

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

В І С Н И К

НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ
“ЛЬВІВСЬКА ПОЛІТЕХНІКА”

Видається з 1964 р.

№ 800

КОМП’ЮТЕРНІ НАУКИ
ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

Відповідальний редактор – д-р техн. наук, проф. Ю.М. Рашкевич

Львів
Видавництво Львівської політехніки
2014

Вісник Національного університету “Львівська політехніка” “Комп’ютерні науки та інформаційні технології” входить до переліку видань ВАК, в яких друкуються матеріали дисертаційних робіт у галузі технічних наук

У віснику публікуються статті за результатами досліджень з актуальних питань комп’ютерних наук та інформаційних технологій, виконаних професорсько-викладацьким складом Національного університету “Львівська політехніка” та провідними вченими України і зарубіжними вченими в галузі проектування архітектури і компонентів комп’ютерних систем, моделювання складних об’єктів, процесів і систем та розроблення і використання новітніх інформаційних технологій.

Для науковців, викладачів вищих закладів освіти, інженерів, що спеціалізуються у сфері новітніх обчислювальних систем, мереж, комп’ютеризованих засобів розв’язання задач цифрового опрацювання сигналів і зображень, автоматизованого проектування та керування, а також докторантів, аспірантів та студентів старших курсів відповідних спеціальностей.

*Рекомендувала Вчена рада Національного університету “Львівська політехніка”
(протокол № 69 від 25.02.2014 р.)*

*Свідоцтво про державну реєстрацію друкованого засобу масової інформації
(серія КВ № 13038-1922Р від 20.07.2007 р.)*

Редакційна колегія:

д-р техн. наук, проф. Ю.М. Рашкевич (відп. редактор);
д-р техн. наук, проф. Р.А. Мельник (заст. відп. редактора);
д-р техн. наук, проф. Д.Д. Пелешко (відп. секретар);
д-р фіз-мат. наук, проф. Я.П. Драган;
д-р техн. наук, проф. Р.П. Базилевич;
д-р техн. наук, проф. М.О. Медиковський;
д-р техн. наук, проф. В.В. Пасічник;
д-р техн. наук, проф. В.В. Різник;
д-р техн. наук, проф. Д.В. Федасюк;
д-р техн. наук, проф. Р.О. Ткаченко;
д-р техн. наук, проф. М.В. Лобур

Адреса редколегії:

Національний університет “Львівська політехніка”
вул. С. Бандери, 12, Львів-13, 79013,
корп. № 5, кімн. 515
akt805@ukr.net

АРХІТЕКТУРА ТА КОМПОНЕНТИ КОМП'ЮТЕРНИХ СИСТЕМ

004.032.26

O. Tyshchenko, I. Pliss
Kharkiv National University of Radioelectronics,
Control System Research Laboratory

AN EVOLVING RESERVOIR NEO-FUZZY NETWORK FOR TIME SERIES PREDICTION

© Tyshchenko O., Pliss I., 2014

Компютерне сховище – це парадигма навчання періодичних нейронних мереж на основі використання рекуррентної частини (так званого “резервуара”) інших показників. Еволюція систем визначила новий підхід, який фокусується на навчанні нечітких систем, що мають своїми параметрами їх структуру адаптації он-лайн. У даній роботі розвивається сховище нео-нечітка мережі, побудованої з використанням елементів затримки і нелінійних нео-нечітких синапсів, що означає еволюціонування систем і м'які обчислення об'єднуються в нову обчислювальну систему.

Ключові слова: сховище обчислювальне, еволюціонуєча система, гібридна система, нео-нечіткої нейрон, онлайн процедура навчання, прогнозування.

Reservoir Computing is a paradigm of training Recurrent Neural Networks based on treating the recurrent part (the so-called “reservoir”) differently from the readouts. This paradigm has become so popular recently due to its computational efficiency and the fact that it's enough to train only a supervised readout. Meanwhile Evolving Systems define a new approach which focuses on learning fuzzy systems that have both their parameters and their structure adapting on-line. In this paper an evolving reservoir neo-fuzzy network is built using time delay elements and nonlinear neo-fuzzy synapses which means that Reservoir Computing, Evolving Systems and Soft Computing are combined in a new computational system.

Keywords: reservoir computing, evolving systems, hybrid systems, neo-fuzzy neuron, online learning procedure, prediction.

Introduction

Reservoir Neural Networks (RNN) have been widely used in emulation and prediction tasks due to training a single linear output layer. Reservoirs are usually created randomly. So when the results don't have a required accuracy we have to start over training our network.

Evolving Systems (ES) are usually associated with streaming data and on-line (often real-time) modes of operation. They can be seen as adaptive intelligent systems with low-computational complexity. Evolving Systems assume on-line adaptation of system structure in addition to the parameter adaptation which is usually associated with the term “incremental”.

Due to the implementation of a wide variety of adaptive, evolving and dynamic methodologies, they represent an important cornerstone within the field of “data-driven learning in non-stationary environments” [1,2]. Combining the favorable properties of (data-driven) fuzzy systems (especially universal approximation capabilities in connection with comprehensibility and understandability aspects) with the concept of evolving modelling approaches may widen the applicability of fuzzy systems to on-line processing and modelling tasks. Under this scope, fuzzy systems inherit all the benefits and merits that evolving models have over batch modelling approaches [3-8].

Reservoir neural networks were basically designed to have the same computational power as traditional recurrent neural networks except the fact that there's no need to train internal weights [9-11]. The reservoir is a recurrent neural network which has n inputs, h internal units of the hidden layer and m output units. Input elements at a discrete time point k form a vector $x(k) = (x_1(k) \mathbf{K} x_n(k))^T$, internal units form a vector $\mathcal{X}(k) = (\mathcal{X}_1(k) \mathbf{K} \mathcal{X}_h(k))^T$ and finally output units form a vector $\mathcal{Y}(k) = (\mathcal{Y}_1(k) \mathbf{K} \mathcal{Y}_m(k))^T$. We take into consideration the simplest case when m equals to 1. So in this case the output signal is a scalar value.

Real-valued connection weights are collected in a $(h \times n)$ weight matrix W_{in} for the input weights, in an $(h \times h)$ matrix W_{res} for the internal connections, in an $m \times (h + n + m)$ -matrix W_{out} for the connections to the output units, and in a $(n \times m)$ -matrix W_{back} for the connections that feed back from the output to the internal units. The weight matrix W_{in} is usually created randomly and can be either full or sparsed. The weight matrix W_{res} is usually randomly created according to Gaussian distribution. It can be easily explained by the fact that a network creates a reservoir full of different nonlinear current and preceding values (the so-called "echo"). Connections directly from the input to the output units and connections between output units are allowed. A reservoir should have a suitable perceptibility. The simplest way to make it is to tune a spectral radius (the biggest eigenvalue) of weight matrix W_{res} .

The activation of internal units is updated according to expression

$$\mathcal{X}(k+1) = f(W_{in}x(k+1) + W_{res}\mathcal{X}(k) + W_{back}\mathcal{Y}(k)), \quad (1)$$

where $\mathcal{X}(k+1)$ is a vector of reservoir states at a time point k , where $f = (f_1 \mathbf{K} f_p)$ – the internal unit's output functions (typically sigmoid functions).

