т.е. $c_{il}^P = 1$. Введем коэффициенты r_{il} следующим образом

$$r_{il} = \sum_{k=1}^K d_{ij}(P, P^{(k)}) = \sum_{k=1}^K |c_{il}^k - c_{il}^P| = \sum_{k=1}^K |c_{il}^k - 1|.$$

Они называются коэффициентами потерь, а матрица $R = [r_{il}]$ – матрицей потерь. Предполагается, что $r_{il} = 0$ для всех $i = l$.

Можно показать [11], что

$$d = \sum_{i=1}^n \sum_{j=1}^n r_{ij} + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^K |c_{ij}^k|,$$

$$(i,j) \in I_p^{(1)} \quad (i,j) \in I_p^{(2)}$$

где $I_p^{(1)}$ – множество индексов (i, j) , для которых $A_i \succ A_j$ в ранжировании P (или множество индексов, для которых $c_{ij}^P = 1$), $I_p^{(2)}$ – множество индексов (i, j) , для которых $A_i \sim A_j$ в ранжировании P (или множество индексов, для которых $c_{ij}^P = 0$). Ранжирование P^M , такое, что $P^M(P_1, P_2, \dots, P^K) = \arg \min_P d(P, P^K)$, называется медианой Кемени.

Таким образом, известны три консенсусных медианы, которые представляют собой решение задачи поиска ранжирования, минимально удаленного от индивидуальных ранжирований агентов в соответствующих пространствах:

- медиана Кука-Сейфорда в пространстве позиций альтернатив;
- медиана Литвака в пространстве векторов предпочтений;
- медиана Кемени в пространстве парных сравнений.

Остановимся на практически важном вопросе вычислительной сложности построения консенсусного рангового агрегирования на основе медианных методов. Доказано, что вычисление всех медианных агрегированных ранжирований есть NP-полные задачи при $n > 5$ [12,13]. Далее будем называть задачи поиска консенсусного медианного ранжирования с $n > 5$ задачами большой размерности. Для достижения возможности решения задачи консенсусного агрегирования за реальное время в таких задачах был предложен ряд алгоритмов, главным образом основанных на решении задачи целочисленного программирования, к которым приводятся задачи рангового агрегирования. В частности, в [14] доказана теорема о том, что задача отыскания медианы Кемени на множестве индивидуальных предпочтений эквивалентна задаче о назначениях. Показано, что решение задачи консенсусного медианного ранжирования может быть получено венгерским методом или как частный случай транспортной задачи с бинарными переменными с вычислительной сложностью $O(n^3)$. Рассмотрен ряд алгоритмов, позволяющих решить задачу медианного консенсусного ранжирования методом ветвей и границ путем приведения вычислительной сложности к полиномиальной [15]. В [16] отмечено, что всего известно 104 алгоритма вычисления, однако указано, что все они нарушают исходную медианную аксиоматику введением тех или иных предположений и тем самым уже являются приближенными или эвристическими алгоритмами поиска консенсусного медианного ранжирования. В процессе исследования проведен анализ более 40 методов, из них наиболее простым представляется приближенный метод вычисления медианы Кемени, предложенный Литваком [9] и сочетающий в себе точность, отсутствие жестких требований к производительности вычислительных ресурсов и главное – алгоритмическую логичность и простоту.

Эвристический метод расчета медианного консенсуса Кемени по Литваку в виде пошагового алгоритма выглядит следующим образом:

Шаг 1. Построить матрицу потерь R , где $r_{ij} = d_{ij}(P_1, P_x) + d_{ij}(P_2, P_x) + \dots + d_{ij}(P_n, P_x)$; P_x – ранжирование, у которого $a_{ij}^x = a_{ij \max}^x$, т.е. $a_{ij}^x = 1$. Если в матрицах всех ранжирований стоит 1 или -1, то соответствующий элемент матрицы потерь будет равен 0.

Шаг 2. Определить матрицу $Q^n = E^n R^n E^n$, где

$$E^n = \begin{pmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 0 \end{pmatrix}.$$

Шаг 3. Определить наименьший элемент, исключением элементов из главной диагонали матрицы $Q^n - n = q_{ij \min}^n$.

Определить значения: $v^n = 2K \frac{(n-2)(n-3)}{2}$; $d_{ji \min}^{(n)} = d_{ji \min}^{(n)} - v^n$.

Альтернативы, определяющие строку и столбец элемента $q_{ij \min}^n$, обозначаются как A_{i_n} и A_{j_1} соответственно.

Шаг 4. Строки и столбцы, относящиеся к альтернативам A_{i_n} и A_{j_1} , удалить из матрицы R^n . Результат – новая матрица R^{n-2} .

Шаг 5. Определить матрицу Q^{n-2} : $Q^{n-2} = E^{n-2} R^{n-2} E^{n-2}$.

Шаг 6. Определить наименьший элемент исключением элементов из главной диагонали матрицы $Q^{n-2} - n = q_{ij \min}^{n-1}$. Определить значения: $v^n = 2K \frac{(n-4)(n-5)}{2}$;

$d_{ji \min}^{(n-2)} = d_{ji \min}^{(n-2)} - v^{(n-2)}$.

Шаг 7. Строки и столбцы, относящиеся к альтернативам A_{i_n} и A_{j_1} , удалить из матрицы R^{n-2} .

Шаги 1 - 4 повторяются, до тех пор, пока не будет получено конечное ранжирование по условию $n-2=0$ или $n-2=1$.

Пример практической реализации поиска консенсусного ранжирования в задаче большой размерности. Для разработки серверной части системы поддержки принятия решений (СППР) абитуриентами, поступающими на ИТ-специальности института компьютерных систем Одесского национального политехнического университета, была проведена коллективная экспертная оценка качеств, знаний и умений (далее – необходимых качеств) необходимых абитуриенту для успешного обучения на соответствующей специальности. Клиентская часть выполнена в виде Android-приложения и содержит вопросы для абитуриента, ответы на которые позволяют оценить у него наличие тех или иных качеств, знаний и навыков. В результате функционирования СППР подсказывает абитуриенту рекомендуемую для него ИТ-специальность [17].

Для оценки необходимых качеств для пяти специальностей: 113 – «Прикладная математика», 121 – «Инженерия программного обеспечения», 122 – «Компьютерные науки», 123 – «Компьютерная инженерия», 126 – «Информационные системы и технологии», 151 – «Автоматизация и компьютерно-интегрированные технологии», были сформированы шесть коллективов агентов – экспертов численностью соответственно 8, 15, 16, 12, 16, 10 человек из ведущих преподавателей выпускающих

кафедр. Каждому из экспертов был предложен список альтернатив, состоящий из 30 требуемых качеств (TKi), приведенный в таблице 1.

Таблица 1.
Список альтернатив (знаний, умений, личных качеств абитуриента)

№	Качества абитуриента	Обозначение
1	Хорошая школьная подготовка по математике	TK1
2	Хорошая школьная подготовка по физике	TK2
3	Хорошая школьная подготовка по информатике	TK3
4	Успешное участие в олимпиадах по математике	TK4
5	Успешное участие в олимпиадах по физике	TK5
6	Успешное участие в олимпиадах по информатике	TK6
7	Успешное участие в работе МАН	TK7
8	Опыт разработки программного продукта	TK8
9	Практическое владение английским языком	TK9
10	Практическое владение другим иностранным языком	TK10
11	Практическое владение одним из языков программирования	TK11
12	Опыт работы с приложениями под ОС Windows	TK12
13	Опыт работы с приложениями под ОС Android	TK13
14	Опыт работы с приложениями под ОС iOS	TK14
15	Практические знания аппаратной части компьютера	TK15
16	Практические умения по установке ОС на компьютер	TK16
17	Практические умения по настройке домашнего Wi-Fi роутера	TK17
18	Умение работать в команде	TK18
19	Умение быстро найти необходимую информацию в интернете	TK19
20	Умение руководить небольшой группой	TK20
21	Умение решать нестандартные задачи	TK21
22	Умение самостоятельно разобраться в незнакомой теме	TK22
23	Умение убедительно объяснить свою точку зрения	TK23
24	Стрессоустойчивость	TK24
25	Усидчивость	TK25
26	Дисциплинированность	TK26
27	Широкий кругозор	TK27
28	Стремление овладеть новыми знаниями и навыками	TK28
29	Ораторское искусство	TK29
30	Умение слушать	TK30

Для формирования консенсусного агрегирования результатов коллективной экспертизы по специальности 151 – «Автоматизация и компьютерно-интегрированные технологии», в которой участвовали 10 экспертов, применим изложенный выше эвристический алгоритм Литвака для приближенного расчета медианы Кемени.

Проведем предварительно ряд процедур предварительной обработки ранжирований [18]:

1. Стандартизация рангов: индивидуальные предпочтения, содержащие связанные ранги, обрабатываются с помощью процедуры стандартизации рангов для приведения ранговых оценок экспертов к сопоставимому виду и приводятся к измерению в одной и той же шкале для всех экспертов;

2. Оценка коэффициента конкордации W : перед формированием групповой оценки необходимо выяснить, насколько непротиворечивы индивидуальные экспертные ранжирования и можно ли их агрегировать в расчете на состоятельную консенсусную оценку:

$$W = \frac{S}{m^2(n^3 - n)/12 - m \sum_{j=1}^m T_j},$$

где S – сумма квадратов разностей между членами суммарного ранжирования и членами ряда, составленного из средних значений; m – число экспертов; n – число ранжируемых объектов;

$$T_j = \frac{1}{12} \sum_{i=1}^n (t_{ij}^3 - t_{ij}),$$

где t_{ij} – число повторений ранга t в j -том ряду.

Выполнена оценка статистической значимости коэффициента конкордации по критерию χ^2 . Показано, что оценке $W = 0,64$ при уровне значимости 0,95 соответствует $\chi^2 = 220,12$, поскольку $\chi^2 > \chi_{табл}^2$, $\chi_{табл}^2$ – табличное значение распределения χ^2 для $(n-2)$ степеней свободы, гипотеза о согласованности мнений всей группы экспертов принимается.

3. Выделение значимых альтернатив: для выявления значимости каждой альтернативы определяются коэффициенты значимости (K_j) по формуле

$$K_j = \frac{mn - S_j}{0,5mn(n-1)}, \text{ где } S_j - \text{сумма рангов для каждой альтернативы.}$$

Для выделения из n альтернатив наиболее значимых, введем порог значимости коэффициентов $K_i \geq 1/n$ и выделим наиболее значимые альтернативы. В результате вычислений выявлено 10 наиболее значимых альтернатив: ТК1, ТК2, ТК3, ТК28, ТК22, ТК25, ТК23, ТК26, ТК27, ТК11, приведенных в таблице 1. Стандартизованные ранги для них представлены в таблице 2.

Таблица 2.

Стандартизованные ранговые оценки для значимых альтернатив

	ТК1	ТК2	ТК3	ТК28	ТК22	ТК25	ТК23	ТК26	ТК27	ТК11
Эксперты	1	2	3	4	5	6	7	8	9	10
Э1	33,0	33,0	33,0	21,0	33,0	21,0	28,5	21,0	28,5	21,0
Э2	34,0	34,0	30,0	24,5	13,0	30,0	24,5	34,0	24,5	24,5
Э3	34,5	34,5	32,0	26,5	15,5	26,5	26,5	26,5	21,5	26,5
Э4	34,5	30,0	23,5	32,5	34,5	23,5	28,0	23,5	23,5	14,5
Э5	33,0	33,0	33,0	33,0	28,0	28,0	19,0	28,0	33,0	19,0
Э6	33,0	33,0	33,0	33,0	26,5	26,5	26,5	26,5	33,0	15,5
Э7	33,0	33,0	33,0	33,0	27,0	33,0	30,0	27,0	16,5	27,0
Э8	33,5	33,5	26,5	33,5	26,5	33,5	26,5	26,5	26,5	26,5
Э9	29,5	29,5	29,5	29,5	29,5	21,0	29,5	21,0	16,0	29,5
Э10	33,5	21,5	33,5	21,5	33,5	21,5	21,5	21,5	21,5	30,5

Каждую индивидуальную систему предпочтений надо преобразовать в матрицу парных сравнений, элементы которой вычисляются по соотношению (8).

Далее из матриц парных сравнений рассчитываем элементы матрицы потерь R (табл. 3), как указано на шаге 1 эвристического метода Литвака.

Таблица 3.

Матрица потерь R для построения медианы Кемени

		TK1	TK2	TK3	TK28	TK22	TK25	TK23	TK26	TK27	TK11
		1	2	3	4	5	6	7	8	9	10
TK1	1	0	8	6	5	4	2	1	1	2	1
TK2	2	12	0	7	8	6	3	2	2	3	3
TK3	3	14	13	0	8	6	5	4	4	4	2
TK28	4	15	12	12	0	7	7	6	5	6	6
TK22	5	16	14	14	13	0	10	9	8	9	7
TK25	6	18	17	15	13	10	0	9	9	8	6
TK23	7	19	18	16	14	11	11	0	8	8	7
TK26	8	19	18	16	15	12	11	12	0	9	8
TK27	9	18	17	16	14	11	12	12	11	0	8
TK11	10	19	17	18	14	13	14	13	12	10	0

Матрица потерь обрабатывалась по алгоритму Литвака в системе Microsoft Excel. В результате получено итоговое консенсусное ранжирование альтернатив:

$TK1 \succ TK2 \succ TK3 \succ TK28 \succ TK25 \succ TK22 \succ TK23 \succ TK26 \succ TK27 \succ TK11$.

Правильность расчетов была проверена с использованием комплекса Matlab-программ, основанных на решении задачи поиска медианы Кемени как задачи целочисленного программирования с использованием метода ветвей и границ [19]. Результаты расчета полностью совпадают, при этом при расчете с помощью Matlab-программ затраты процессорного времени примерно в 1,8-2 раза больше, чем при расчетах в Microsoft Excel. Для сравнения тот же набор индивидуальных ранжирований был обработан по методу Борда [6]. Затраты процессорного времени при этом сопоставимы с расчетами в Microsoft Excel. Полученное агрегированное консенсусное ранжирование выглядит так:

$TK1 \succ TK2 \succ TK3 \succ TK28 \succ TK22 \succ TK25 \succ TK23 \succ TK26 \succ TK27 \succ TK11$.

Полученный результат показывает, что медиана Кемени, рассчитанная по эвристическому приближенному алгоритму Литвака, позволяет получить более точное консенсусное агрегирование – видно, что счет Борда не позволяет различить превосходство альтернативы 6 над альтернативой 5.

Выводы

В работе исследованы медианные методы построения консенсусных предпочтений в ранговых шкалах. Систематизированы данные по медианам Кука-Сейфорда, Литвака и Кемени. Поскольку все медианные методы являются NP-полными задачами, проведен анализ приближенных методов оценивания консенсусных ранжирований в задачах большой размерности. Установлено, что приближенный алгоритм расчета медианы Кемени, предложенный Литваком, позволяет получить достаточно точное консенсусное ранжирование с минимальными вычислительными

затратами. Это позволяет рекомендовать алгоритм Литвака для расчета рангового консенсуса в многоагентных задачах большой размерности.