The output is computed according to equation

$$\mathcal{Y}(k+1) = f_{out} \left(W_{out} \begin{bmatrix} \mathcal{X}(k+1) \\ x(k+1) \end{bmatrix} \right), \quad (2)$$

where $f_{out} = (f_{out1} \mathbf{K} f_{outT})$ – the output unit's output functions.

The most interesting thing about reservoirs is that all weight matrices to the reservoir are initialized randomly, while all connections to the output are trained. When using the system after training with the "reservoir-output" connections, the computed output by is fed back into the reservoir.

Reservoir construction is largely driven by a series of randomized model-building stages, which rely on a series of trials and errors. Typical model construction decision of an echo state network involves setting the reservoir size, the sparsity of the reservoir and input connections, the ranges for random input and reservoir weights, and the reservoir matrix scaling parameter. A simple, deterministically constructed cycle reservoir is comparable to the standard echo state network methodology [12]. The short-term memory capacity of linear cyclic reservoirs can be made arbitrarily close to the proved optimal value.

In this case the reservoir is created with the help of time delay elements and bell membership functions instead of traditional sigmoid functions. The reservoir output depends linearly on the inputs. So we propose to use neo-fuzzy neurons [13-15] in the output layer of the reservoir. We also propose a rather simple and quick learning algorithm.

A disadvantage of the NFN model is that it assumes that the nonlinearities of the inputs are separable. So the NFN model is not a universal approximator, in contrast to the neural nets or to the conventional fuzzy systems. But it possesses better approximation properties comparing to linear models, as it has more degrees of freedom allowing good piecewise-linear approximation for many processes and imposing very low requirements on the computational resources. The most important advantage of the NFN model is its extremely fast learning which can be performed in an online mode with very simple weight update rules, or in just a single operation with the linear least-squares formula.

An evolving network architecture

The proposed evolving reservoir neo-fuzzy architecture sets up both network parameters and network architecture.

Neo-fuzzy neurons were proposed by T. Yamakawa and co-authors [15]. These constructions are neuronal models with nonlinear synapses. The output of the nonlinear neuron is obtained by sum of the synapses outputs represented by nonlinear functions. They can approximate a nonlinear input-output relationship by one neuron, and there is no local minimum problem in learning [13, 14]. Fig.1 shows a structure of the conventional NFN. An input signal $x(k)$ is fed into the NFN layer. It should be mentioned that this construction has back connections which come through the layer of delay elements from NFN outputs back to their inputs [16-18].

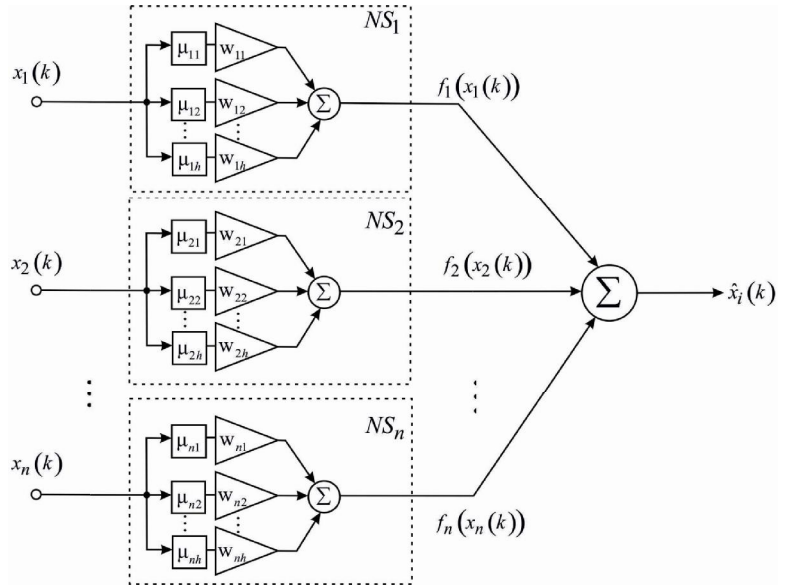


Fig.1. The structure of a conventional neo-fuzzy neuron which consists of nonlinear neo-fuzzy synapses

We could get a MISO-object (multi-input single-output) schema by joining advantages of the above-mentioned reservoir and evolving computing approaches and using a neo-fuzzy synapse as a basic element of our new system. The reservoir is formed with the help of time delay elements and nonlinear synapses. If the required accuracy is not high enough while processing data one more nonlinear synapse is added to the network which turns a system into a NARMA(1,1)-object. We could also get in a simple way a NARMA(2,1)-object.

The simplest NAR first-order architecture [15] based on neo-fuzzy synapses is in fig.1. $x(k)$ is an input value of the network, $\hat{x}(k)$ is a network output, $k = 1, 2, \dots, N, \mathbf{K}$ - current discrete time.

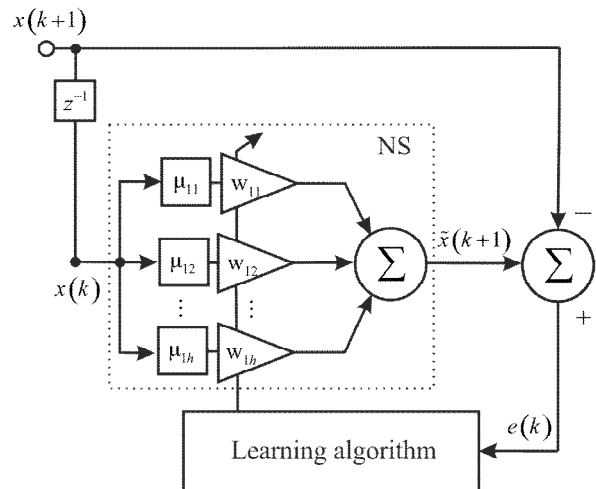


Fig.2. The simplest "reservoir" architecture

$$\mathcal{X}(k+1) = f(x(k)) = \sum_{j=1}^h m_{1j}(x(k))w_{1j}(k). \quad (3)$$

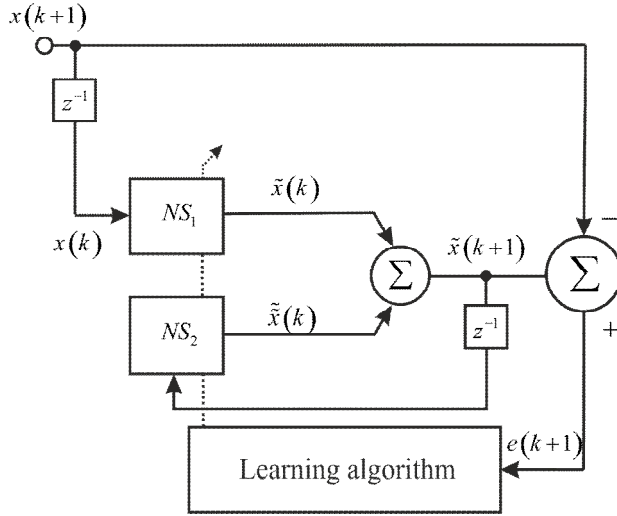


Fig. 3. NARMA(1,1)-object

$$\mathcal{X}(k+1) = f(\mathcal{X}(k)) + f(\tilde{\mathcal{X}}(k)), \quad (4)$$

$$f(\mathcal{X}(k)) = \sum_{j=1}^h m_{1j}(x(k))w_{1j}(k), \quad (5)$$

$$f(\tilde{\mathcal{X}}(k)) = \sum_{j=1}^h m_{2j}(\mathcal{X}(k))w_{2j}(k). \quad (6)$$

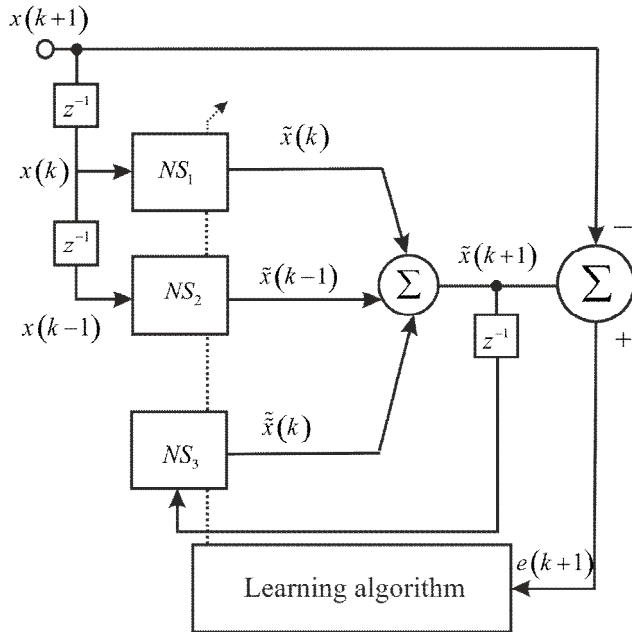


Fig. 4. NARMA(2,1)-object

$$\mathcal{X}(k+1) = f(\mathcal{X}(k)) + f(\mathcal{X}(k-1)) + f(\tilde{\mathcal{X}}(k)), \quad (7)$$

$$f(\mathcal{X}(k-1)) = \sum_{j=1}^h m_{2j}(x(k-1))w_{2j}(k). \quad (8)$$

Taking into consideration a NARMA(n,1)-object the output can be written

$$\mathcal{Y}(k+1) = f(\mathcal{Y}(k)) + f(\mathcal{Y}(k-1)) + \dots + f(\mathcal{Y}(k-n+1)) + f(\mathcal{Y}(k)). \quad (9)$$

Membership functions usually form a set of functions similar to the function array shown in fig. 5. In our case membership functions are evenly distributed in the range $[0,1]$.

It should be noticed that triangular membership functions provide piecewise-linear approximation which can lead to the results accuracy deterioration. To minimize this effect one may increase the amount of membership functions but this will increase the amount of synaptic weights and make a system architecture much more complicated as well as its learning algorithm. Cubic splines should be used as membership functions to get rid of the above-mentioned situations which can be written in the form:

$$m_j(x_i) = \begin{cases} 0.25 \left(2 + 3 \frac{2x_i - c_{ij} - c_{i,j-1}}{c_{ij} - c_{i,j-1}} - \left(\frac{2x_i - c_{ij} - c_{i,j-1}}{c_{ij} - c_{i,j-1}} \right)^3 \right), & x \in [c_{i,j-1}, c_{ij}], \\ 0.25 \left(2 - 3 \frac{2x_i - c_{i,j+1} - c_{ij}}{c_{i,j+1} - c_{ij}} + \left(\frac{2x_i - c_{i,j+1} - c_{ij}}{c_{i,j+1} - c_{ij}} \right)^3 \right), & x \in (c_{ij}, c_{i,j+1}], \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

Cubic splines satisfy Ruspini partition too and improve approximation characteristics of the fuzzy inference process. On the other hand, the usage of cubic splines provides smooth polynomial approximation and allows to model nonstationary signals with high accuracy results (fig. 5).

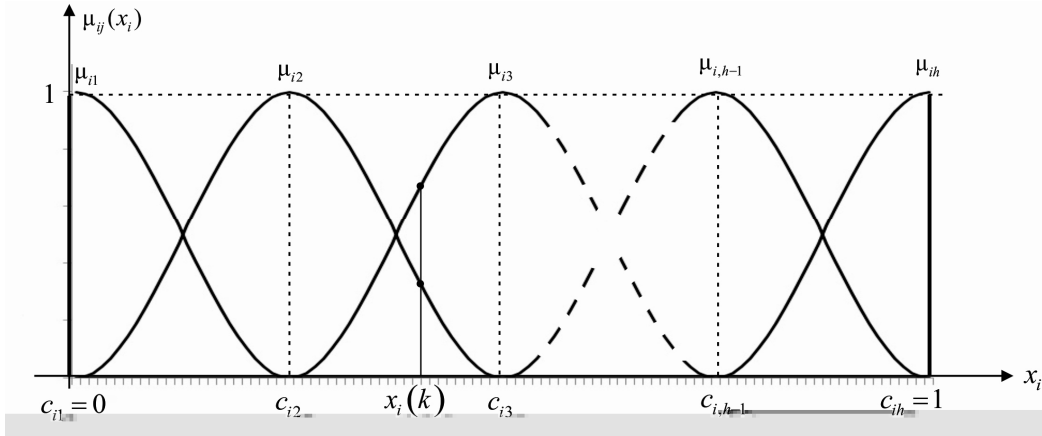


Fig. 5 – Cubic spline membership functions

Taking into consideration a $((n+1)h) \times 1$ -vector of membership functions values

$$\mathbf{m}(k) = \left(\left(m_{11}^{[q-1]}(x(k)); \mathbf{K}; m_{21}^{[q-1]}(x(k-1)); \mathbf{K}; m_{31}^{[q-1]}(x(k-2)); \right. \right. \\ \left. \left. m_{n1}^{[q-1]}(x(k-n+1)); \mathbf{K}; m_{n+1,h}^{[q]}(\mathcal{Y}(k)) \right) \right)^T,$$

where q stands for an amount of neo-fuzzy synapses in the evolving reservoir network architecture ($q-1$ says that the architecture hasn't been changed, q says that the architecture has been updated which means that new synaptic weights and membership functions have been added to the network architecture), and a synaptic weights vector of the same dimensionality

$$w(k) = \left(\left(w_{11}^{[q-1]}(x(k)); \mathbf{K}; w_{21}^{[q-1]}(x(k-1)); \mathbf{K}; w_{31}^{[q-1]}(x(k-2)); \right. \right. \\ \left. \left. w_{n1}^{[q-1]}(x(k-n+1)); \mathbf{K}; w_{n+1,h}^{[q]}(\mathcal{X}(k)) \right) \right)^T,$$

the network output can be presented in a vector form

$$\mathcal{Y} = w^T m(k). \quad (11)$$

A Learning Procedure

To find optimal values of synaptic weights the conventional learning criterion

$$E = \frac{1}{2} \sum_{k=1}^N \|y(k) - \hat{y}(k)\|^2 a^{N-k} \quad (12)$$

and an adaptive procedure are used:

$$\left\{ \begin{array}{l} w^q(k+1) = \begin{pmatrix} w_1^{q-1}(k) \\ \mathbf{L} \\ w_{n+1}^q(k) \end{pmatrix} + h^q(k) e(k-1) \begin{pmatrix} m_1^{q-1}(k+1) \\ \mathbf{L} \\ m_{n+1}^q(k+1) \end{pmatrix}, \\ h^q(k) = (r^q(k))^{-1}, \\ r^q(k+1) = ar^{q-1}(k) + \|m_1^{q-1}(k+1)\|^2 + \|m_{n+1}^q(k+1)\|^2, \\ 0 < a \leq 1. \end{array} \right. \quad (13)$$