Список литературы

1. Elkind, E. Distance rationalization of voting rules / E. Elkind, P. Faliszewski, A. Slinko. // *Social Choice and Welfare*. — 2015. — Vol. 45, No. 2. — Pp. 345–377.
2. D'Ambrosio, A. Kemeny's axiomatic approach to end consensus ranking in tourist satisfaction / A. D'Ambrosio, V.A. Tutore. // *Statistica Applicata*. — 2008. — Vol. 20, No. 1. — Pp. 21–32.
3. Bury, H. Group Judgement With Ties. Distance-Based Methods / H. Bury, D. Vagner. // In: H. Aschemann (ed.). *New Approaches in Automation and Robotics*. — Vienna: I-Tech I-Tech Education and Publishing. — 2008. — Pp. 153–172.
4. Borda, J. C. M'emoire sur les 'elections au scrutin [Electronic resource] // *Histoire de l'Acad'emie Royale des Sciences*. — Paris, 1781. Mode of access: http://gerardgreco.free.fr/IMG/pdf/MA_c_moire-Borda-1781.pdf
5. Condorcet, J. A. N. Essai sur L'Application de L'Analyse a la Probabilite des Decisions Rendues a la Pluralite des Voix [Electronic resource] // A Paris: de l'Imprimerie royale, 1785. — Paris, 1781. Mode of access: <http://www.e-rara.ch/zut/content/titleinfo/1175327>
6. Петровский, А. Б. Теория принятия решений. / А. Б. Петровский. — М.: Академия, 2009. — 400 с.
7. Cook, W.D. Distance-based and adhoc consensus models in ordinal preference ranking. / W.D. Cook. // *European Journal of Operational Research*. — 2006. — No. 172. — Pp. 369–385.
8. Писарук, Н. Н. Исследование операций / Н. Н. Писарук. — Минск: БГУ, 2015. — 304 с.
9. Литвак, Б.Г. Экспертная информация: Методы получения и анализа / Б. Г. Литвак. — М.: Радио и связь, 1982. — 184 с.
10. Bury, H. Zastosowanie mediany Litvaka do wyznaczania oceny grupowej w przypadku występowania obiektów równoważnych / H. Bury, D. Wagner. // *Studia i Materiały Polskiego Stowarzyszenia Zarządzania Wiedzą*. — 2007. — Vol. 10. — Pp. 19–34.
11. List, C. Judgement aggregation: a survey / C. List, C. Puppe. // In: *Oxford handbook of rational and social choice*. — Oxford: Oxford University Press. — 2009. — Pp. 457–482.
12. Hudry, O. Complexity of computing median linear orders and variants / O. Hungry. // *Electronic Notes in Discrete Mathematics*. — 2013. — Vol. 42. — Pp. 57–64.
13. Hudry, O. NP-hardness results on the aggregation of linear orders into median orders / O. Hungry. // *Annals of Operations Research*. — 2008. — Vol. 163, No. 1. — Pp. 63–88.
14. Самохвалов, Ю.Я. Экспертное оценивание. Методический аспект / Ю. Я. Самохвалов, Е.М. Наumenko. — Киев: ДУИКТ, 2007. — 262 с.
15. Amodio, S. Accurate algorithms for identifying the median ranking when dealing with weak and partial rankings under the Kemeny axiomatic approach // S. Amodio, A. D'Ambrosio, R. Siciliano. // *European Journal of Operational Research*. — 2015. — No. 249. — Pp. 667–676.
16. Alnur, A. Experiments with Kemeny Ranking: What Works When? / A. Alnur, M. Meila. // *Mathematical Social Sciences*. — 2012. — Vol. 64, No. 1. — Pp. 28–40.
17. Куваева, В.И. Применение методов экспертного оценивания при построении систем поддержки принятия решений / В.И. Куваева, А.В. Позняк, В.А. Болтенков. // *Системы та засоби штучного інтелекту: тези доповідей Міжнародної наукової молодіжної школи*. — Київ: ІППІ «Наука і освіта». — 2017. — С. 104–108.
18. Писарева, О.М. Методы прогнозирования развития социально-экономических систем / О.М. Писарева. — М: Высшая школа, 2007. — 591 с.
19. Compute the median ranking according to the Kemeny axiomatic approach // [Electronic resource]: Mode of access: <https://www.mathworks.com/matlabcentral/fileexchange/52235-compute-the-median-ranking-according-to-the-kemenyaxiomaticapproach?focused=3889946&tab=function&requestedDomain=true>.

АНАЛІЗ МЕТОДІВ КОНСЕНСУСНОГО АГРЕГУВАННЯ РАНГОВИХ ПЕРЕВАГ

В.О. Болтънков, В.І. Куваєва, О.В. Позняк

Одеський національний політехнічний університет,
просп. Шевченка, 1, Одеса, 65044, Україна; e-mail: vaboltentkov@gmail.com

У роботі досліджені медіанні методи консенсусного агрегування багатоагентних індивідуальних переваг, виконаних в рангових шкалах. Розглянуто три медіанні ранжирування: медіана Кука-Сейфорда в просторі позицій альтернатив, медіана Литвака в просторі векторів переваг, медіана Кемені в просторі парних порівнянь. Оскільки всі медіанні ранжирування є NP-повними задачами, проведена оцінка обчислювальної складності різних варіантів наближених алгоритмів медіанного агрегування. Введено поняття задач агрегування великої розмірності з числом альтернатив більше п'яти. Встановлено, що для задач рангового агрегування великої розмірності точне рішення, при мінімальних обчислювальних витратах, забезпечує евристичний алгоритм побудови медіани Кемені, запропонований Б.Г. Литваком. Евристичний алгоритм Литвака реалізований в електронних таблицях Microsoft Excel для практичної задачі побудови системи підтримки прийняття рішень при виборі абітурієнтом ІТ-спеціальності для навчання в університеті. Отримані результати дозволяють рекомендувати евристичний алгоритм Литвака для обчислення медіанного консенсусу за Кемені для практичного застосування в задачах агрегування рангів великої розмірності.

Ключові слова: агрегування рангових переваг, рангові шкали, консенсусні ранжирування, медіанні алгоритми, системи підтримки прийняття рішень

ANALYSIS OF MEDIAN METHODS FOR CONSENSUS RANK PREFERENCES AGGREGATION

V.A. Boltentkov, V.I. Kuvayeva, A.V. Pozniak

Odesa National Polytechnic University,
1, Shevchenko Ave., Odesa, 65044, Ukraine; e-mail: vaboltentkov@gmail.com

The median methods of consensus aggregation of multi-agent individual preferences performed in rank scales has been investigated. Three median rankings were considered: the Cook-Seyford's median in the space of alternatives positions, the Litvak's median in the space of preference vectors, the of Kemeny's median in the space of paired comparisons. Since all median rankings are NP-complete problems, the computational complexity of various variants of approximate median aggregation algorithms has been studied. The notion of large-dimensional aggregation problems with the number of alternatives more than five has been introduced. It is established that for the problems of rank-based aggregation of large dimension, the exact solution with minimal computational costs is provided by the heuristic algorithm for constructing the Kemeny's median proposed by B.G. Litvak. The heuristic algorithm of Litvak is implemented in Microsoft Excel spreadsheets for the practical task of constructing a decision support system for applicants selecting an IT specialty for studying at a university. The obtained results make it possible to recommend the heuristic algorithm of Litvak for calculating the median consensus according to Kemeny for practical application in problems of large dimension ranks aggregation.

Keywords: aggregation of rank preferences, rank scales, consensus ranking, median algorithms, decision support systems

ОПТИМИЗАЦИЯ ОТБОРА И АНАЛИЗА ИНФОРМАЦИИ В РАЗНОСТРУКТУРНЫХ ХРАНИЛИЩАХ ДАННЫХ

Д.С. Шибает, В.В. Вычужанин, Н.О. Шибаета, Н.Д. Рудниченко

Одесский национальный морской университет,
ул. Мечникова, 34, Одесса, 65029, Украина; e-mail: denshibaev@outlook.com

В работе предложена методология обработки большого объема информации, целью которой является уменьшение времени на прогнозирование состояния сложных технических систем, уменьшение затрат на передачу данных через спутниковые линии связи и возможность использования модульного программного продукта в качестве практического решения проблемы передачи большого объема информации через высоконагруженные спутниковые линии связи.

Использование предложенной методологии позволит сократить затраты, возникающие при передаче данных через спутниковые линии связи, а также повысить эффективность обработки большого объема информации о состоянии оборудования сложных технических систем.

Ключевые слова: анализ данных, большие объемы данных, компьютерные системы, базы данных, модульное программное обеспечение, SCADA-системы, передача данных, протоколы передачи данных, VSAT

Введение

Современные алгоритмические решения, направленные на работу с динамически-расширяемыми массивами информации, представляют собой сложные информационно-технические системы, разделенные на автономные подсистемы, основной задачей которых является реплицирование информации в центральные узлы обработки и управления данными [1,2]. Такие решения позволяют модульно наращивать хранилища данных в реальном времени, с целью дистанционной проверки сохраняемой информации удаленными аналитиками.

Сохранность целостности данных, обеспечение качественной масштабируемости системы, а также быстрый доступ к информации является актуальной задачей, решение которой позволит улучшить систему обработки больших объемов данных из сложных технических систем (СТС) [3,4].

СТС оснащается автоматизированными системами, выполняющими анализ работы их оборудования в реальном времени. Сбор данных с таких систем выполняет сервер контроля и обработки информации. Использование централизованных комплексов, интегрированных в СТС, обеспечит:

- контроль технологических процессов при взаимодействии оборудования;
- анализ состояния подсистем СТС;
- оптимизацию обслуживания систем;
- уменьшение нагрузки на обслуживающий персонал.

Известные сложные компьютерные системы (СКС) разделяются на категории, зависящие от области их применения. Одним из направлений в диагностировании с использованием СКС является анализ работоспособности СТС на транспорте [5]. Такая СКС состоит из активного сервера или пары серверов, а также линии передачи данных. Датчики параметров подсистем СТС выполняют передачу данных в строго

обозначенные периоды времени, которые могут быть заданы индивидуально для каждого типа оборудования СТС.

Накопление информации, полученной в результате работы СКС, позволяет контролировать качество и точность работы СТС, а также выполнять прогнозирование технического состояния оборудования систем. При этом возникают проблема хранения и сортировки информации, их анализа и контроля, а также сложности в накоплении информации для использования ее в различных аналитических целях [6].

Одним из оптимальных решений по хранению и накоплению информации является организация сервера баз данных, способного принимать информацию от оборудования СТС. Такая система позволяет хранить большое количество информации на протяжении длительного времени и быть доступной техническому персоналу в любое время. Для этого используют дисковые массивы большого объема, защищенные от внешнего воздействия и позволяющие добиться целостности данных.

Современные СКС, используемые на водном транспорте, представляют собой сочетания модульных интегрированных комплексов и вспомогательных контролеров управления. Такие комплексные решения являются SCADA-системами. Однако в СКС возникают проблемы, связанные со сбором и обработкой информации. Это связано с большими массивами информации, формируемыми и передаваемыми по внутренней локальной сети транспортного средства и требующими детального анализа для своевременного выявления неисправностей, анализа потребления топливных и горюче-смазочных материалов СТС [7,8].

Работа SCADA-систем основывается на промежуточном считывании информации с датчиков параметров СТС, занесением показаний в базу данных (БД) с дальнейшим архивированием показаний для анализа и прогнозирования технического состояния систем.

В связи с этим возникает необходимость создания единого решения, способного гарантировать качественную запись данных с датчиков контроля параметров подсистем СТС в централизованную БД. Такое решение должно обладать повышенной отказоустойчивостью, гарантировать проверку целостности данных, а также обеспечивать простое масштабирование системы.

Одним из классических решений такой задачи является использование реляционных БД, обладающих набором возможностей, способных улучшить качество процесса записи информации. Это позволяет:

- упростить проектирование модели данных;
- упорядочить структуру таблиц;
- сохранить целостность данных;
- контролировать дублирование данных.

Такие БД отказоустойчивы и обеспечивают хранение больших объемов информации без дополнительного контроля. Использование их в качестве центрального хранилища СКС улучшит показатели сохранности информации, а также гарантирует целостность данных при передаче информации от мест контроля технического состояния оборудования до записи информации в реляционную БД. Результатом использования реляционной БД в составе SCADA-системы (рис. 1) станет возможность сохранения [9]. При этом возникает проблема с передачей данных в удаленные центры обработки информации, которые обладают достаточной вычислительной мощностью, необходимой для работы с большими объемами накопленной информации за время эксплуатации транспортного средства. Передача большого объема данных, хранимых в реляционных БД, требует сохранения структуры таблиц данных для удаленных аналитических центров, а также необходимость в высокоскоростной линии передачи данных.

Из анализа систем сбора и хранения информации о состоянии СТС следует, что проблемой современных СКС, работающих по архитектурному принципу SCADA,

является невозможность обеспечения быстрой передачи данных для их обработки. Причина заключается в монолитности используемых архитектур хранения информации и большой нагрузке на каналы передачи данных. Это делает невозможным анализировать данные о техническом состоянии оборудования СТС в режиме реального времени [10-12].

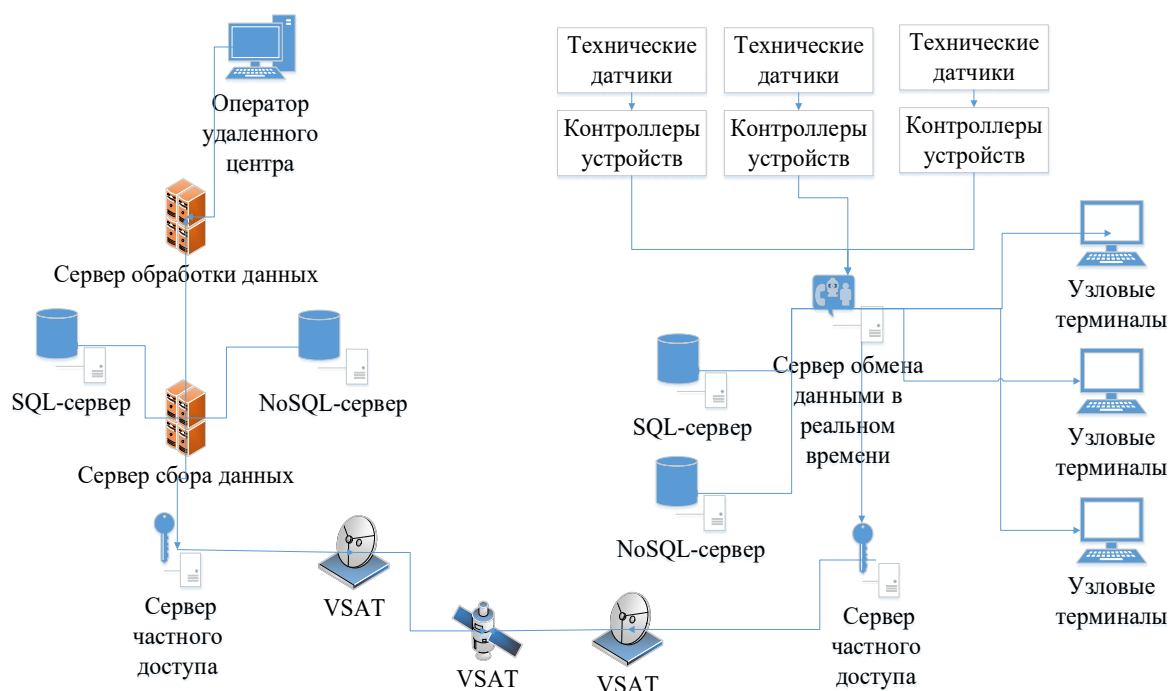


Рис. 1. Схема распределения хранения и анализа данных в SCADA СКС

Целью работы является разработка методологии передачи и обработки большого объема информации в реальном времени, включающей поиск, отбор и анализ данных, необходимых для прогнозирования технического состояния оборудования СТС.

Основная часть

Первоначальным этапом разработки решения по передаче и обработке большого объема информации, является формализация задач с использованием известного программного решения. Основные решаемые задачи:

- разработка алгоритма перехвата и промежуточного анализа потока передаваемой информации технического состояния СТС к серверу хранения результатов;
- подсчет исходных метаданных в пакете;
- разработка алгоритма распределения пакетов согласно считанным метаданным;
- разработка алгоритма архивирования данных согласно нормам передачи через спутниковые линии;
- построение модели передачи данных, основанной на минимизации размера передаваемого пакета;
- разработка алгоритма проведения анализа данных, необходимых для построения прогноза состояния СТС;
- прогнозирование состояния СТС, основанное на полученных данных из SCADA-систем.

В качестве программной реализации задач, сформированных при разработке передачи и обработки большого объема информации, необходимо создать модульное

программное решение, допускающее его использование как части функционального модуля SCADA-системы.