Conclusion

The proposed architecture forms an evolving "reservoir" which is built using neo-fuzzy synapses. This network is designed to deal with nonstationary time series. The network has such advantages as numerical simplicity and high processing speed while comparing it to traditional forecasting neural networks and neuro-fuzzy systems.

1. Lughofer E. *On-line incremental feature weighting in evolving fuzzy classifiers // Fuzzy Sets and Systems*. – 163(1). – 2011. – P. 1-23. 2. Lughofer E. *Evolving Fuzzy Systems – Methodologies, Advanced Concepts and Applications // Studies in Fuzziness and Soft Computing*. – Springer, 2011. – 410p. 3. Angelov P., Lughofer E., Zhou X. *Evolving fuzzy classifiers using different model architectures // Fuzzy Sets and Systems*. – 159(23). – 2008. – P. 3160-3182. 4. Baruah R.D., Angelov P. *Evolving fuzzy systems for data streams: a survey // Wiley Interdisc. Rev.: Data Mining and Knowledge Discovery*. – 1(6). – 2011. – P. 461-476. 5. Angelov P., Zhou X. *Evolving Fuzzy-Rule-Based Classifiers From Data Streams // IEEE Fuzzy Systems*. – 16(6). – 2008. – P. 1462-1475. 6. Angelov P., Zhou X. *Evolving Fuzzy Classifier for*

Novelty Detection and Landmark Recognition by Mobile Robots // Mobile Robots. – 2007. – P. 89-118.

7. Angelov P. *A fuzzy controller with evolving structure // Information Science.* – 161(1-2). – 2004. – P. 21-35.

8. Kasabov N. *Evolving Connectionist Systems: The Knowledge Engineering Approach.* – Springer London, 2007. – 451p.

9. Alexandre L.A., Embrechts M.J. *Reservoir size, spectral radius and connectivity in static classification problems // ICANN 2009, Part I, LNCS 5768.* – Springer-Verlag, Berlin Heidelberg, 2009. – P. 1015-1024.

10. Jaeger H. *The echo-state approach to analysing and training recurrent neural networks // Technical report, German National Research Centre for Information Technology.* – 2001. – 45p.

11. Jaeger H. *Short-term memory in echo state networks // Technical report, German National Research Centre for Information Technology.* – 2001. – 36p.

12. Rodan A., Tino P. *Minimum complexity echo state network // IEEE Transactions on Neural Networks.* – 22(1). – 2011. – P. 131–144.

13. Miki T., Yamakawa T. *Analog implementation of neo-fuzzy neuron and its on-board learning // Computational Intelligence and Applications.* – Piraeus: WSES Press, 1999. – P. 144-149.

14. Uchino E., Yamakawa T. *Soft computing based signal prediction, restoration and filtering // Intelligent Hybrid Systems: Fuzzy Logic, Neural Networks and Genetic Algorithms.* – Boston: Kluwer Academic Publisher, 1997. – P. 331-349.

15. Yamakawa T., Uchino E., Miki T., H. Kusanagi. *A neo fuzzy neuron and its applications to system identification and prediction of the system behavior // Proc. 2-nd Int. Conf. on Fuzzy Logic and Neural Networks "IIZUKA-92".* – Iizuka, Japan, 1992. –P. 477-483.

16. Bodyanskiy Ye., Tyshchenko O. *Neo-fuzzy forecasting echo state network // Proc. 4th Int. Conf. ACSN-2009.* – Lviv: NVF "Ukrainski Tehnologii", 2009. – P. 95-96.

17. Bodyanskiy Ye., Tyshchenko O. *The reservoir predictive neuro-fuzzy network // Proc. Int. Conf. on Intellectual Systems for Decision Making and Problems of Computational Intelligence.* – Vol.1. – Kherson: KhNTU, 2010. – P. 279-282.

18. Tyshchenko O., Pliss I. *The forecasting neuro-neo-fuzzy network based on reservoir computing // Control Systems, Navigation and Connection Systems.* – No.1(21), Vol.2. – Kyiv, 2012. – P. 123-126.

ABOUT SOME DESIGN PRINCIPLES OF INFORMATION-RETRIEVAL SYSTEM AND PROCESSING OF ELECTRONIC DOCUMENTS IN INTERNET

Ó Osypenko V., Shvorov A., 2014

У статті аналізуються основні тенденції пошуку і основний напрямок виявлення пріоритету електронних документів при обробці в Інтернеті. Технічні принципи побудови інформаційно-пошукової системи і здатність використовувати процесор і нейронні мережі для вирішення завдань, визначають важливість розгляду вхідних документів об'єктивно. Об'єктом дослідження є системи, що підтримують рішення, призначені для визначення пріоритету в розгляді документів, які надходять. Пропонована система може активно використовуватися в інформаційній підтримці блоків індуктивної технології системи інформаційно-аналітичного дослідження.

Ключові слова: електронний документ, вхідний документ, лінгвістичний процесор, нейронна мережа.

The paper analyzes the main trends of search the basic direction and identifying the priority of electronic documents handling in the Internet. The technical principles of construction of information retrieval system and ability to use linguistic processor and neural networks for solving problems determining the importance of consideration the input documents were grounded. The object of the study was the decision support systems designed to determine priority of review incoming documents. The proposed system can be actively used in information support blocks of inductive technology of system information-analytical research.

Keywords: electronic documents, incoming documents, linguistic processor, neural network.

Introduction

One of the complex and specific types of operator activity in a information retrieval systems (IRS) in the Internet is an activity that is related to the analysis of information flows in order to highlight and primary handling the most important the input electronic documents (IED). This raises a need to in the preceding of automatic abstracting of each IED as well as for analyzing of complex information parameters to assess their importance.

1. The analysis of recent research and purpose of papers

Our analysis reveals that one of the efficient processing ways of IED given their importance is widely used of decision support systems (DSS) to determine the sequence of processing IED [1-3]. However, currently not fully in modern literature is considered the preliminary automatic abstracting of each IEDs as well as an analysis of complex of informative parameters to assess their importance by using the linguistic processor (LP) as a component of the DSS.

The purpose of this paper is to determine the basic elements of design the search engine in the Internet and the structure of DSS in terms of determining the sequence of IED consideration in view of their importance.

2. The main material of research

In developing the IRS, which is an integrated expert system, the following main stages of its creation can be allocated: identification, conceptualization, formalization, implementation, testing, maintenance testing [3].

At the stage of identification are defined the tasks to be resolved using IRS during its operation, and turns out the goal of development as well as the necessary resources. Stage conceptualizing involves the meaningful analysis and justification of methods for solving functional tasks IRS. At the stage of formalization are determined the detection methods and interpretation of all kinds of knowledge, simulated the operation of the system, executed the assessing the degree conformity of planned and achieved goals IRS operation, assessed the adequacy of representation and manipulation of knowledge. Stage performance is characterized by obtaining of knowledge from experts.

At the stage of testing by an expert group is carried out the check of possibility of using an integrated expert system in the process of finding relevant information in the Internet. Testing continues for as long as the expert group concludes that system reached the required level of competence.

At the stage of exploitation testing is verified the suitability of IRS to solve functional problems and the degree of goal achievement. As a result of this phase you may need the substantial modification of system. The most difficult is the initial stages of IRS construction of, that requires further research.

So, on considered the stages provided justification of approaches to implementation the following basic functions of IRS: IED-search by keywords and parameters; forming the abstract of IED; determining the sequence of issuance abstracts IED to the operator depending on their importance; IRS opportunity for self-learning, as well as the convenient interaction of operator with IRS.

In order to form abstract IED there is a need for a special linguistic support ICS, based on a linguistic database (LDB), which includes various dictionaries specified format [4]. With LDB linguistic processor to solve the problem of decomposition of the text IED for these abstract components of the IED.

The sequence of IED abstracting includes the following stages:

- nomination of the previous hypotheses about the meaning of all IED;
- determining the values of obscure words (special terms);
- create a common hypothesis (about knowledge);
- clarify the meaning of terms and interpretation and interpretation of certain text fragments under the influence of the general hypothesis (from the whole to particle);
- forming semantic structure of the text IED through the installation of internal relations between key words and fragments, as well as due to the formation of abstract concepts, generalizing particular fragments of knowledge;
- adjusting the overall hypothesis about the meaning of IED and interpretation of individual pieces of text under the influence of the general hypothesis (from particle to whole);
- adoption of basic hypotheses.

Thus, on the basis through the use of both deductive (from whole to particle) and inductive (from particle to whole) components provided the abstracting of IED.

In addition, by using LP determined the importance of IED that characterizes the content and urgency of working IED, as well as the priority of destination from which received the IED. To solve the classification problem, i.e., classifying IED, characterized by a set of indicators of how important or most important, it is proposed to use the so-called hybrid fuzzy classifier [5-7]. This classifier is a system that combines both structural and in functional terms, the principles of neural network models and fuzzy logic data processing, respectively.

The structure of the neuron consists of multipliers (synapses) an adder and nonlinear (linear) converter. Synapses provide the connection between neurons and multiple the input signal on number that characterizes the strength of connection - synapse weight. The adder unites signal fed by the synaptic connections from other neurons and external inputs. Nonlinear (linear) converter realizes the function of one argument – the output of the adder. It is called as “activation function” or “transfer function” of the neuron. Neuron generally realizes the scalar function of a vector argument.

A mathematical model of a neuron is described by the relations:

$$s = \sum_{i=1}^n w_i \cdot x_i + b, \quad (1)$$

where: W_i , ($i=1, \dots, n$) – the weight of the synapse; b – bias value; x_i – component of the input vector (input), ($i=1, \dots, n$); s – the result of the addition; y – output of neuron; n – the number of neuron inputs; f – transformation (activation function or transfer function).

Thus, the neuron is completely described by its weights w_i and transfer function $f(s)$. After receiving a set of numbers (vector) x_i , the neuron produces a some number y on output. That is, neural networks work is to convert the input vector X in the output vector Y , and this transformation is given by network weights.

Let us consider in more details the process of functioning of the neural network to determine the importance of the IED. The neural network generates an output signal Y according to the input signal X , realizing the some functions $Y = G(X)$. Type of function G determined by the values of synaptic weights and network displacements. Then, let the solution of some problem is a function $Y = F(X)$ defined by pairs of input-output data $(X^1, Y^1), (X^2, Y^2), \dots, (X^N, Y^N)$ for whom $Y^k = F(X^k), k=1, 2, \dots, N$. Accordingly, the study is the synthesis function G , which is close to F in the sense of a function error E .

If the set of training examples are selected, i.e. pairs of $(X^k, Y^k), k=1, 2, \dots, N$ and method of calculating of error function E is determined, then the training of the neural network is transformed into a multi-dimensional optimization problem, which has a very large dimension. In this case, since the function E can be arbitrary form, in general – it is many-extreme nonconvex optimization problem.

Classically, neural networks are used to: the classification of images, clustering (categorization), approximation, prediction/forecasting, optimization tasks, management etc. However, in view of the ability of neural networks to adequately operate under uncertainty and incompleteness of information, it is advisable to use them to acquire new knowledge about the objects.

The method of such application of neural networks shall test to recognize of importance of IED keywords.

According to the scoring method, the input network parameters are:

- density of the keyword in the title IED ($Z1$);
- weight of keywords in IED ($Z2$);
- frequency of keywords on a web page ($Z3$);
- frequency of keywords in IED ($Z4$);
- factor taking into account the location keyword ($X1$): the title page (*title*); location of keyword in meta tags; appear on the page; replaces the text in the picture, as well as in a hyperlink.
- clearance of keyword: style; font size; font fat content ($X2$);
- site redesign ($X3$);
- correction factor ($X4$).

During the study, other factors were not used. The output of the network will generate the importance of keywords divided by the expert of importance in standard units (s.u.).

To construct a mathematical model used the linear neural network with the corresponding simple analytical expression:

$$Y = X * W + B, \quad (2)$$

where: Y – the network output; W – weight ratio; X – input; B – bias.

In demos Statistica 6.1 software environment were synthesized the linear neural network IRS (Fig. 1) with characteristics: training error (RMS of error for all outputs) - 0.103727 s.u., control error – 0.151430 s.u., test error – 0.123600 s.u.