Начальным этапом реализации модульного программного решения является формулирование программных требований, выполняемых средствами разработки программных продуктов. Для этого используются инструменты проектирования, в состав которых входят CASE-средства. Было выполнено построение диаграммы вариантов использования, отражающей программное взаимодействие подсистем CTC в модульном программном решении (рис. 2).

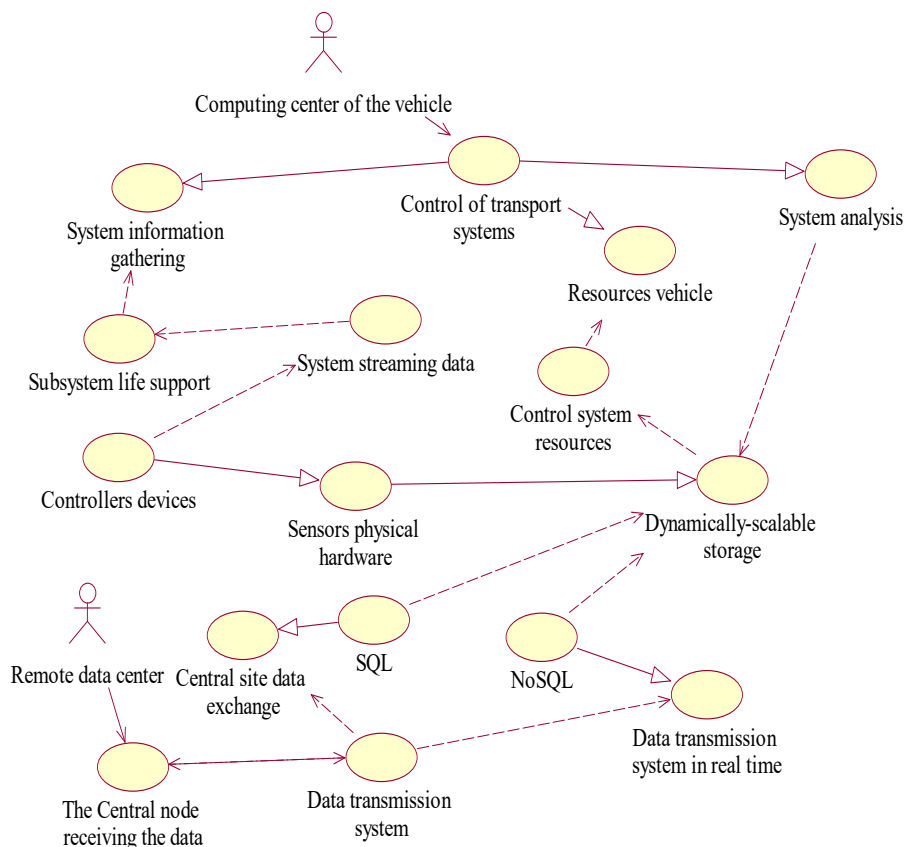


Рис. 2. Диаграмма взаимодействия подсистем CTC в модульном программном решении

Под модульностью программного решения подразумевается использование отдельных функциональных частей, осуществляющих работу с потоками информации, передаваемой по сетям, считывание и идентификацию метаданных пакетов информации, использование метода распределения информации согласно полученным метаданным и т.д. Такая комплексная реализация создаст объемное, монолитное решение, которое сложно интегрировать в SCADA-систему из-за разнообразия использования технологического оборудования CTC и средств контроля за их работоспособностью.

Для определения последовательности использования модулей программного решения был разработан обобщенный алгоритм обработки большого объема информации в реальном времени (рис. 3).

Для данных, которые являются приоритетно важными, необходимо предоставить отдельное хранилище, построенное на нереляционной архитектуре (NoSQL). Такое хранилище позволит добиться максимальной скорости записи информации, а также минимизировать сжатый архив информации для дальнейшей передачи по спутниковым каналам. В качестве используемого типа NoSQL хранилища предлагается использовать «bigtable-подобную» базу данных, использующую семейства колонок в качестве

хранилища информации. Такая архитектура БД похожа на классическую реляционную базу данных, однако самым большим отличием является представление колонок в виде разреженной матрицы, строки и столбцы которой используются в качестве ключей. Использование «bigtable-подобной» БД подходит для взаимодействия с большим объемом информации, что делает модульную систему достаточно сложным средством обработки больших массивов данных.

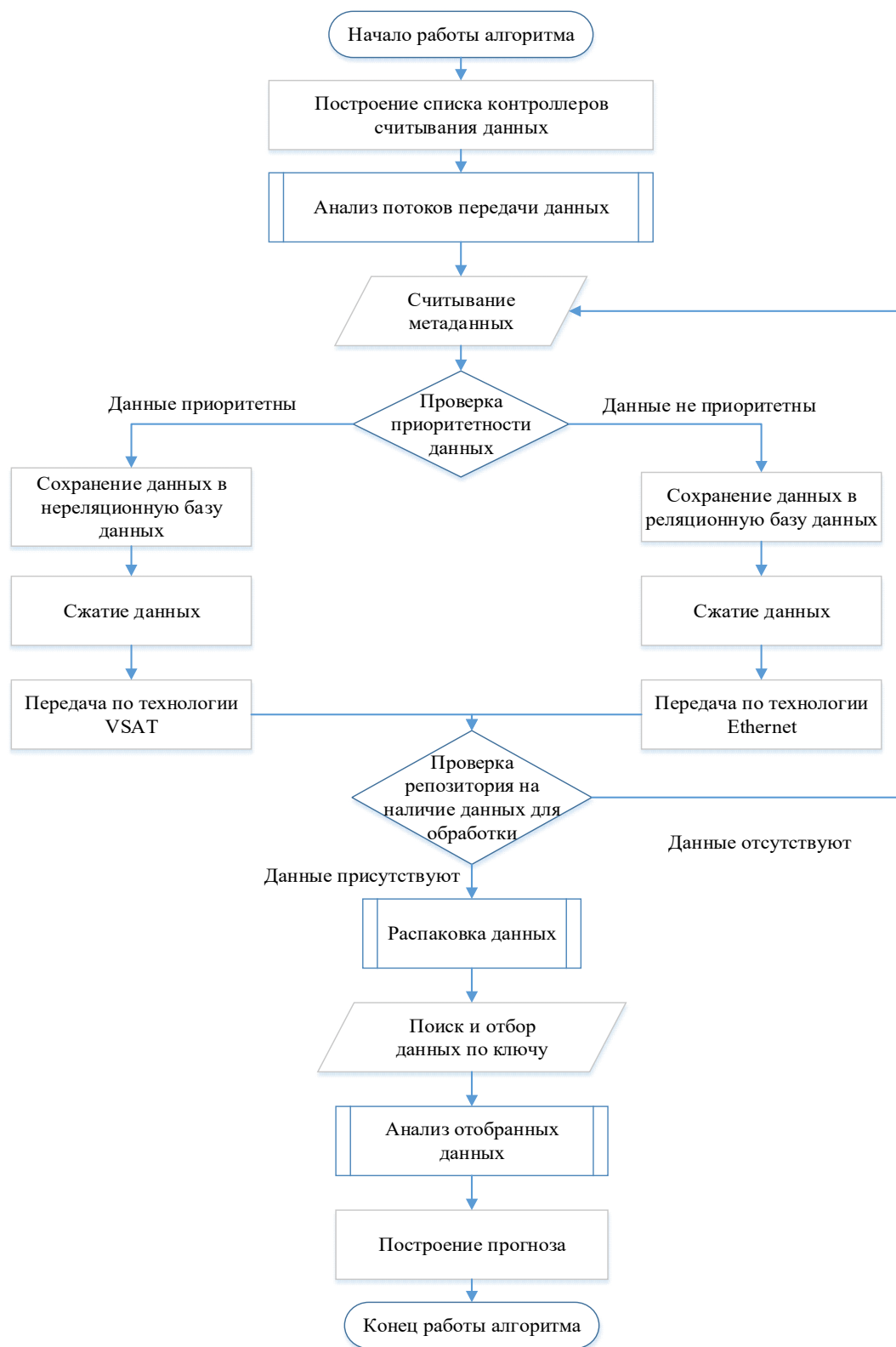


Рис. 3. Обобщенный алгоритм работы модульного программного обеспечения

В качестве способа распределения метаданных предлагается использовать прямую классификацию, позволяющую выполнять сортировку по типам датчиков сбора информации. Это позволит увеличить точность отбора данных, а также сфокусировать прогноз состояния СТС на конкретном оборудовании.

Для минимализации передаваемых данных был выбран алгоритм - деревья Хаффмана, который является одним из самых подходящих алгоритмов сжатия с контролем целостности данных. Он основывается на использовании кодов Хаффмана, при помощи которых можно сократить названия и системные маркировки датчиков, что уменьшит размер передаваемого массива.

В качестве спутниковой системы передачи данных рассматривается архитектура, основывающаяся на VSAT, способная передавать данные в реальном времени на скорости 4 мбит/сек. Такое решение анализирует данные, полученные от контроллеров датчиков физических устройств СТС, и использует такую информацию в качестве входных данных для построения различных прогнозов.

Выводы

Из анализа устройства СКС, входящих в состав SCADA архитектур, можно сделать вывод о необходимости разработки методологии обработки большого объема информации. Результатом является реализация модульного программного решения, способного улучшить итоговый процесс прогнозирования состояния СТС. Такое решение позволит внедрить NoSQL в средства оценки технического состояния систем, а также использовать улучшенные средства обработки больших массивов информации.

Список литературы

1. Кормен, Т. Алгоритмы: построение и анализ / Т. Кормен, Ч. Лейзерсон, Р. Ривест. — М: МЦНМО, 2002. — 960 с.
2. Кнут, Д.Э. Искусство программирования. Т.3. – Сортировка и поиск / Д. Кнут. — М.: Изд. Дом «Вильямс», 2010. — 788 с.
3. Вычужанин, В.В. Повышение эффективности эксплуатации судовой системы комфортного кондиционирования воздуха при переменных нагрузках / В.В. Вычужанин. — М: Одесса, ОНМУ, 2009. — 206 с.
4. Vychuzhanin, V. Assessment of risks structurally and functionally complex technical systems / V. Vychuzhanin, N. Rudnichenko. // Eastern-European Journal of Enterprise Technologies. — 2014. — Т. 1, № 2 (67). — С. 18–22.
5. Vychuzhanin, V. Devising a method for the estimation and prediction of technical condition of ship complex systems / V. Vychuzhanin, N. Rudnichenko, V. Boyko, N. Shibaeva, S. Kononov. // Eastern-European Journal of Enterprise Technologies. — 2016. — Т. 6, № 9 (84). — С. 4–11
6. Finkel, R A. Quad trees a data structure for retrieval on composite keys / R.A. Finkel, J.L. Bentley. // Acta Informatica. — 1974. — Pp. 1–9.
7. Андреев, Е.Б. SCADA-системы: взгляд изнутри / Е.Б. Андреев, Н.А. Куцевич, О.В. Синенко. — М.: РТСОфт, 2004. — 176 с.
8. Прошин, Д.И. Проблемы выбора инструментальных средств построения SCADA-систем / Д.И. Прошин, Л.В. Гурьянов // Информатизация и Системы Управления в Промышленности. — 2010. — № 1 (25). — С. 21–25.
9. Ефимов, И.П. SCADA-система TraceMode / И.П. Ефимов, Д.А. Солуянов. — Ульяновск: УлГТУ, 2010. — 158 с.
10. Кудрявцев, К.Я. Методы повышения скорости поиска информации в базах данных / К.Я. Кудрявцев, А.Е. Коротков. // LAP Lambert Academic Publishing, 2012. — 84 с.
11. Samet, H. The Quadtree and Related Hierarchical Data Structures / H. Samet. // ACM Comput. Surv. — 1984. — Pp. 187–260.

12. Arasu, A. Efficient exact set-similarity joins / A. Arasu, V. Ganti, R. Kaushik. // Proceedings of the 32nd international conference on Very large data bases. VLDB '06. VLDB Endowment. — 2006. — Рр. 918–929.

ОПТИМІЗАЦІЯ ВІДБОРУ ТА АНАЛІЗУ ІНФОРМАЦІЇ В РІЗНОСТРУКТУРНИХ СХОВИЩАХ ДАНИХ

Д.С. Шибасєв, В.В. Вичужанин, Н.О. Шибасєва, Н.Д. Рудніченко

Одеський національний морський університет,
вул. Мечнікова, 34, Одеса, 65029, Україна, denshibaev@outlook.com

В роботі запропоновано методологію обробки великого обсягу інформації, результатом якої є зменшення часу на прогнозування стану складних технічних систем, зменшення витрат на передачу даних через супутникові лінії зв'язку і можливість використання модульного програмного продукту в якості практичного вирішення проблеми передачі великого обсягу інформації через високонавантажені супутникові лінії зв'язку. Використання запропонованої методології дозволить скоротити витрати, що виникають при передачі даних через супутникові лінії зв'язку, а також підвищити ефективність обробки великого обсягу інформації, отриманої в результаті передачі масиву даних від контролерів станів обладнання складних технічних систем.

Ключові слова: аналіз даних, великі об'єми даних, комп'ютерні системи, бази даних, модульне програмне забезпечення SCADA-системи, передача даних, протоколи передачі даних, VSAT

OPTIMIZATION OF SELECTION AND ANALYSIS OF INFORMATION IN DATA WAREHOUSES RESTRUCTURING

D.S. Shibaiev, V.V. Vichuzhanin, N.O. Shibaieva, N.D. Rudnichenko

Odesa National Maritime University,
34, Mechnikova Str., Odesa, 65029, Ukraine, denshibaev@outlook.com

Proposed methodology processing of a large volume of information, the result of which is to reduce the time to forecast the state of complex technical systems, reducing costs for data transmission via satellite links and the ability to use modular software product as a practical solution to the problem of delivering large amounts of high-load information via satellite links. Using the proposed methodologies will reduce the costs incurred when transmitting data via satellite links, as well as to improve the efficiency of processing a large amount of information obtained as a result of transfer of array data from controller condition of the equipment complex technical systems.

Keywords: data mining, big data, computer systems, databases, modular software, SCADA, data transmission, data transfer protocols, VSAT

АЛГОРИТМ ФУНКЦІОНУВАННЯ СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ В ЗЕМЕЛЬНИХ ВІДНОСИНАХ

М.А. Кухар

Харківський національний університет міського господарства ім. О.М. Бекетова,
вул. Маршала Бажанова, 17, Харків, 61002, Україна; e-mail: office@kname.edu.ua

Розглядається актуальна задача використання онтологічного інжинірингу земельних відносин. Розроблені алгоритм функціонування системи підтримки прийняття рішень в земельних відносинах, математичні конструкції теорії категорій і логіки предикатів. Вони відображають земельні відносини в логічному, структурованому вигляді і дозволяють отримати чітке розуміння функціонування системи підтримки прийняття рішень при використанні цих моделей як елементів бази знань. У перспективі, удосконалюючи і деталізуючи запропонований алгоритм, можливо сформувати більш ефективну систему підтримки прийняття рішень в сфері земельних відносин. Формалізовані у математичному вигляді закони, якщо їх зібрати у базі знань, можна використовувати як базис для регулювання земельних відносин чітко, згідно закону; зменшити час вирішення конкретних питань у сфері земельних відносин; збільшити ефективність формування нових норм і правил земельного законодавства, з урахуванням прийнятих раніше законів; зменшити кількість незаконних дій посадових осіб в сфері земельного законодавства та ін. Результати роботи можуть бути впроваджені на підприємствах, які займаються моніторингом землі та, загалом, проводять свою діяльність у заявку з земельними відношеннями, а також можуть використовуватись в системах земельного адміністрування.

Ключові слова: алгоритм, формалізація, теорія категорій, логіка предикатів, земельні відносини

Вступ

Аналіз законодавчої бази з урахуванням рішень органів державної влади та наукових працівників, які займаються відповідними базами, показує, що вони мають ряд недоліків: окремі закони про землю слабо пов'язані між собою, окремі положення викликають протиріччя в трактуванні закону, відсутність сучасних ІТ-технологій, спрямованих на організацію земельних відносин, та ін. Інформаційні технології використовуються в плануванні та управлінні земельними ресурсами, а також є необхідним інструментом при розробці систем підтримки прийняття рішень у стабільному управлінні земельними ресурсами та розробці політики їх виконання [1].

Сьогодні в Україні актуальним і перспективним завданням є систематизація та приведення земельних відносин до європейських стандартів, в основі яких лежить автоматизація та інформатизація. У сфері земельних відносин автоматизація та інформатизація зменшують негативні процеси і явища, пов'язані з зловживанням правовими недоліками в земельному законодавстві [2].

Дослідження даної предметної області [3-6] показують, що створення високопродуктивних систем автоматизації та інформатизації має здійснюватися на базі інтелектуальних інформаційних технологій, основу яких складають системи підтримки прийняття рішень. Дані інтелектуальні системи повинні володіти потужним математичним забезпеченням у вигляді відповідних методів і моделей. Системи підтримки прийняття рішень використовуються сьогодні в різноманітних сферах: від виробництва на конкретному підприємстві до прогнозування масштабних природних

катастроф [7-9], та мають різні математичні базиси, дозволяють замінити людські ресурси на етапі обробки і структурування вихідних даних. Важливо підкреслити, що роль людини в системах підтримки прийняття рішень залишається пріоритетною.

Будь-яку систему підтримки прийняття рішень можливо розробити лише при наявності бази даних і бази знань. Бази даних представляють собою набір даних моніторингу, певних розрахунків та досліджень, а бази знань містять правила виводу, інформацію про людський досвід і знання в деякій предметній області, які дозволяють виконувати операції над даними бази даних.

При заповненні баз знань в системах підтримки прийняття рішень в сфері земельних відносин важливу роль відіграють моделі формалізації права, бо землеустрій базується на земельному законодавстві. Ці моделі запропоновані в роботах [10-12], які показали, що ця сфера недостатньо досліджена.

Особливістю правових формальних моделей є те, що вони повинні бути строгими і не допускати неоднозначного тлумачення. Застосування моделей формалізації в правознавстві сприяє повному і правильному сприйняттю інформації та забезпеченню строгих правових приписів. Використання цих моделей дозволить чітко визначати всі базові поняття і сформулювати аксіоми земельних відносин у вигляді відповідних математичних конструкцій, дозволить вирішувати конкретні задачі в земельних відносинах за допомогою програмних засобів, розробляти програмне забезпечення з використанням математичних моделей формалізації та ін.

Мета роботи

Метою роботи є отримання моделей функціонування системи підтримки прийняття рішень і розробка узагальненого алгоритму функціонування системи підтримки прийняття рішень в сфері земельних відносин за умови заповнених баз даних і баз знань відповідно до предметної області.

Основна частина

Для досягнення поставленої мети необхідно розуміти структуру і логіку прийняття рішень в сфері земельних відносин, основою яких є вихідні дані про землю, власників і відношення між ними. Вони регулюються земельним законодавством України.

Система підтримки прийняття рішень при виробленні рішень і рекомендацій для тої чи іншої задачі в сфері земельних відносин використовує інформацію з центрального сховища даних (ЦСД), яке є частиною ієрархічної структури системи підтримки прийняття рішень і включає до свого складу базу даних (БД) і базу знань (БЗ). В даному випадку база даних повинна містити атрибутивну інформацію про землю, власників і відносинах між ними. База знань повинна складатися з моделей формалізації права, які можуть бути побудовані або логічними, або наближеними методами, або їх комбінаціями.

При розробці системи підтримки прийняття рішень необхідно враховувати представлення знань у базі знань онтологічними моделями. Тому, для формального представлення даних в базі знань обрані методи теорії категорій. Дані методи дозволяють у вигляді об'єктів категорії використовувати такі математичні конструкції, як висловлювання, предикати і т.д. Ця особливість теорії категорій при необхідності дозволяє комбінувати математичні моделі, які описують предметну область, тобто земельні відносини.

Позначимо вміст бази знань у вигляді багаторівневої моделі – категорії малих категорій – Cat , які поєднані між собою функторами – $\bar{F}$. Особливістю даної категорії є те, що її об'єкти $\{Ob_1^{Cat}, Ob_2^{Cat}\} \subset Cat$ – це малі категорії.

Виділена категорія Cat складається з двох малих категорій, які інтерпретуються наступним чином:

– $Ob_1^{Cat} \equiv K_1$ – сукупність законів у сфері земельних відносин;

– $Ob_2^{Cat} \equiv K_2$ – сукупність схем вирішення задач в сфері землеустрою, які можуть бути обрані відповідно до заданих критеріїв (правил) (блок 2).

Представлені в базі знань об'єкти категорій K_1 і K_2 дозволяють здійснити послідовний відбір закону і моделі, а потім поставити їх у відповідність $G \subseteq Ob(K_1) \times Ob(K_2)$ (рис. 1).

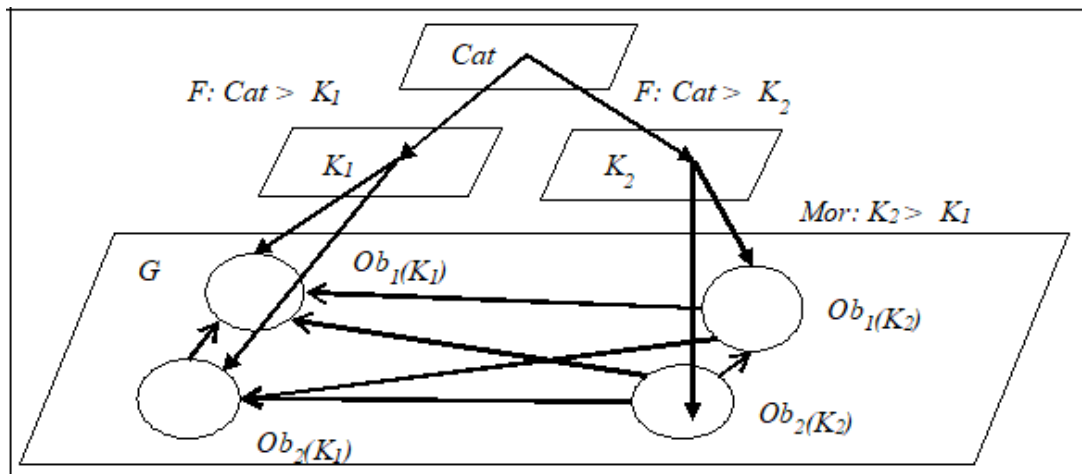


Рис. 1. Багаторівнева комутативна діаграма відношень між категоріями і об'єктами малих категорій при вирішенні задач в системі підтримки прийняття рішень

Взаємодія між об'єктами категорій K_1 і K_2 є представленням реальних земельними відносинами (рис.2), в яких:

– $\{Ob_1^{11}, ..., \{Ob_1^{12}, ..., \{Ob_1^{13}, ..., \{Ob_1^{14}, ..., Ob_m^{14}\}\}\}\} \subset K_1$ – вежа множин категорної моделі K_1 , де $Ob_1^{11}, Ob_1^{12}, Ob_1^{13}, Ob_1^{14}$ – закони, розділи, глави і статті земельного законодавства, відповідно;

– $\sum Mor(K_1, K_2)$ – множина морфізмів між малими категоріями;

– $M = \{m_1, m_2, ..., m_\xi\}, \xi = \overline{1, h}$ – множина схем вирішення задач в земельних відносинах;

– $D = \{d_j\}, j = \overline{1, y}$ – множина органів державної влади;

– $S = \{s_i\}, i = \overline{1, t}$ – множина власників, користувачів та розпорядників земельною ділянкою;

– $Z = \{z_p\}, p = \overline{1, l}$ – множина земельних ділянок;

– $G \subseteq ((D \times S \times Z) = \{(d_j, s_i, z_p) | d_j \in D, s_i \in S, z_p \in P\})$ – відносини між власниками, користувачами, розпорядниками, органами державної влади та земельними ділянками;

– $G = \{g_q\}, q = \overline{1, \delta}$ – множина відносин між власниками, користувачами, розпорядниками, органами державної влади та земельними ділянками;

– $m_h = \langle s_h, d_h, z_h, g_h \rangle$ – схема вирішення задачі в земельних відносинах.

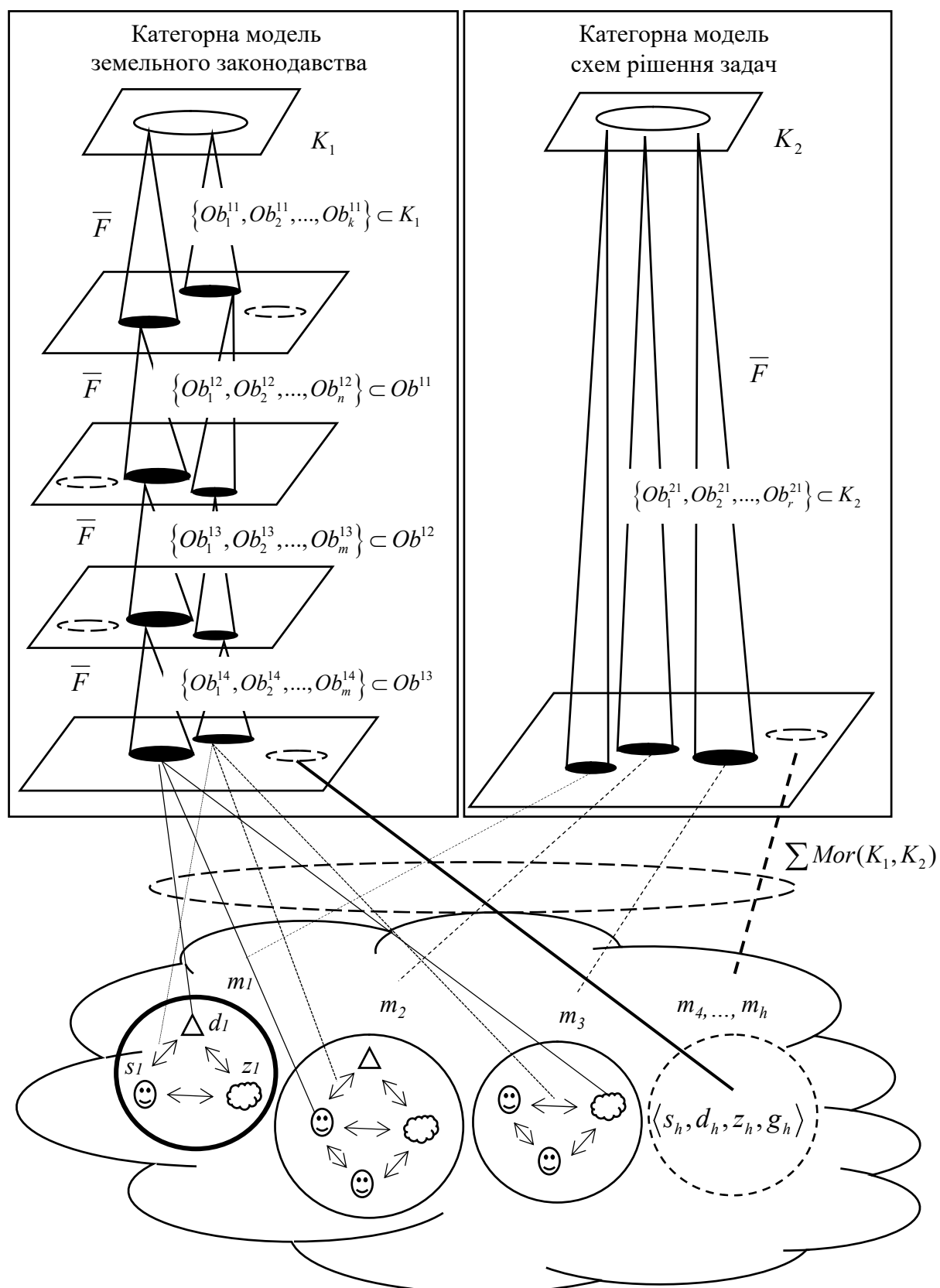


Рис. 2. Багаторівнева комутативна діаграма представлення земельних відносин в системі підтримки прийняття рішень

На основі моделей знань об'єктів та суб'єктів земельних відносин здійснюються процеси підтримки прийняття рішень, які представлені у вигляді алгоритму на рис.3.

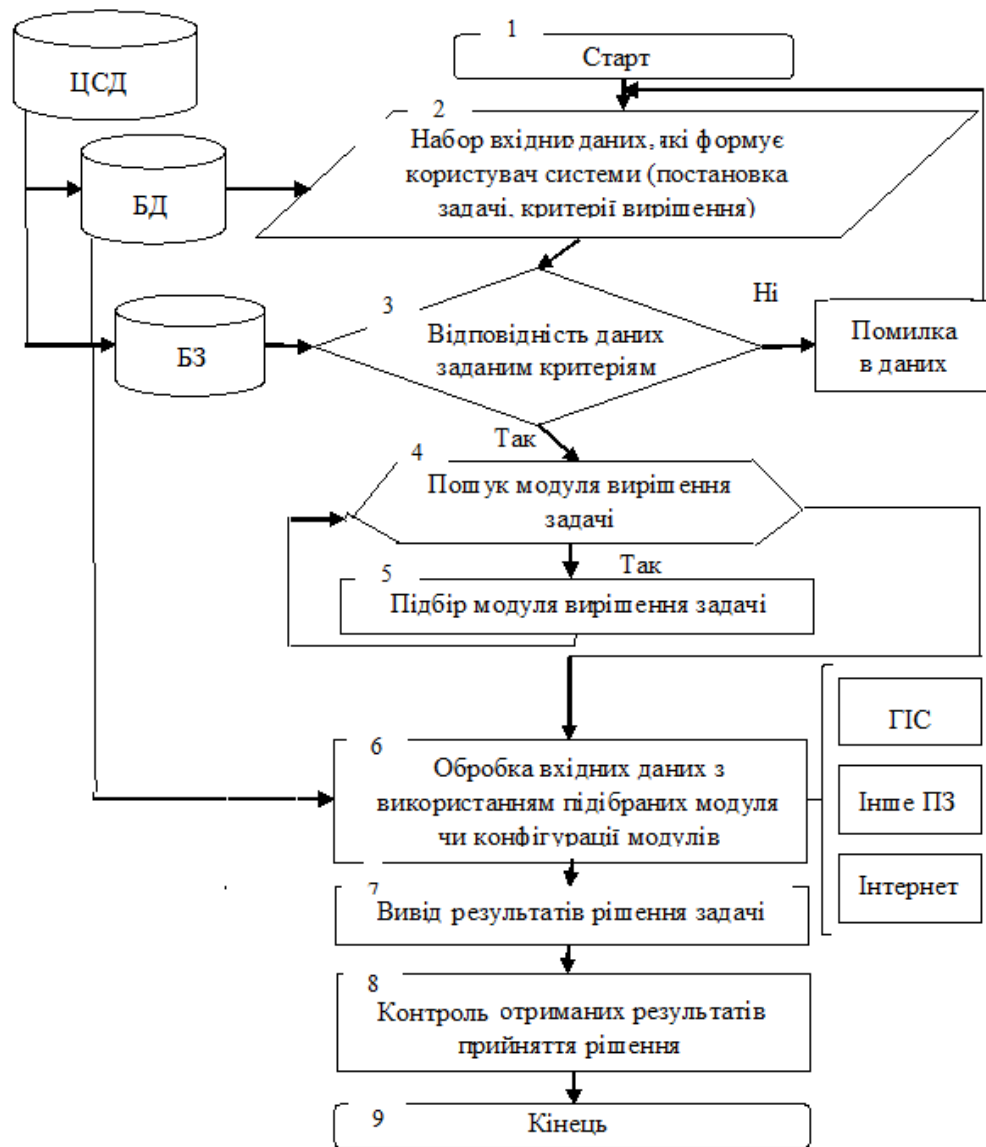


Рис. 3. Узагальнений алгоритм процесу підтримки прийняття рішень задач в сфері земельних відносин

Послідовність етапів реалізації алгоритму системи підтримки прийняття рішень наступна.