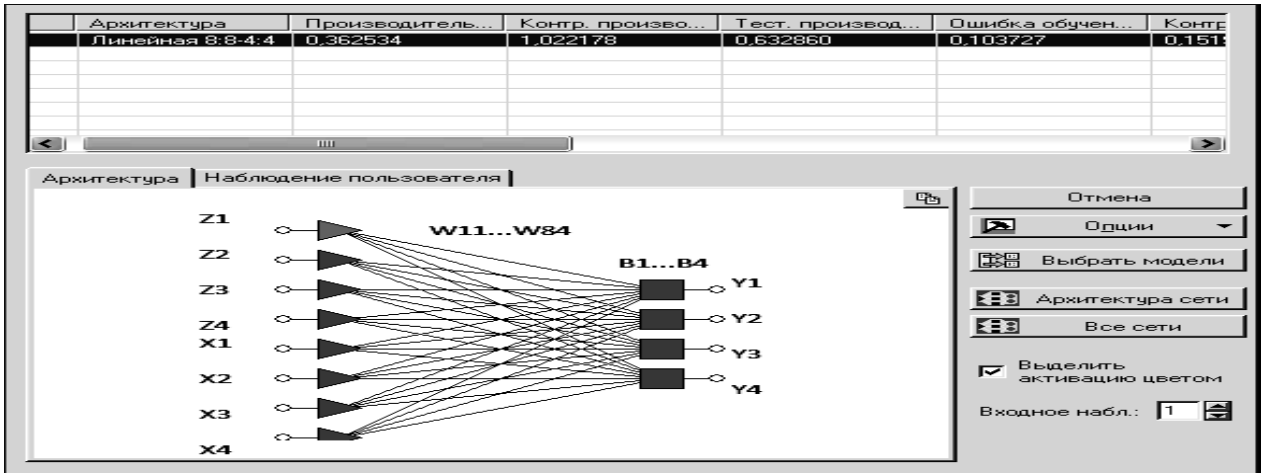


Fig. 1. Neural network architecture

On Fig. 1: $W11 \dots W84$ – weights (all of them – 32, the first number – is the number of input parameters, the second - weight coefficients of output); $B1 \dots B4$ - biases.

Then, analytical expressions for each of the 4 outputs of the neural network are:

$$Y1 = W11 * Z1 + W21 * Z2 + W31 * Z3 + W41 * Z4 + W51 * X1 + ; \quad (3)$$

$$+ W61 * X2 + W71 * X3 + W81 * X4 + B1$$

$$Y2 = W12 * Z1 + W22 * Z2 + W32 * Z3 + W42 * Z4 + W52 * X1 + ; \quad (4)$$

$$+ W62 * X2 + W72 * X3 + W82 * X4 + B2$$

$$Y3 = W13 * Z1 + W23 * Z2 + W33 * Z3 + W43 * Z4 + W53 * X1 + ; \quad (5)$$

$$+ W63 * X2 + W73 * X3 + W83 * X4 + B3$$

$$Y4 = W14 * Z1 + W24 * Z2 + W34 * Z3 + W44 * Z4 + W54 * X1 + \quad (6)$$

$$+ W64 * X2 + W74 * X3 + W84 * X4 + B4$$

The expression of the objective function for the optimization problem in analytical form is as follows:

$$F = \sqrt{(Y1 - R1)^2} + \sqrt{(Y2 - R2)^2} + \sqrt{(Y3 - R3)^2} + \sqrt{(Y4 - R4)^2} \rightarrow 0, \quad (7)$$

where: $R1 \dots R4$ – are the procedural values of output variables.

Given the nature of the tasks that the operator decides, it is obvious that this is the most effective organization of its activities, which provides analysis of the maximum number of the most important of IED. Efficiency of the operator of ICS will be as follows:

$$W(t_a) = \sum_{i=1}^n C_i P_i(t_i^{prc} < T^{avl}), \quad (8)$$

where t_a – the analysis time; C_i – the importance of the information contained in IED, $\sum_{i=1}^n C_i = 1$; $P_i(t_i^{prc} < T^{avl})$ – probability of correct processing (maintenance) IED for the time that is not exceed of allowable (T^{avl}); n – number of species IED that analyzed by the operator during operation time of t_a [1].

The proposed performance indicator actually describes weighed probability of application maintenance IED. For a queuing system, where there is an operator position [1], the probability of maintenance is one of the most important performance indicators of IRS. It confirms correctness of the

chosen performance indicator, as actually describes the weighed probability of timely analysis of all types of IED per a fixed time, in view of their importance and the average analysis time. Therefore, to take into account the importance of IED, it is necessary to find an efficient algorithm that would provide a choice from stream of VED in which there are both the greatest importance and lowest average analysis time as well as providing of the solution the operator in the form of recommendations for sequence their processing. It is possible to implement in the DSS, which automates the organization of the operator activity by means of additional information and linguistic support IRS.

The block diagram of this IPS DSS has the form shown in Fig. 2. The main elements of the system are the decision-making block (DMB) and linguistic support block (LSB). The decision-making unit evaluates the importance of IED and determined the average time of IED analysis of this type. In the operator control unit in accordance with the scheduling algorithm produced a solutions that are displayed to the operator in the form of recommendations. Block of linguistic support (LS) consists of a linguistic database that retains the details of IED dictionaries and linguistic processor that is used for forming of abstracts of IED [1].

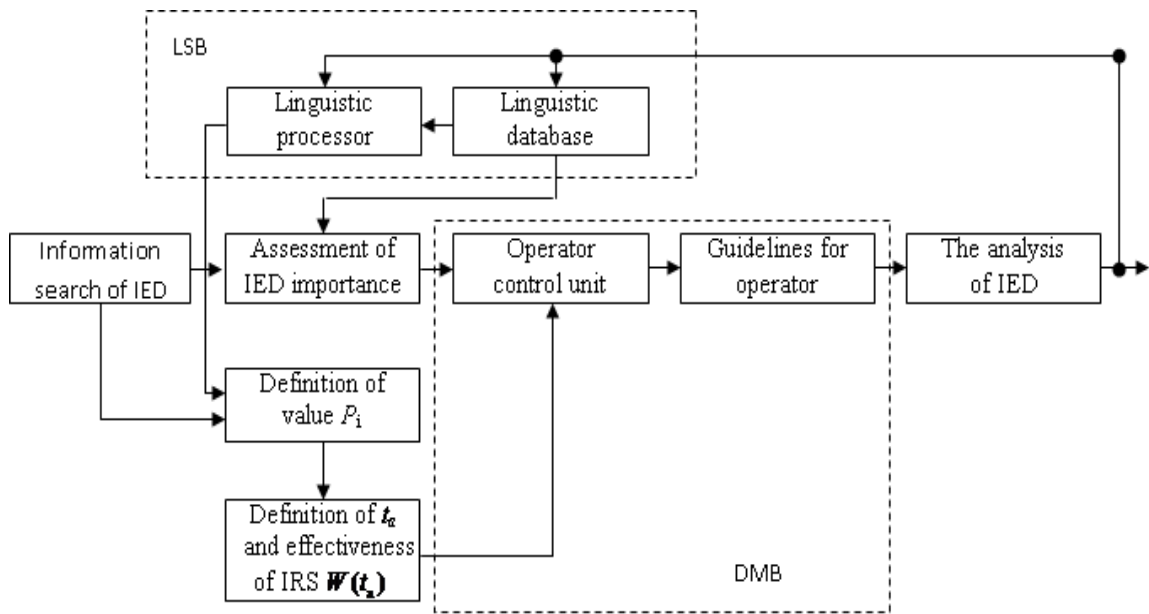


Fig. 2. Structural diagram of DSS in IRS.

How show the results of experimental studies, the application of the proposed apparatus in decision making support system enhances the efficiency of the operator to determine the sequence of processing IED in IRS 10–15%.

The described above approach can be actively used in other applications, such as inductive technologies of system information-analytical research (IT SIAR), particularly, in blocks of information-analytical support at execution of integrated innovative projects [8]. Search and processing of relevant information in such tasks could be more effective under applying of balance information bases criterion $CR_{inf}(I_b)$ [8] when using at least two analytical groups.