Блок №1. Старт роботи алгоритму. Визначається необхідність і доцільність використання системи підтримки прийняття рішень для вирішення поставленої задачі. При їх наявності процес переходить на наступний блок алгоритму.

Блок №2. Формування запиту користувача відповідно до атрибутивних даних, наявних в БД. Визначення критеріїв вирішення типових задач.

Блок №3. Перевірка відповідності даних запиту критеріям їх обробки і можливості їх використання тією чи іншою моделлю БЗ. У разі виявлення невідповідності формується повідомлення (наприклад, «Помилка в даних») – здійснюється перехід на блок №2 для коригування вихідного запиту.

Блок №4. Пошук модуля або конфігурації модулів системи підтримки прийняття рішень, що надходить на обслуговування задачі.

Блок №5. Вибір адміністратором модулів або їх конфігурації для вирішення поставленої задачі, що забезпечується на основі онтологічних моделей, представлених в БЗ.

Блок №6. Обробка вхідних даних з використанням допоміжного програмного забезпечення, яке представлено модулем або конфігурацією модулів в системі підтримки прийняття рішень наданого адміністратором.

Блок №7. Представлення результатів у вигляді і форматі, зручному для їх аналізу користувачем.

Блок №8. Здійснення контролю запропонованих рішень. При необхідності рішення коригуються.

Блок №9. Закінчення роботи алгоритму. Формується остаточний набір результуючих даних за рішенням задачі, заданої користувачем.

Система підтримки прийняття рішень, яка реалізує запропонований алгоритм, видає кінцеву множину рекомендацій щодо рішення поставленої задачі і критеріїв її рішення. У свою чергу, користувач має можливість використовувати таку систему в якості засобу автоматизації складного процесу впорядкування земельних відносин. Це дозволить: розвантажити і впорядкувати роботу служб, пов'язаних із земельними відносинами, послабити і в перспективі виключити повну залежність вирішення земельних питань від людського фактору, підвищити ступінь об'єктивності прийнятих рішень.

Як приклад можна запропонувати вирішення завдання аналізу деградованості земельних ділянок з метою їх моніторингу для виявлення якості ґрунту. Для цього введемо відповідні обмеження. Будемо вважати, що мала категорія K_1 складається з 10 об'єктів (земельної ділянки - z_1 , двох суб'єкти конфлікту - s_1 і s_2 , органу державної влади, який приймає рішення з приводу конфлікту d_1 і відносин між ними $Mor_1, Mor_2, Mor_3, Mor_4, Mor_5, Mor_6$, і K_2 складаються з одного об'єкта - M_1 , що описує схему рішення задачі (рис.4).

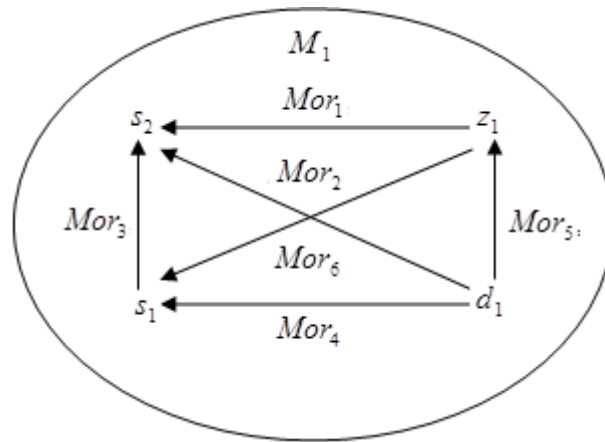


Рис. 4. Ілюстрація узагальненої схеми прикладу земельних відносин за допомогою теорії категорій

Для більш детальної розробки системи підтримки прийняття рішень кожен морфізм категорії в цій базі знань повинен являти собою конкретний процес, який буде ефективно розпізнаватися програмним кодом на комп'ютері. Для цього морфізми можна представити за допомогою логіки числення предикатів.

Так, наприклад, морфізм між власником і земельною ділянкою «ЯКЩО земельна ділянка є деградованою, ТО вона непридатна для використання, АБО земельна ділянка не є деградованою, ТО вона придатна для використання», можна записати в такий спосіб:

$$(Mor : s_1 \rightarrow z_1) \sim (((z_1 \in z_{\text{Iddeg}}) \rightarrow (z_1 \notin s_1)) \vee ((z_1 \notin z_{\text{Iddeg}}) \rightarrow (z_1 \in s_1))),$$

де z_{ldcg} – показник деградованої ділянки.

Висновки

Таким чином, отриманий узагальнений алгоритм, який покроково описує рішення задачі в сфері земельних відносин. У перспективі, удосконалюючи і деталізуючи запропонований алгоритм, можливо сформувати більш ефективну систему підтримки прийняття рішень в сфері земельних відносин. Формалізовані у математичному вигляді закони, якщо їх зібрати у базі знань, можна використовувати як базис для регулювання земельних відносин чітко, згідно закону; зменшити час вирішення конкретних питань у сфері земельних відносин; збільшити ефективність формування нових норм і правил земельного законодавства, з урахуванням прийнятих раніше законів; зменшити кількість незаконних дій посадових осіб в сфері земельного законодавства та ін. Результати роботи можуть бути впроваджені на підприємствах, які займаються моніторингом землі та діяльність яких взагалі пов'язана з земельними відносинами, а також можуть використовуватись в системах земельного адміністрування.

Список літератури

1. Merem, E.C. Using Geospatial Information Technology in Natural Resources Management: The Case of Urban Land Management In West / E.C. Merem, Y.A. Twumasi. // *Africa. Scientific Journal Sensors* (Basel). — 2008. — № 8 (2). — С. 607–619.
2. Шипулін, В.Д. Система земельного адміністрування: основи сучасної теорії: навч. посібник / В.Д. Шипулін. — Харків: ХНУМГ ім. О. М. Бекетова, 2016. — 220 с.
3. Лорьер, Ж.Л. Системы искусственного интеллекта / Ж.Л. Лорьер. — М.: Мир, 1991. — 568 с.
4. Метакидес, Г. Принципы логики и логического программирования / Г. Метакидес, А. Нероуд. — Т.: Факториал, 1998. — 288 с.
5. Компаниец, Р.И. Системное программирование. Основы построения трансляторов / Р.И. Компаниец, Е.В. Маньков, Н.Е. Филатов. — СПб: Корона принт, 2000. — 256 с.
6. Каверина, И.А. Курс лекций по математической логике и теории алгоритмов [Електронний ресурс] / И.А. Каверина. — М.: Ногинский колледж, 2016. Режим доступу: <https://infourok.ru/kurs-lekciy-po-matematicheskoy-logike-i-teorii-algoritmov-1147481.html> (Дата звернення: 28.11.2017).
7. Tripathi, K.P. Decision support system is a tool for making better decisions in the organization / K.P. Tripathi. // *Indian Journal of Computer Science and Engineering*. — 2011. — Vol. 2, No. 1. — Pp.112–117.
8. Ališić, I. Decision Support System to Urban Land Readjustment / I. Ališić, S. Mastelić Ivić, H. Tomić, M. Tomljenović. // 6th International Conference on Engineering Surveying Prague, Czech republic. — 2014. — Pp. 143-148.
9. Hou, J. Development of a decision support system for tsunami evacuation: application to the Jiyang District of Sanya city in China // J. Hou, Y. Yuan, W. Peitao, R. Zhiyuan, L. Xiaojuan. // *Nat. Hazards Earth Syst. Sci.* — 2017. — Pp.335-343.
10. Манько, Д.Г. Технології формалізації права / Д.Г. Манько. // *Науковий вісник Міжнародного гуманітарного університету*. — 2013. — № 5. — С. 18–21.
11. Радейко, Р.І. Формалізація як метод дослідження правових явищ [Текст] / Р. І. Радейко. — Львів: Інститут права та психології Національного університету «Львівська політехніка». — 2014. — № 8 (10). — С. 86–93.
12. Радейко, Р.І. Теоретико-правові аспекти вирішення проблеми формалізації права [Текст] / Р.І. Радейко. — Львів: Право і суспільство. — 2013. — № 6 (2). — С. 42–46.

АЛГОРИТМ ФУНКЦИОНИРОВАНИЯ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ В ЗЕМЕЛЬНЫХ ОТНОШЕНИЯХ

М.А. Кухар

Харьковский национальный университет городского хозяйства им. А.Н. Бекетова,
ул. Маршала Бажанова, 17, Харьков, 61002, Украина; e-mail: office@kname.edu.ua

В этом исследовании предложена актуальная задача использования онтологического инжиниринга земельных отношений, которая представлена методами алгоритмизации и системного анализа. С их помощью разработан алгоритм функционирования системы поддержки принятия решений в земельных отношениях, математические конструкции теории категорий и логики предикатов. Они отражают земельные отношения в логическом, структурированном виде и позволяют получить четкое понимание функционирования системы поддержки принятия решений при использовании этих моделей как элементов базы знаний. В перспективе, совершенствуя и детализируя предложенный алгоритм, возможно сформировать более эффективную систему поддержки принятия решений в сфере земельных отношений. Формализованные в математическом виде законы, если их собрать в базе знаний, можно использовать как базис для регулирования земельных отношений четко, согласно закону; уменьшить время решения конкретных вопросов в сфере земельных отношений; повысить эффективность формирования новых норм и правил земельного законодательства, с учетом принятых ранее законов; уменьшить количество незаконных действий должностных лиц в сфере земельного законодательства и т.д. Результаты работы могут быть внедрены на предприятиях, занимающихся мониторингом земли и, в общем, проводят свою деятельность в связи с земельными отношениями, а также могут использоваться в системах земельного администрирования.

Ключевые слова: алгоритм, формализация, теория категорий, логика предикатов, земельные отношения

ALGORITHM OF THE DECISION SUPPORT SYSTEM IN LAND RELATIONS

M.A. Kukhar

Kharkov National University of Municipal Economy named after A.N. Beketov,
17, Marshala Bazhanova Str, Kharkov, 61002, Ukraine; e-mail: office@kname.edu.ua

In the research proposed the actual task of using ontological engineering of land relations, which is represented by methods of algorithmization and system analysis. With their help the algorithm of functioning system of decision support in land relations, mathematical constructions of category theory and predicate logic. They reflect land relations in a logical, structured manner and provide a clear understanding of the functioning of the decision support system when using these models as elements of the knowledge base. In perspective, perfecting and detailing the algorithm, it is possible to form a more effective system of decision support in the sphere of land relations. Formalized in mathematical form laws, if they are collected in the knowledge base, can be used as a basis for regulating land relations clearly, according to the law; reduce the time of solving specific issues in the sphere of land relations; to increase the effectiveness of the formation of new norms and rules of land legislation, taking into account previously adopted laws; reduce the number of illegal actions of officials in the sphere of land legislation, etc. The results of the work can be implemented at enterprises, engaged in monitoring the land and, generally, conduct their activities in connection with land relations, and can be used in land administration systems.

Keywords: algorithm, formalization, category theory, predicate logic, land relations

МЕТОДЫ И АЛГОРИТМЫ ПОКРЫТИЯ

(Часть 2)

О.Н. Паулин

Одесский национальный политехнический университет,
просп. Шевченко, 1, Одесса, 65044, Украина; e-mail: paolenic@yandex.ua

В статье продолжается анализ методов и алгоритмов решения задачи о покрытии на предмет выделения макроопераций (МО), как моделей вычислительного процесса нахождения покрытия. Рассматриваются методы и алгоритмы с использованием предварительного сокращения таблицы покрытия (ТП). Такое сокращение возможно, если ТП обладает определёнными свойствами, а именно: наличие особенной «1» в строке, наличие нулевой строки, возможность поглощения строки/столбца. Приводятся словесные описания алгоритмов, их схемы, а также описания и схемы вычислительных модулей (частных процедур). Выделяются МО как вычислительные модели, которые частично обобщаются.

Ключевые слова: методы и алгоритмы покрытия, вычислительный процесс, макрооперация, вычислительная модель, особенная единица, нулевая строка, поглощение строки/столбца.

Введение

Данная статья является продолжением статьи [1], в которой рассмотрены 2 варианта перебора элементов последовательности: полный перебор и перебор по вогнутому множеству. Результат работы представленного ниже алгоритма – список выявленных ядерных строк и циклический остаток, который может быть обработан методом граничного перебора.

Идея рассмотренных ниже метода и алгоритма заключается в предварительном сокращении таблицы покрытия (ТП) на основе её свойств [2], выраженных соответствующими теоремами, выстроенных в определённом порядке. ТП может обладать (не обладать) следующими свойствами: наличие особенной «1» в строке, наличие нулевой строки, возможность поглощения строки/столбца.

Цель работы – сокращение времени анализа программ за счёт выделения макроопераций (МО) из вычислительного процесса обработки ТП.

Задача выделения МО решается аналогично [3].

Основная часть

Описание метода и алгоритма обработки ТП на основе её свойств. ТП может обладать (не обладать) следующими свойствами: строка – ядерная или антиядерная, столбец – поглощающий, строка – поглощаемая.

Метод состоит в последовательном рассмотрении перечисленных свойств и, при их наличии, – в сокращении ТП за счёт удаления определённых строк и/или столбцов. Эти свойства определяются соответствующими теоремами, представленными ниже.

Структуры данных:

M – матрица принадлежности или таблица покрытия, $M[i, j]$ – её элемент;

P – общий признак наличия изменений в ТП, $P = p_1 \vee p_2 \vee p_3 \vee p_4$;

p_i – частный признак наличия изменений в ТП для подпроцесса, $i=1..4$;

L – список ядерных строк.

Словесное описание алгоритма (COA), реализующего данный метод:

1. Общая процедура состоит в строгой последовательности применения теорем (1-я, 2-я, 3-я, 4-я и снова 1-я, 2-я, и т.д.).

2. Результат процедуры для каждой теоремы – удаление строк или/и столбцов ТП. Но возможны случаи неприменимости теорем.

3. Процедура для каждой теоремы заканчивается определением значения признака p_i ($p_i = 1$ в случае наличия изменений в ТП).

4. Формируется общий признак P наличия изменений в ТП ($P = p_1 \vee p_2 \vee p_3 \vee p_4$). Если в результате прохода вычислительного процесса по всем частным процедурам ТП не изменилась ($P = 0$), то вычислительный процесс закончен. Иначе общая процедура повторяется.

На рис. 1(а) приведена общая схема алгоритма (СА), построенная по COA.

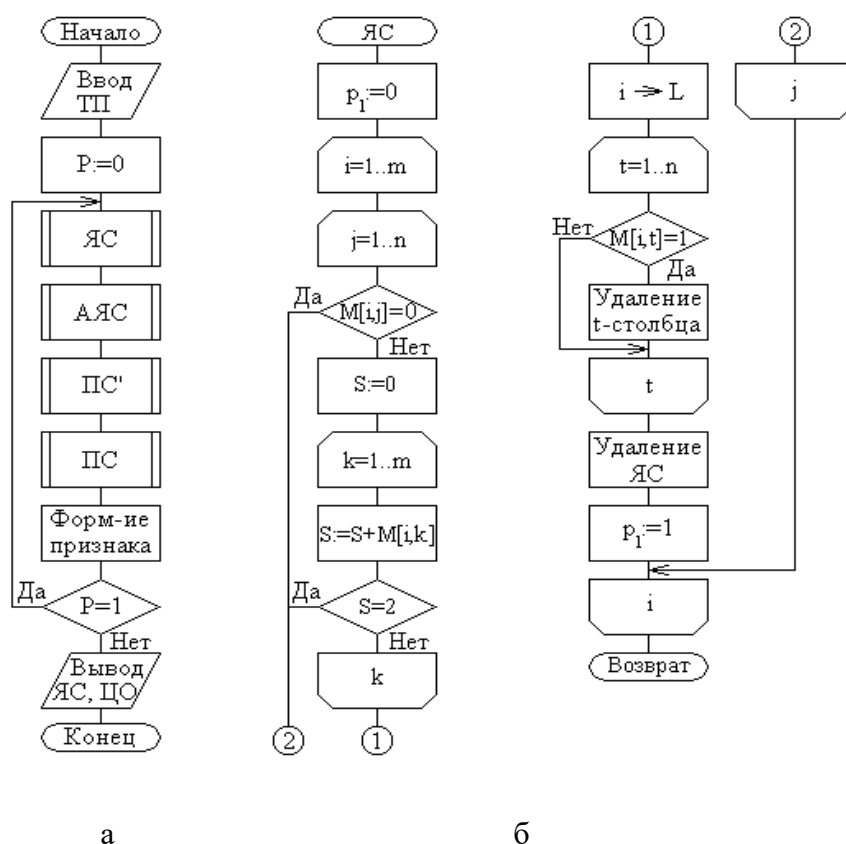


Рис. 1. СА: а – общей процедуры сокращения ТП ; б – нахождения ядерной строки

Рассмотрим частные процедуры (подпроцессы вычислений).

Сокращение ТП на основе теоремы о ядерной строке.

Теорема 1. Если в столбце ТП содержится единственная единица, то строка, содержащая эту единицу, входит во все покрытия и называется ядерной. Ядерная строка (ЯС) вычёркивается из ТП с запоминанием, вычёркиваются также все столбцы, покрываемые единицами ЯС [4].

Назовём единицу строки ТП особенной, если она в своём столбце является единственной – без неё покрытие принципиально невозможно.

Словесное описание процедуры нахождения ЯС:

1. Признаку p_1 присваивается значение 0.

2. Перебираются строки ТП в цикле по i , $i=1..m$

3. В текущей строке перебираются её элементы. Если $M[i, j]=0$, то переход на п.3.
 4. Для очередной «1» строки просматривается весь столбец ТП, содержащий эту «1» ($k=1..m$). При этом нулевые элементы игнорируются, а единичные суммируются ($k=1..m$). Если $S=2$, то переход на п. 3.
 5. Если она особенная ($S=1$), то текущая строка является ядерной; её номер заносится в список ЯС.
 6. Удаляются покрывающие столбцы ЯС ($t=1..m$), при $M[i, t]=1$.
 7. Удаляются ЯС. Признаку p_1 присваивается значение 1.
 8. Если перебраны все строки, то возврат в основную программу.
- На рис. 1(б) приведена СА нахождения ЯС.

Сокращение ТП на основе теоремы об антиядерной строке.

Теорема 2. (об антиядерной строке). Антиядерной называется строка, содержащая нулевые элементы. Эту строку можно удалить из ТП без запоминания [4].

Словесное описание процедуры нахождения антиядерной строки (АЯС):

1. Признак p_2 принимает значение 0.
2. Перебираются строки ТП. Если перебраны все строки, то переход на п. 5.
3. Просматриваются элементы текущей строки ($j=1..n$). При этом возможны два варианта: досрочный выход из цикла при $M[i, j]=1$ с переходом на п. 2 либо полный перебор при всех $M[i, j]=0$, т.е. получается антиядерная строка.
4. АЯС удаляется из ТП; $p_2 := 1$. Переход на п. 2.
5. Возврат.

На рис. 2(а) представлена СА частной процедуры «Нахождение АЯС»

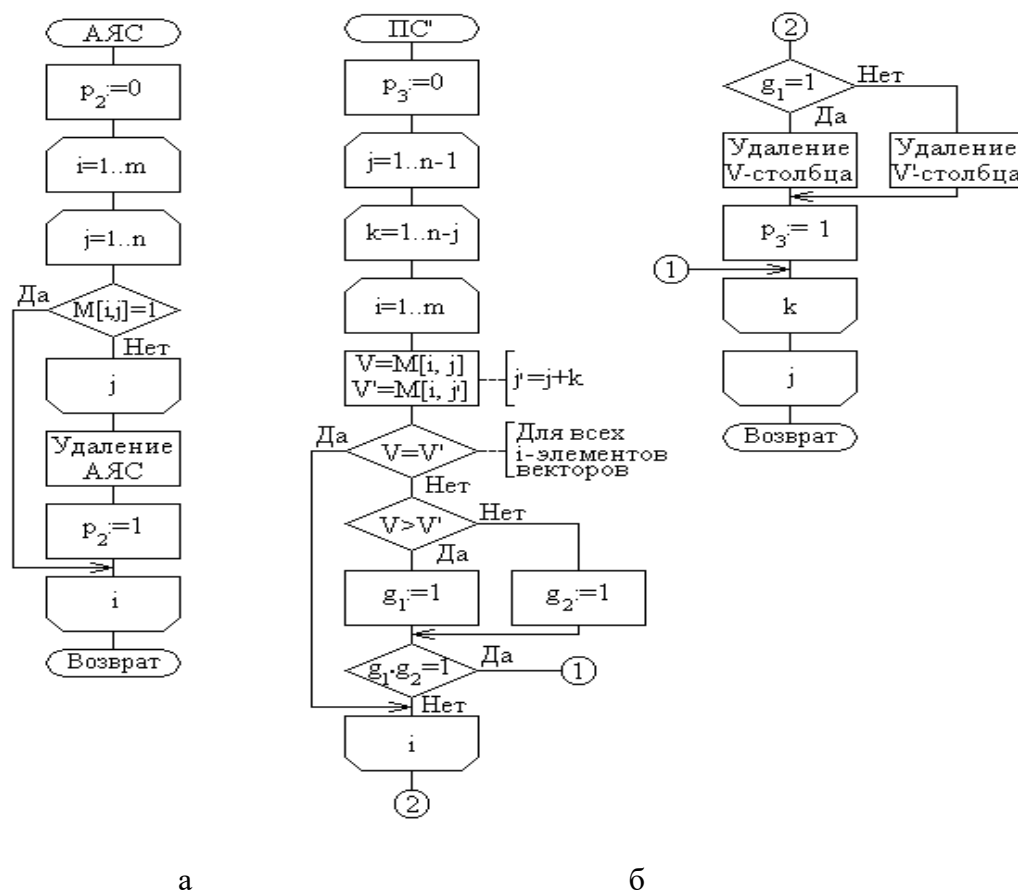


Рис.2. СА: а – нахождения антиядерной строки; б – нахождения поглощающего столбца

Прежде чем перейти к рассмотрению третьей и четвёртой теорем, познакомимся с понятием «поглощение».

Определение 1. Вектор $E = (e_1, \dots, e_m)$ поглощает вектор $F = (f_1, \dots, f_m)$, $E \geq F$, если для всех компонент e_i, f_i этих векторов можно одновременно записать $e_i \geq f_i$. В противном случае векторы называются *несравнимыми*.

Пример 1. Пусть заданы два вектора $E = (1, 1, 0, 1)$, $F = (0, 1, 0, 1)$. Сравнивая их поэлементно (например, слева направо), заключаем, что $E \geq F$.

Пример 2. Пусть заданы два вектора $E = (1, 1, 0, 1)$, $F = (0, 1, 1, 1)$.

Возможны 4 варианта результата сравнения элементов пары векторов (V, V') ; для отражения этих вариантов предлагается использовать 2 признака: g_1 и g_2 (табл. 1).

Таблица 1.

Признаки результата сравнения элементов векторов

Отношения элементов векторов	g_1	g_2
$V[i] = V'[i]$	0	0
$V[i] < V'[i]$	0	1
$V[i] > V'[i]$	1	0
Вектора несравнимы	1	1

Сокращение ТП на основе теоремы о поглощающих столбцах.

Теорема 3. В ТП могут быть вычеркнуты все поглощающие столбцы (рассматриваемые как двоичные векторы) без ущерба для построения всех безызбыточных покрытий [4].

СОА «Нахождение поглощающего столбца»:

1. Признак p_3 принимает значение 0; $g_1 = 0; g_2 = 0$.
2. Формируем пары столбцов. Для этого:
 - 2.1. Фиксируется j -й столбец в качестве первого компонента пары, $j = 1, \dots, n-1$.
 - 2.2. Перебираются остальные столбцы в качестве второго компонента пары с номером j' , $j' = j+k$, $k = 1, \dots, n-j$.
3. Представляем столбцы пары в виде двоичных векторов (V, V') с номерами j и j' соответственно.
4. В цикле по i , $i = 1, \dots, m$, сравниваем элементы пары векторов.

Если $V[i] = V'[i]$, то переход на п.4.

Если $V[i] < V'[i]$, то $g_1 = 1$, иначе $g_2 = 1$.

Если $g_1 \& g_2 = 1$, то переход на п. 2.2.
5. По окончании поэлементного сравнения проверяем $g_1 = 1$? Если – да, то удаляем столбец с номером j , иначе удаляем столбец с номером j' $p_3 = 1$. Переходим на п. 3.

6. Если просмотрены все пары, то возврат в основную программу.

На рис. 2(б) приведена СА нахождения поглощающего столбца.

Теорема 4. (о поглощаемых строках при поиске одного кратчайшего покрытия). Если при решении задачи о покрытии достаточно гарантировать получение хотя бы одного кратчайшего покрытия, то можно удалить все поглощаемые строки [4].

СОА «Нахождение поглощаемой строки»:

1. $p_4 := 0, g_1 := 0, g_2 := 0$.
2. Формируются пары строк. Для этого:
 - 2.1. Фиксируется i -я строка в качестве первого компонента пары, $i = 1, \dots, m-1$.

2.2. Перебираются остальные строки в качестве второго компонента пары с номером i' , $i'=i+k$, $k=1, \dots, m-i$.

3. Представляем строки пары в виде двоичных векторов (V, V') с номерами i и i' соответственно.

4. В цикле по $j, j=1, \dots, n$, сравниваем элементы пары векторов.

Если $V[i] = V'[i]$, то переход на п.4.

Если $V[i] \neq V'[i]$, то $g_1 = 1$, иначе $g_2 = 1$.

Если $g_1 \& g_2 = 1$, то переход на п. 2.2.

5. По окончании поэлементного сравнения проверяем $g_1 = 1$? Если – да, то удаляем строку с номером i' , иначе удаляем строку с номером i $p_4 = 1$. Переходим на п. 2.2.

6. Если просмотрены все пары, то возврат в основную программу.

СА нахождения поглощаемой строки не приводится, поскольку она в значительной степени повторяет рис. 4 с той разницей, что параметры i и j следует поменять местами. Кроме того, в случае проверки $g_1 = 1$? надо поменять местами вектора V и V' с номерами i и i' соответственно.

Разновидностью теоремы 4 является теорема 5.

Теорема 5. (о поглощающих строках при построении минимальных покрытий). Если при решении задачи о покрытии достаточно гарантировать построение всех (хотя бы одного) минимального покрытия, то можно вычеркивать поглощаемую строку, если цена её больше или равна цене поглощающей строки [4].

В этом случае п. 5 должен включать в себя проверку по цене строк.

Используя теоремы 1-4, упрощаем ТП. При этом возможны два исхода.

1. ТП после упрощения становится пустой (вычеркнуты все столбцы). В этом случае множество ядерных строк – требуемое покрытие.

2. Остаток ТП более не упрощается приёмами из теорем 1-4. Получаем циклический остаток (ЦО) таблицы покрытий. Покрытие ЦО таблицы можно строить методами граничного перебора либо разложения по столбцу.

Полное решение (покрытие исходной таблицы) состоит из ядерных строк и строк покрытия ЦО.

Общий список МО. Поскольку данный метод основан на методе полного перебора, то все его МО [1] сохраняются и в этом случае. Однако здесь появляются новые МО, связанные с новыми подпроцессами на основе свойств ТП и определением понятия «поглощения».

Общая процедура может быть представлена как МО высшего уровня, включающая в себя частные процедуры, которые могут быть представлены МО более низкого уровня: «Ядерная строка», «Антиядерная строка», «Поглощающие столбцы», «Поглощаемая строка», а также элементарные МО: «Формирование признака (логической формулы) окончания/прерывания вычислительного процесса» и «Альтернатива»

В свою очередь, МО для частных процедур также могут быть представлены через МО более низкого уровня: «Счётный цикл», «Условный цикл», «Поиск особенной «1» в строке», «Занесение номера ЯС в список», «Удаление строки (с двумя вариантами: ядерная и поглощаемая)», «Удаление столбца (с двумя вариантами: покрывающий и поглощающий)», «Присвоение переменной (признаку, флагу) определённого значения», «Поиск нулевой строки», «Представление строки/столбца двоичным вектором», «Сравнение элементов пары векторов», «Суммирование элементов строки/столбца и векторов», «Формирование циклического остатка».

Выводы

Рассмотрены методы и алгоритмы решения задачи о покрытии для случая предварительного сокращения таблицы покрытия с использованием её особенностей (свойств), сформулированных в виде теорем. Выделены макрооперации для каждого подпроцесса вычислений, которые собраны в единый список с частичным обобщением.

Список литературы

1. Паулин, О.Н. Вычислительные модели алгоритмов покрытия / О.Н. Паулин. // Информатика та математичні методи в моделюванні. — Одеса, ОНПУ. — 2016. — Т. 6, № 4. — С. 385–396.
2. Кнут, Д. Искусство программирования. Т.3. – Сортировка и поиск / Д. Кнут. — М.: Изд. Дом «Вильямс», 2000. — 832 с.
3. Паулин, О.Н. О выделении макроопераций из вычислительных процессов сортировки массивов данных [Текст] / О.Н. Паулин, Н.О. Комлевая, С.Ю. Марулин. // Проблеми програмування. — 2016. — № 2-3. — С. 87–95.
4. Новоселов, В.Г. Прикладная математика для инженеров-системотехников. Дискретная математика в задачах и примерах: учебное пособие / В.Г. Новоселов, А.В. Скاتков. — К.: УМК ВО, 1992. — 200 с.

МЕТОДИ І АЛГОРИТМИ ПОКРИТТЯ (Частина 2)

О.М. Паулін

Одеський національний політехнічний університет,
просп. Шевченка, 1, Одеса, 65044, Україна; e-mail: paolenic@yandex.ua

У статті продовжується аналіз методів і алгоритмів вирішення задачі про покриття на предмет виділення макрооперацій (МО), як моделей обчислювального процесу знаходження покриття. Розглядаються методи та алгоритми з використанням попереднього скорочення таблиці покриття (ТП). Таке скорочення можливо, якщо ТП володіє певними властивостями, а саме: наявністю особливої «1» в рядку, наявністю нульового рядка, можливість поглинання рядка/стовпчика. Наводяться словесні описи алгоритмів, їх схеми, а також опис і схеми обчислювальних модулів (окремих процедур). Виділяються МО як обчислювальні моделі, які частково узагальнюються.

Ключові слова: методи і алгоритми покриття, обчислювальний процес, макрооперація, обчислювальна модель, особлива одиниця, нульовий рядок, поглинання рядка\стовпчик

COVERING METHODS AND ALGORITHMS (Part 2)

O.N. Paulin

Odesa National Polytechnic University,
1, Shevchenko Ave., Odesa, 65044, Ukraine; e-mail: paolenic@yandex.ua

The analysis of methods and algorithms for solving the covering problem for the purpose identification of macrooperations (MO) finding as models of the computational process of finding the cover continues in the article. Methods and algorithms are considered using preliminary reduction of the coverage table (CT). Such a reduction is possible if the CT has certain properties, namely: the presence of a special "1" in the line, the presence of a null string, the possibility of absorbing a row/column. The verbal descriptions of algorithms, their schemes, as well as descriptions and schemes of computational modules (private procedures) are given. The MO are identified as computational models, which are partially generalized.