This criterion is responsible for the informative component in IT SIAR and is designed to select on s -step of inductive procedures of system research only such information I_b^{s+} , requests for which should have resemblance context of both analytical groups, closing to the current information base to optimal I_b^* . The final optimal information basis I_b^* of inductive SIAR-technology is an information basis for the creation of the final document of innovative research. Briefly remind the principle of action of balance information bases criterion $CR_{inf}(I_b)$.

At each step of the study, it compares the information requests (in practice it can be a numbered list of questions analysts in any form, questionnaire) from each groups A and B. For a single step of research:

$$CR_{inf}(I_b) = \sum_{k=1}^K d_{sk}, \quad s=1, \dots, S \quad (9)$$

$$\text{where: } d_{sk} = \begin{cases} 0, & \text{if } I_b^{sk(A)} = I_b^{sk(B)} \\ 1, & \text{if } I_b^{sk(A)} \neq I_b^{sk(B)} \end{cases} ,$$

K – is the number of allowable questions in requests on s -step of research and, thus, on s -th step the information support group offers for innovative project analytical group a some information "portion" I_b^{s+} , which minimizes the criterion (9):

$$I_b^{s+} = \arg \min_{I_b^s \in \mathcal{J}_b} CR_{inf} \{ I_b^{sA}, I_b^{sB} \}, \quad (10)$$

where \mathcal{J}_b – all information requests from both groups in the s -th step. For the entire analytical cycle of the project:

$$CR_{inf}(I_b) = \sum_{s=1}^S \sum_{k=1}^K d_{sk} \rightarrow \min . \quad (11)$$

Actually, selection of relevant pieces of information and could be under the responsibility of decision support systems, discussed above, that is, such system could helps to produce the relevant pieces of information in research under IT SIAR.

Conclusions

Considered the technical and methodological principles of gaining knowledge with the architecture using neural networks is advisable to apply when determining the sequence of processing documents with regard to the criteria of importance. The practical value of the proposed approach has been demonstrated. The results can be used at construction of information retrieval structures based on the calculated intelligence systems. The proposed system can be actively also used in information support blocks of inductive technology of system information-analytical research.

1. *Intelligent decision support systems: theory, synthesis, performance* / V.A. Tarasov, B.M. Gerasimov, I.A. Levin, V.A. Korniiichuk. – K.: MAKNS, 2007. – 255 p. [In Ukraine]. 2. *Gerasimov B.M. Fuzzy sets in problems of design, management and information processing* // B.M. Gerasimov, G. Grabowski, N.A. Ryumshyn. – K.: Technica, 2002. – 140 p. [In Russian]. 3. *Gerasimov B.M. Decision support system in automated of real-time control system* / B.M. Gerasimov, V.M. Glutsky, A.A. Rabchun // *Artificial Intelligence*. – № 3. – 2000. – P. 39-47. [In Russian]. 4. *The model of linguistic database in systems of automatic processing of natural language text data* / Zamaruyeva IV, VB Tolubko etc. // *Computer and math. methods in modeling*. – 2013. – № 1. – P. 75-81. [In Ukraine]. 5. *Shvorov A.S. Parametric method of information processing in the information-analytical systems* / A.S. Shvorov // *Proceedings of the Military Institute of National Taras Shevchenko University*. – 2013. – Vol. 43. – P. 128-133. [In Ukraine]. 6. *Rutkovska D. Neural networks, genetic algorithms and fuzzy systems* / D. Rutkovska, Pilinsky M., L. Rutkowski. – M.: Hotline – Telecom, 2004. – 452 p. [In Russian]. 7. *Melin P., Urias J., Solano D., Soto M., Lopez M., Castillo O., Voice Recognition with Neural Networks, Type-2 Fuzzy Logic and Genetic Algorithms*. *Engineering Letters*, 13:2, 2006. 8. *Osypenko V.V. System of criteria in inductive procedures of system-information-analytical researches* / V.V. Osypenko // *System technologies*. – No. 6 (71). – Dnipropetrovs'k. – 2011. – Pp. 106-113. [In Ukraine].

MARKING SPEECH SIGNAL BASED ON FACTOR HOLDER SMOOTHNESS AND FAST FOURIER TRANSFORM

© Pelekh Y., Peleshko D., 2014

За допомогою використання показника гладкості Гельдера на спектрі мовного сигналу запропоновано покращений метод маркування мовних сигналів. Особливістю методу є висока точність декодування водяного знаку та інваріантність від оригінальних даних, сигналу та водяного знаку.

Ключові слова: мовний сигнал, цифрови водяний знак, показник Гельдера, квазістаціонарна ділянка, гладкість функції, спектр сигналу, швидке перетворення Фур'є.

A new improved method of speech signals marking is presented. The method is based on employing of Hölder condition to speech signal spectrum. The key features of this method are its high accuracy during watermark decoding and independence from original data like speech signal and watermark.

Keywords: speech signal, digital watermark, Hölder condition, quasistationary, smooth function, signal spectrum, fast Fourier transform.

Introduction and analysis of the literature

Access to information and its rapid spread is a major factor in the development of new technologies in the past decade. Audio information has become more widely used over the Internet. With the increasing use of audio content on the Internet has serious problems, such as forgery, fraud and piracy. Virtually any user of modern electronic devices (gadgets) can copy the audio file and use it for their own purposes, such as use in presentations or for marketing campaigns, etc.

Thus, abuse of copyright is widespread among users of multimedia content. This, in turn, was the motivating factor in the development of new technologies protect audio information. Therefore the problem of piracy prevention information, counterfeiting and copyright infringement is urgent and requires the development of new and effective methods.

Digital markeruvannya speech signals has several options for application [9]. The most common are withholding information, copy protection and identification of personal data.

One of the technologies to protect speech signals is the use of digital watermarking. Digital watermarking methods are considered to be an effective solution to the problems of copyright. In order to confirm the copyright of the digital content, watermarking methods embed the secret information about the holder, using the masking effect of Human Visual System or Human Audio System.

Digital watermarks have proven to be one of the best ways to protect intellectual property from illegal copying. A digital watermark is a signal that is embedded in the digital data (audio, image, video and text) and to verify the identity of the data to be extracted only the specialized.

Given the relatively recent spread of technology, with the object of using watermarks can be divided into two classes: the watermark image and sound signals. The main object of research is in the speech signal. In this respect the method of implementation of algorithms technology in audio watermarking techniques can be divided into two groups: methods based on processing time series (rows) and methods based on execution sequences of transformations.

Usually the basic principle markeruvannya digital watermark is a hidden need replacement components watermarked audio signal. With technology Perceptual Audio Coding (MPEG Layer-On (IGC), MPEG-2 and MPEG-4) it is known that even slight changes in the speech signal may lead to distortion [8]. Therefore, camouflage replacement should be characterized by stability, transparency and reliability. Here, transparency refers to a perception of the speech signal, which is characterized by the fact that users can not feel the difference with the audio signal from the original watermarked signal.