Keywords: methods and algorithms of coverage, computational process, macrooperation, computational model, singular unit, zero row, row/column absorption

МЕТОД РОЗПІЗНАВАННЯ ОПТИЧНИХ ЗОБРАЖЕНЬ У ХЕММІНГОВОМУ ПРОСТОРІ НА ОСНОВІ КРИТЕРІЮ СТРУКТУРНОЇ СКЛАДНОСТІ

Н.Я. Возна

Тернопільський національний економічний університет
вул. Львівська, 11, Тернопіль, 46020, Україна; e-mail: nvozna@ukr.net

В статті викладено теоретичні засади формування критерію структурної складності бінарних, напівтонових та кольорових оптичних зображень у Хеммінговому просторі, приведені аналітичні вирази оцінок Хеммінгової віддалі на основі різних взаємкореляційних функцій та наведені приклади застосування методу розпізнавання оптичних зображень у Хеммінговому просторі при вирішенні прикладних задач структуризації поліфункціональних даних у розподілених комп'ютерних системах. Результати досліджень мають ефективне застосування у галузі створення процесорів великої розрядності для шифрування даних. В основу таких пристроїв покладена функціональна схема пристрою для обчислення модульної кореляційної функції, завдяки чому досягнуто підвищення швидкодії пристрою визначення Хеммінгової віддалі між двома сигналами.

Ключові слова: структуризація, Хеммінгова віддаль, розпізнавання зображень, коефіцієнт взаємкореляції

Вступ

Розробка теорії структуризації даних є важливою та актуальною науковою задачею у галузі розпізнавання образів та оптичних зображень.

Вирішення прикладних задач у сфері застосування названої теорії тісно пов'язане з теорією розпізнавання образів та цифрового опрацювання зображень [1,2]. При цьому широко застосовуються методи кореляційного та спектрального аналізу у лінійному та квадратичному Хеммінговому просторах [3].

Для порівняння бінарних та багатоградацийних зображень успішно застосовано W -спектр зв'язності у матриці пікселів $m_x m_y m_z$ згідно виразів [1]:

$$W_{\Sigma x,y} = \sum_0^8 \sum_{i,j}^{m_x m_y} a_{i,j}^v; \quad W_{\Sigma x,y,z} = \sum_0^{26} \sum_l^{m_z} \sum_{i,j}^{m_x m_y} a_{i,j,l}^v,$$

де $a_{i,j}^v$, $a_{i,j,l}^v$ – елементи зображення із координатами відповідно i, j та i, j, l і зв'язністю v ($v \in \overline{1,4}$).

Порівняння зображень виконується згідно лінійної оцінки Хеммінгової віддалі (ω):

$$\Delta \omega_{hg} = |\omega_h - \omega_g|,$$

де h, g – відповідні функції зв'язності порівнюваних зображень.

В роботі [2] для кодування та розпізнавання зображень застосований принцип побудови морфологічної піраміди на основі структуруючого елементу та розрахунку математичного очікування, дисперсії, коефіцієнта взаємкореляції та ентропії.

При цьому для порівняння вхідного та профільованого зображень застосована квадратична оцінка Хеммінгової віддалі ймовірнісної ентропії по Шеннону згідно виразу:

$$\omega_{mn} = -10 \log \left(\frac{\frac{1}{2^k} \sum_{m=0}^{2^k-1} \sum_{n=0}^{2^k-1} [f(m,n) - f'(m,n)]^2}{2^k} \right); k = 8,$$

де $f(m,n), f'(m,n)$ – відповідні інтенсивності початкового та фільтрованого зображень, k – об'єм вибірки.

В якості кореляційної оцінки порівняння зображень застосований відомий коефіцієнт взаємочоваріації:

$$K_{mn}(0) = \frac{\sum_{m=0}^{2^k-1} \sum_{n=0}^{2^k-1} f_{mn} \cdot f'_{mn}}{\sqrt{(D_m + M_m^2) \cdot (D_n + M_n^2)}},$$

де D_m, D_n та M_m, M_n відповідні дисперсія та математичне очікування інтенсивностей порівнюваних зображень.

Мета роботи

Метою роботи є розробка теоретичних засад розпізнавання оптичних зображень у Хеммінговому просторі, представленому у теоретико-числовому базисі (ТЧБ) Крестенсона. Застосування оцінки Хеммінгової віддалі забезпечує можливість реалізації цифрових спецпроцесорів з високою швидкістю опрацювання структуризованих даних. Основним завданням даного дослідження є реалізація розробленого методу, що дозволить підвищити ефективність шифрування даних на основі багаторозрядних спецпроцесорів, які виконують операції модульної арифметики базису Крестенсона.

Основна частина

Теоретичні засади кодування зображень у різних теоретико-числових базисах. Існуюча практика структуризації та кодування елементів зображень у Хеммінговому просторі базується виключно на представленні кодів пікселів у двійковій системі числення ТЧБ Радемахера [4] (рис.1).

Кожен структурний елемент зображення певної інтенсивності описується вектором a :

$$a_{i,j}^v \in [m, n, l],$$

де m, n – двійкові коди декартових координат, l – код інтенсивності у випадку бінарного зображення.

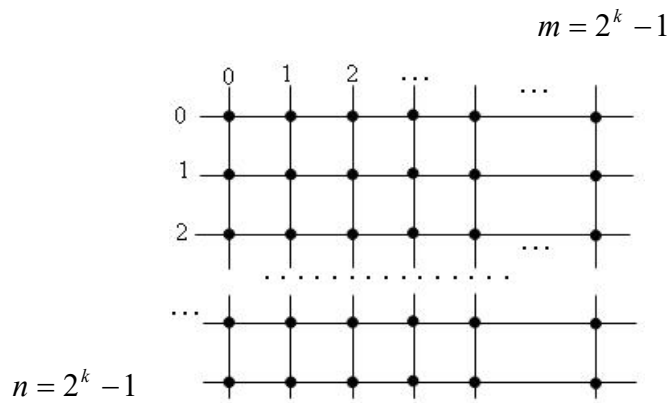


Рис. 1. Структуризація пікселів зображень у двовимірному Хеммінговому просторі

Для кольорових зображень відповідно кодуються інтенсивності трьох базових кольорів - s, t, z – відповідно синього, червоного та зеленого, тобто

$$a_{i,j}^v \in [m, n, s, t, z].$$

Таким чином, існуючі способи структуризації та кодування зображень у Хеммінговому просторі є багатовимірним, що потребує відповідної алгоритмічної та обчислювальної складності при виконанні їх порівняння та розпізнавання.

Більш ефективним способом структуризації та кодування елементів зображень у Хеммінговому просторі є застосування ТЧБ Крестенсона, який породжує систему числення залишкових класів [4,5] (рис.2).

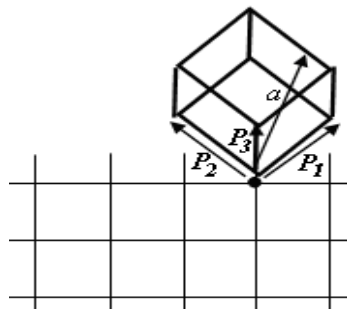


Рис. 2. Структуризація елементів зображень у ТЧБ Крестенсона

На рис.2: P_1, P_2, P_3 – взаємно прості модулі, a – результат кодування.

В основу цілочисельного перетворення системи залишкових класів (СЗК) покладена відома Китайська теорема про залишки, згідно якої будь-яке ціле число може бути представлено набором найменших невід'ємних залишків $(b_1, b_2, ..., b_i, ..., b_k)$ у системі взаємно простих чисел $(P_1, P_2, ..., P_i, ..., P_k)$.

Таким чином вектор структуризованих даних, який кодує елемент зображення у Хеммінговому просторі $a_{i,j}^v \in \langle m, n, z \rangle$ однозначно кодується та декодується у ТЧБ Крестенсона згідно виразів:

$$a_{i,j}^v = (b_1, b_2, ..., b_i, ..., b_k); a_{i,j}^v = \text{res} \sum_{i=1}^k (\text{mod } \rho); \rho = \prod_{i=1}^k P_i; b_{ik} = \frac{\rho}{P_i} \cdot m_i \cdot b_i,$$

де $(P_1, P_2, ..., P_i, ..., P_k)$ – система взаємно простих модулів; ρ – діапазон кодування даних, m_i – обернений елемент, $(1 \leq m_i \leq P_i - 1)$; b_{ik} – скоригований по m_i залишок b_i .

Кодування кольорових зображень RGB-системою в цілочисельній формі системи залишкових класів базису Крестенсона виконується згідно виразу:

$$N_k = res \sum_{i=1}^3 b_i \cdot B_i \pmod{P_0},$$

де N_k – число у позиційній системі числення; res – символ операції знаходження найменшого невід’ємного залишку; b_i – найменший невід’ємний залишок; B_i – базисні числа СЗК; $P_0 = P_1 \cdot P_2 \cdot P_3$, де P_1, P_2, P_3 – взаємно прості модулі.

Для однозначного кодування кольорів RGB-системи обрано наступні модулі:

$$P_1 = 255_{(10)} = 11111111_{(2)};$$

$$P_2 = 256_{(10)} = 100000000_{(2)};$$

$$P_3 = 257_{(10)} = 100000001_{(2)}.$$

$$\text{Звідси: } P_0 = 16776960_{(10)} = 1111111111 \ 1111110000 \ 0001_{(2)}.$$

Для здійснення перевірки знаходимо обернені елементи відповідних модулів:

$$m_1 = 128, \ m_2 = 255, \ m_3 = 129$$

та базисні числа:

$$B_1 = 8421376_{(10)} = 100000001000000000000000_{(2)};$$

$$B_2 = 16711425_{(10)} = 111111101111111100000001_{(2)};$$

$$B_3 = 8421120_{(10)} = 100000000111111100000000_{(2)}.$$

Перевірка правильності розрахунків виконується згідно виразу:

$$(B_1 + B_2 + B_3) \pmod{P_0} = 1$$

При цьому число градацій червоного кольору становить $0 \leq R \leq 254$, зеленого – $0 \leq G \leq 255$, синього – $0 \leq D \leq 256$.

Запропонований спосіб кодування у базисі Крестенсона дозволяє замінити вектори P_1, P_2, P_3 одним вектором P_0 , розрядність якого рівна 24 біти, що дорівнює розрядності кодів RGB-системи.

Практична реалізація визначення Хеммінгової віддалі. Відомий пристрій визначення залишку по модулю багаторозрядного числа [6] містить вхідну і вихідну шини, які є відповідно n -розрядними входами і n -розрядними виходами пристрою. В кожному розряді пристрою міститься однорозрядний суматор, прямий вхід переносу якого підключений до прямого виходу переносу суматора молодшого розряду пристрою, а прямий вихід переносу підключений до прямого виходу переносу суматора старшого розряду пристрою. Такий пристрій дозволяє визначити модульну різницю між двома двійковими кодами.

На рис.3 зображена структурна схема пристрою визначення залишку по модулю багаторозрядного числа, де 1 – вхідна шина, 2 – вихідна шина, 3 – спеціалізований однорозрядний суматор, 4 – D-тригер, 5 – вхід синхронізації, 6 – мультиплексор.

Даний пристрій використовується для виконання операції

$$(a + P_a) \pmod{P},$$

де P_a – доповнюючий код числа P ($P_a = \bar{P} + 1$), який використовується в асиметричних системах шифрування даних. Для отримання доповнюючого коду спочатку записують зворотний код початкового числа, для чого всі його розряди інвертують, а потім до отриманого після інвертування коду додають одиницю. На основі даного коду

проектується однокристална мікроелектронна реалізація багаторозрядного модульного суматора. Для можливості шифрування і розшифрування інформації такі кристали використовуються як апаратні засоби шифрування інформації.

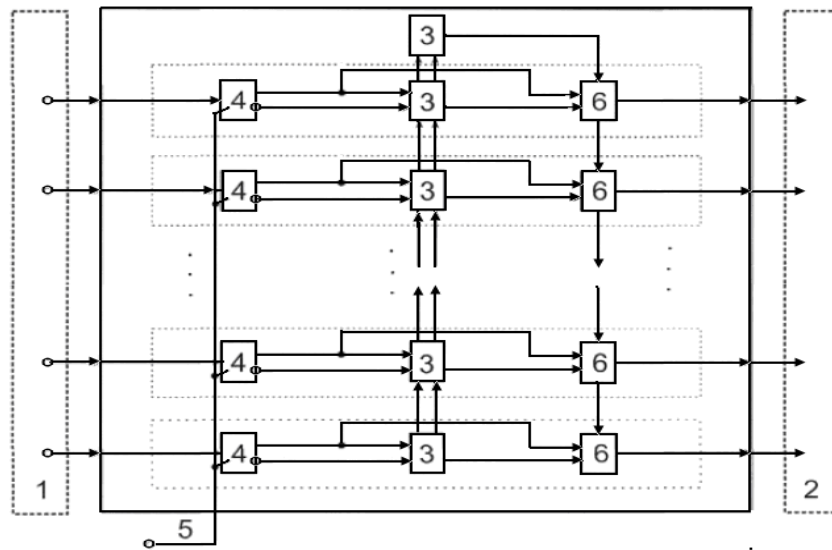


Рис. 3. Структурна схема пристрою визначення залишку по модулю багаторозрядного числа

На початку роботи пристрою (рис.3) вхідний код даних a ($0 < a \leq 2P-1$) подається на вхідну шину пристрою 1. Після подачі сигналу синхронізації у вигляді фронту наростання на вхідну шину 5 вхідний код даних записується в тригер 4 відповідного розряду суматора 3. Вихідні коди тригерів подаються на відповідні входи нульових чи одиничних спеціалізованих однорозрядних суматорів відповідного доповнюючого коду модуля P_a . У результаті підсумовування вхідного коду з кодом P_a та всіх наскрізних переносів у суматорах пристрою на виході $2k$ знакового розряду однорозрядного суматора (S_k) формується потенціал: 0, якщо $a \geq P$, тоді $b = (a + P_0) \bmod P$, інакше $a < P$ і $b = a$. Отриманий код b з виходів мультиплексорів 6 надходить на вихідну шину пристрою 2.

Однак недоліком такого пристрою є обмежені функціональні можливості, оскільки він не дозволяє визначити Хеммінгову віддаль між двома аналоговими сигналами на основі усередненої суми модульних різниць цифрових двійкових кодів. Іншим недоліком такого пристрою є низька швидкодія та висока апаратна складність обумовлена наявністю мультиплексора у його структурі.

Відомий також багатоканальний пристрій для обчислення модульної кореляційної функції [7], який призначений для статистичного аналізу випадкових процесів в реальному часі. Такий пристрій дозволяє визначити усереднену Хеммінгову віддаль між двома поточними аналоговими сигналами та їх Хеммінгову віддаль між двома зсунутими в часі вхідними аналоговими сигналами.

На рис.4 приведена структурна схема багатоканального пристрою для обчислення модульної кореляційної функції, де 1 – блок перетворення аналог-часовий інтервал; 2 – синхронізатор; 3 – група регістрів зсуву; 4 – елементи "Виключне АБО"; 5 – блок пам'яті.

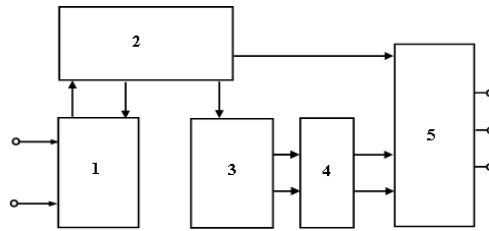


Рис. 4. Структурна схема багатоканального пристрою для обчислення модульної кореляційної функції

Недоліком такого пристрою є низька швидкодія обумовлена наявністю перетворювачів аналог-часовий інтервал, які формують вихідні послідовні цифрові унітарні коди з числом імпульсів 2^n , де n – розрядність відповідних двійкових кодів.

Запропоновано новий пристрій визначення Хеммінгової віддалі між двома аналоговими сигналами, де підвищення швидкодії визначення накопиченого усередненого значення суми модульних різниць оцифрованих аналогових сигналів відповідає оцінці Хеммінгової віддалі між ними згідно виразу:

$$Z = \frac{1}{n} \sum_{i=1}^n |x_i - y_i|,$$

де $i \in \overline{1, n}$ – об'єм вибірки цифрових кодів; x_i та y_i – вхідні оцифровані аналогові сигнали $x(t)$ і $y(t)$.

Такий пристрій визначення Хеммінгової віддалі може бути використаний у якості компонента цифрових процесорів розпізнавання образів.

Удосконалення пристрою полягає в тому, що вхідні аналогові сигнали у двоканальному АЦП паралельного типу перетворюються у паралельні унітарні коди, які однозначно представляють більше та менше значення x_i та y_i , після чого перетворюються у паралельні інверсні розрядно-позиційні коди Хаара, дешифруються у відповідні: прямий двійковий код більшого з них та доповнюючий код меншого з них, які додаються у швидкодіючому багаторозрядному накопичувальному суматорі.

На рис.5 та рис.6 представлена структурна схема пристрою визначення Хеммінгової віддалі між двома сигналами, який містить: 1 – синхронізатор; 2 – перший аналоговий вхід; 3 – перший компаратор i -го розряду; 4 – другий аналоговий вхід; 5 – другий компаратор i -го розряду; 6 – схема формування ступінчастих взірцевих потенціалів; 7 – вхід схеми формування ступінчастих взірцевих потенціалів; 8 – логічний елемент "АБО"; 9 – логічний елемент "І"; 10 – перший логічний елемент "І-НЕ"; 11 – перший логічний елемент "НЕ"; 12 – другий логічний елемент "І-НЕ"; 13 – другий логічний елемент "НЕ"; 14 – третій логічний елемент "І-НЕ"; 15 – перший багаторозрядний суматор; 16 – парафазний вхід логічної 1 0; 17 – накопичуваний суматор; 18 – вхід логічного 0; 19 – вхід скиду в 0 накопичуваного суматора; 20 – вхід синхронізації накопичуваного суматора 17; 21 – вихідна шина пристрою.

На початку циклу роботи пристрою на першому виході синхронізатора 1 формується сигнал скиду у нульовий стан накопичуваного суматора 17 по третьому входу 19, а на другому виході синхронізатора 1 на протязі циклу роботи пристрою формується n - синхронізуючих імпульсів, які поступають на треті входи всіх логічних елементів "АБО" 8 та логічних елементів "І" 9. Вхідні аналогові сигнали $x(t)$ $y(t)$ поступають на відповідні входи пристрою 2, 4, які у вигляді потенціалів U_y та U_x поступають на перші входи відповідних компараторів 3, 5, на другі входи яких з'єднані з відповідними виходами схеми формування взірцевих потенціалів 6, вхід якої з'єднаний з третім входом пристрою 7 взірцевого потенціалу U_0 .

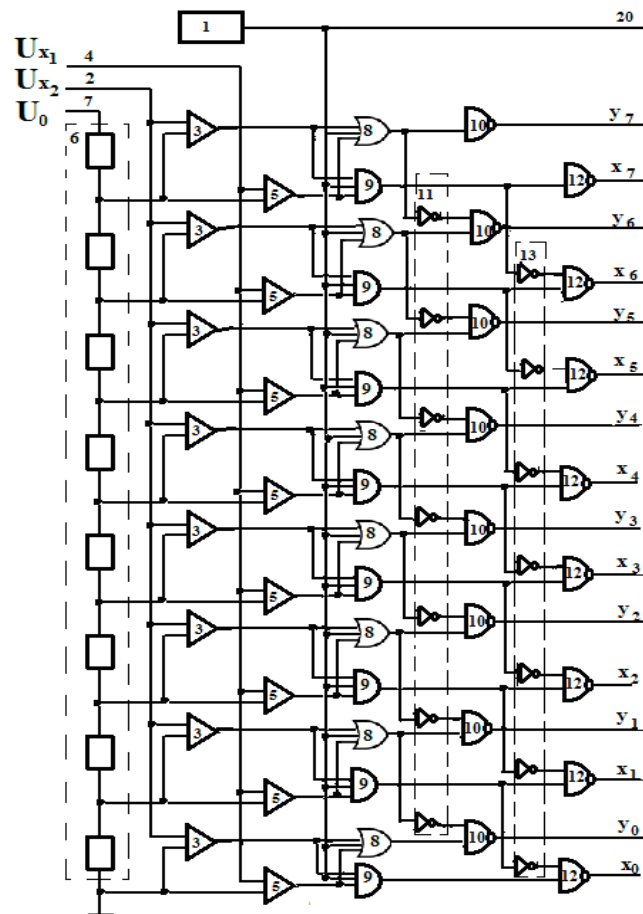


Рис. 5. Структурна схема пристрою визначення Хеммінгової віддалі між двома сигналами

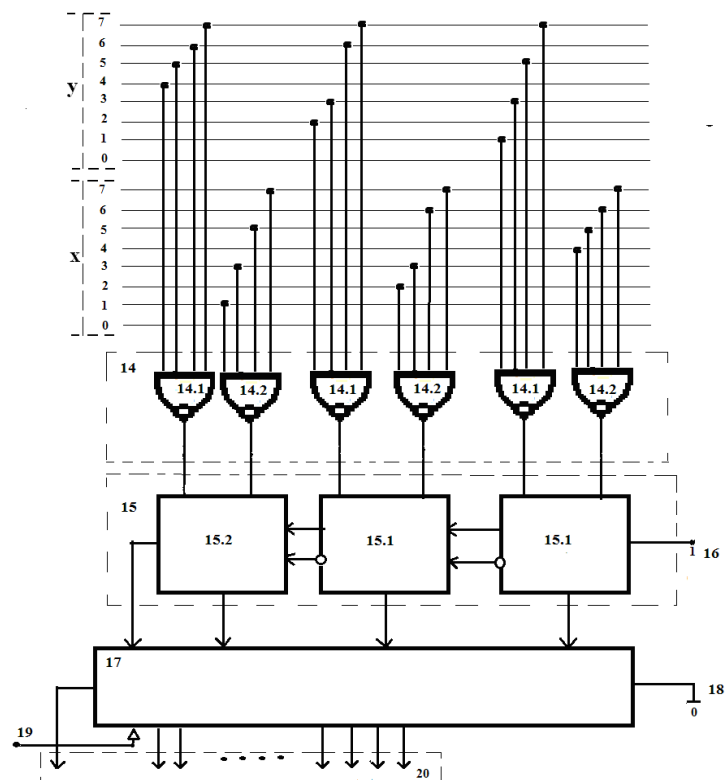


Рис. 6. Структурна схема пристрою визначення Хеммінгової віддалі між двома сигналами

При цьому на виходах компараторів 8 та 9 формуються відповідні паралельні унітарні коди вхідних потенціалів U_y та U_x , які поступають на входи відповідних логічних елементів "АБО" 8, І 9. Логічні елементи "АБО" формують інверсні паралельні унітарні коди більшого з двох цифрових значень x_i або y_i , а на виходах логічних елементів І 9 формуються паралельні інверсні коди меншого з двох цифрових значень x_i або y_i .

При цьому на виходах логічних елементів "І-НЕ" 10 формуються інверсні розрядно-позиційні паралельні коди Хаара більшого з двох цифрових значень x_i або y_i , які поступають на відповідні входи третіх логічних елементів "І-НЕ" 14.1, на виходах яких формуються прямі паралельні двійкові коди більшого із них, на виходах других логічних елементів "І-НЕ" 12 формуються інверсні розрядно-позиційні коди Хаара меншого з двох цифрових значень x_i або y_i , які поступають на відповідні входи третіх логічних елементів "І-НЕ" 14.2, на виходах яких формуються інверсні паралельні двійкові коди меншого із них. Отримані вихідні двійкові коди цифрових значень $\overline{x_i}$ та y_i поступають на входи багаторозрядного двійкового суматора 15 з парафазними входами та наскрізними переносами [8], де додаються з врахуванням логічної одиниці на вході суматора 16. У результаті на виході суматора 15 формується двійковий код модульної різниці між цифровим значенням x_i та y_i , сума яких накопичується в накопичувальному суматорі 17, із старших розрядів якого знімається двійковий код 20 значення Хеммінгової віддалі між двома аналоговими сигналами.

Наприклад, нехай $x_i = 6$, $y_i = 3$, тоді отримаємо наступні значення:

6	3	6 > 3	6	3		
0	0	0	0	1	1	
0	0	0	0	1	1	
1	0	1	0	0	1	
1	0	1	0	1	1	
1	0	1	0	1	1	
1	1	1	1	1	0	
1	1	1	1	1	1	
1	1	1	1	1	1	
x_i	y_i	$x_i > y_i$	$x_i > y_i$	$x_i > y_i$	$x_i > y_i$	
U	U	U	U	$\overline{H}$	$\overline{H}$	
1		2		3		
				6	3	
				1	0	
				1	1	
				0	1	
				$x_i > y_i$	$y_i < \overline{x_i}$	
				R	R	
				4		
						$6 + 4 - 1$
						$1 + 1 + 0 = 0$
						$1 + 0 + 0 = 1 \Rightarrow 3 = x_i - y_i $
						$0 + 0 + 1 = 1$
						$y_i > \overline{x_i}$
						R [R]
						5

У процесі визначення Хеммінгової віддалі між двома аналоговими сигналами виконуються наступні перетворення цифрових кодів x_i та y_i :

1. Перетворення аналогових сигналів $x(t)$ та $y(t)$ у вигляді паралельних унітарних кодів на виході компараторів (U);
2. Формування паралельних унітарних кодів більшого з двох x_i та y_i на виходах логічних елементів "АБО" (U);
3. Формування паралельних інверсних кодів Хаара на виходах логічних елементів "І-НЕ" ($\overline{H}$);
4. Формування паралельного коду Радемахера $y >$ та інверсного паралельного коду Радемахера $x <$ на виходах логічних елементів "І-НЕ" (R);
5. Виконання операції додавання прямого коду y_i , інверсного коду $\overline{x_i}$ та одиниці доповнюючого коду.

Висновки

Отже, в результаті розроблених теоретичних засад методу розпізнавання оптичних зображень у Хеммінговому просторі досягнуто підвищення швидкодії у порівнянні з відомими пристроями в 2^n рази пристроєм визначення Хеммінгової віддалі між двома сигналами, де n – розрядність паралельного АЦП. Це дозволяє пришвидшити шифрування інформаційних потоків, які представляють бінарні та напівтонові зображення.

Список літератури

1. Кожем'яко, В.П. Застосування W-спектру зв'язності для порівняння зображень / В.П. Кожем'яко, Л.І. Тимченко, Ю.Ф. Кутаєв, А.А. Яровий, Н.І. Кокрецька. // Оптико-електронні інформаційно-енергетичні технології. — 2002. — № 1 (3). — С. 10-16.
2. Харатишвили, Н.Г. Кодирование изображений на основе построения морфологической пирамиды / Н.Г. Харатишвили, И.М. Ихеидзе, З.Дж. Гошгашвили. // Оптико-електронні інформаційно-енергетичні технології. — 2002. — № 1 (3). — С. 54-62.
3. Кожем'яко, В.П. Оптико-електронний пристрій для кореляційного аналізу біозображень / В.П. Кожем'яко, О. Семенець, Хані Аль Зубі, Д.Ш. Гельман. // Оптико-електронні інформаційно-енергетичні технології. — 2002. — № 1 (3). — С. 101-103.
4. Акушский, И.Я. Машинная арифметика в остаточных классах / И.Я. Акушский, Д.И. Юдицкий. — М.: Сов. радио, 1978. — 256 с.
5. Николайчук, Я.М. Коды поля Галуа: теория та застосування. — М: Тернопіль: ТзОВ "Терно-граф", 2012. — 320 с.
6. Пат. 90144 Україна МПК G06F 7/00 (2014.01) Пристрій визначення залишку по модулю багаторозрядного числа / Я.М. Николайчук, В.Л. Кімак, О.І. Волинський, Б.Б. Круліковський № u201315351; заявл. 27.12.2013; опубл. 12.05.2014, Бюл. № 9, 2014.
7. Николайчук, Я.Н., Ищеряков С.М., Ластовецкий А.Д. Многоканальное устройство для вычисления модульной корреляционной функции SU 1686433, Бюл. № 39, 1991.
8. Пат. 109142 Україна МПК G06F 7/00 (2016.01) Однорозрядний суматор / А.Я. Давлетова, Б.Б. Круліковський, Н.Я. Возна, Я.М. Николайчук № u201602165; заявл. 04.03.2016; опубл. 10.08.2016, Бюл. № 15, 2016.

МЕТОД РАСПОЗНАВАНИЯ ОПТИЧЕСКИХ ИЗОБРАЖЕНИЙ В ХЕММИНГОВОМ ПРОСТРАНСТВЕ НА ОСНОВЕ КРИТЕРИЕВ СТРУКТУРНОЙ СЛОЖНОСТИ

Н.Я. Возна

Тернопольский национальный экономический университет,
ул. Львовская, 11, Тернополь, 46020, Украина; E-mail: nvozna@ukr.net

В статье изложены теоретические основы формирования критерия структурной сложности бинарных, полутоновых и цветных оптических изображений в Хемминговом пространстве, приведены аналитические выражения оценок Хеммингового расстояния на основе различных взаимокорреляционных функций и приведены примеры применения метода распознавания оптических изображений в Хемминговом пространстве при решении прикладных задач структуризации полифункциональных данных в распределенных компьютерных системах. Результаты исследований имеют эффективное применение в области создания процессоров большой разрядности для шифрования данных. В основу таких устройств положена функциональная схема устройства для вычисления модульной корреляционной функции, благодаря чему достигнуто повышение быстродействия устройства определения Хеммингового расстояния между двумя сигналами.

Ключевые слова: структуризация, Хеммингово расстояние, распознавание изображений, коэффициент взаимокорреляции

METHOD FOR RECOGNITION OF OPTICAL IMAGES IN THE HAMMING SPACE BASED ON THE STRUCTURAL COMPLEXITY CRITERION

N.Ya. Vozna

Ternopil National Economic University,
11, Lvivska Str., Ternopil, 46020, Ukraine; E-mail: nvozna@ukr.net

The theoretical principles of the formation of the criterion of structural complexity of binary, halftone and color optical images in the Hamming space are highlighted, analytical expressions of the Hamming distance estimates are given on the basis of different cross-correlation functions, and the examples of applying the method for recognition of optical images in the Hamming space while solving applied problems of structuring multifunctional data in distributed computer systems are presented in this paper. The research results can be used while developing high capacity processors for data encryption. These devices are based on a functional scheme of the device for calculating the modular correlation function. This leads to increased efficiency of the device for estimating the Hamming distance between two signals.

Keywords: structuring, the Hamming distance, image recognition, cross-correlation coefficient

ІНФОРМАТИКА ТА МАТЕМАТИЧНІ МЕТОДИ В МОДЕЛЮВАННІ

Том 7, номер 4, 2017. Одеса – 99 с., іл.

ИНФОРМАТИКА И МАТЕМАТИЧЕСКИЕ МЕТОДЫ В МОДЕЛИРОВАНИИ

Том 7, номер 4, 2017. Одесса – 99 с., ил.

INFORMATICS AND MATHEMATICAL METHODS IN SIMULATION

Volume 7, No. 4, 2017. Odesa – 99 p.

Засновник: Одеський національний політехнічний університет

Зареєстровано Міністерством юстиції України 04.04.2011р.

Свідоцтво: серія КВ № 17610 - 6460Р

Друкується за рішенням Вченої ради Одеського національного політехнічного університету (протокол №1 від 26.09.2017)

Адреса редакції: Одеський національний політехнічний університет,
проспект Шевченка, 1, Одеса, 65044 Україна

Web: <http://www.immm.opu.ua>

E-mail: immm.ukraine@gmail.com

Автори опублікованих матеріалів несуть повну відповідальність за підбір, точність наведених фактів, цитат, економіко-статистичних даних, власних імен та інших відомостей. Редколегія залишає за собою право скорочувати та редагувати подані матеріали

© Одеський національний політехнічний університет, 2017