Transparency perception is one of the major requirements for watermarking technology. In the process of embedding prohibited the introduction of significant tangible artifacts. That watermark should not affect the quality perception of the original audio signal. On the other hand, to ensure the stability and reliability, the power of the watermark is to be maximum which is in some contradiction to transparency. It defines the problem of finding some balance between transparency on the one hand, and stability and reliability of the other.

It should be noted that the stability of the watermark determines the protection against unauthorized access and means that the watermark can be read only by an authorized person. With this condition the technology watermark imposed another condition, which is the need to ensure sufficient capacity is integrated in such a number of the alarm bits that would ensure sustainable security.

This paper proposes a new watermarking scheme based on fast Fourier transform (FFT) for markeruvannya speech signals. The watermark is embedded in a given form of peaks with a magnitude spectrum of each frame is overlapping audio.

Once it should be noted that the proposed labeling scheme of voice signals ensures the stability and reliability of several types of attacks, such as noise, cropping, oversampling, re-quantization, compression and lowpass filtering.

Statement of the problem

The aim is to improve the method of “blind” labeling speech signal [1], which is based on calculating the Hölder exponent for quasi-stationary areas. The method of constructing quasi areas based singular decomposition based on digital signals [2, 4, 5].

1. Presentation signal

By its nature, the speech signals are continuous. However, in this paper we consider only discrete signals. Below is a representation of the signal and its division into quasi-stationary areas that do not intersect.

Suppose we have a signal $x(t)$, де $t = 0, 1, 2, \dots, N$ – splitting functions. After constructing quasi areas [4] we obtain:

$$x(t) \leftrightarrow \bigcup_{i=1}^m Y_i, . \quad (1)$$

where Y_i – quasi-stationary area serving associations successive elementary areas x_i .

$$Y_i \leftrightarrow \bigcup_{j=l_i}^{m_i+l_i-1} x_j . \quad (2)$$

From the formula (2), m – number of quasi areas Y_i , l_i – initial quasi-stationary area code, m_i – number of elementary areas x_i in association (1). Thus, it follows that $m \leq t$.

To solve problem (2) the method proposed in [2]. The application of this method we obtain a new representation of the speech signal as a series of quasi-stationary areas Y_i , which do not intersect.

After initial segmentation labeling problem is to find the corresponding signal samples for encoding / decoding a watermark on the site. It is necessary to find the most “favorable” signal sections which receive minimal distortion in the reconstruction. Search areas offered himself done through calculating Holder, which in its physical sense would indicate a maximum amplitude at intervals (kvazistatsionarah). To improve the stability of the method to noise, filtering and other undesirable distortion is proposed to calculate the Hölder exponent spectrum of the input signal, which is built using fast Fourier transform.

2. Algorithm and embedding scheme

2.1. Calculating Holder exponent

Suppose that the speech signal $x(t)$ in the interval Y_i is a smooth manner k . Then, by [6], the function $x(t)$ at the point of l , $l \in Y_i$ satisfies the Holder exponent α , $0 \leq \alpha \leq 1$, if there is a constant $A \in \mathbf{R}$ and for all points t with neighborhood l performed condition

$$|x(t) - x(l)| \leq A|t - l|^\alpha. \quad (3)$$

According to [6] it is known that the function $x(t)$ at the point of l , $l \in Y_i$ satisfies the Hölder exponent α , $\alpha \geq 1$, if and only if there exists a constant $A \in \mathbf{R}$ and polynomial $p_l(t)$ degree $m = \lfloor \alpha \rfloor$ such that for all points t with neighborhood l performed condition

$$|x(t) - p_l(t)| \leq A|t - l|^\alpha. \quad (4)$$

According to [6] function $x(t)$ in $[a, b]$ satisfies the Hölder exponent α , if $x(t)$ satisfies the Hölder exponent α for all $l \in Y_i$, with a constant $A \in \mathbf{R}$, which is independent of l .

At each interval Y_i polynomial $p_l(t)$ determined only way. If $x(t)$ is continuously differentiable $k = \lfloor \alpha \rfloor$ once a interval Y_i , then $p_l(t)$ is Taylor scheduled $x(t)$ at the point of l . If $x(t)$ satisfies a uniform Hölder exponent condition $\alpha > m$ in the interval Y_i , then, as shown in [6], the function $x(t)$ necessarily k -times continuously differentiable on this interval. Limited, but bursting at the point l function satisfies Hölder exponent 0 to l . If smoothness is Hölder exponent is $\alpha < 1$ at the point of l , then $x(t)$ not differentiated in l , and α determines the type of rupture [6]. Uniform Hölder exponent smooth functions $x(t)$ in Y_i associated with asymptotic decline in its Fourier transform [6].

2.2. Discrete Fourier transform and discrete spectrum signals

It is known that any complex periodic signal can be fed via a Fourier series as a sum of simple harmonic oscillation. The combination of simple harmonic oscillation, which can be decomposed complex periodic signal is called its spectrum.

The distribution of the amplitudes of harmonics in frequency is called the amplitude-frequency spectrum or amplitude spectrum, and the distribution of their initial phases of the frequency - phase-frequency spectrum and phase spectrum.

Lines discrete spectrum with the dimension of the signal amplitude. Continuous spectrum indicates the distribution of the amplitudes across the spectrum and has the dimension of density signal amplitudes.

Fourier transforms - integral transforms one complex-significant functions of a real variable to another, closely related to the Laplace transform and equivalent decomposition in Fourier series for nonperiodic functions. Under this transformation function can be represented in the form of fluctuating additive operator functions.

The Fourier transform of the speech signal functions $x(t)$ mathematically defined as a complex function $F(\omega)$, given by the integral

$$F(\omega) = \int_{-\infty}^{\infty} x(t)e^{-i\omega t} dt. \quad (5)$$

The inverse Fourier transform given by

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{-i\omega t} d\omega = f(t). \quad (6)$$

It is known that periodic signals $x(t) = x(t + mT_0)$ represented by Fourier series as a sum of harmonic components, as in this case, the analyzed signal is only the fundamental harmonic $\omega_0 = 2\pi/T_0$ and its multiple components $k - \omega_0$. In the exponential form of Fourier series has the form

$$x(t) = \dots + \xi_{-2}e^{-j2\omega_0 t} + \xi_{-1}e^{-j\omega_0 t} + \xi_0 + \xi_1e^{j\omega_0 t} + \xi_2e^{j2\omega_0 t} + \dots = \sum_{k=-\infty}^{\infty} \xi_k e^{jk\omega_0 t}. \quad (7)$$

Coefficients of the series ξ_k are complex variables and are determined from the ratio

$$\xi_k = C_k e^{j\psi_k} = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)e^{jk\omega_0 t} dt = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \cos(k\omega_0 t) dt - j \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \sin(k\omega_0 t) dt. \quad (8)$$

The set of coefficients of the spectrum is a periodic signal. The range of amplitudes C_k and phase spectrum ψ_k uniquely determine the signal $x(t)$ determine the part each frequency harmonic components in

the composition of the resulting vibrations. Often, only limited consideration C_k which determines the energy properties of the signal, and ψ_k relates only to the waveform.

Discrete signals $x(n)$ formed by sampling (taking samples) continuous signal $x(t)$ in time- interval T_s . By theorem of Kotelnikov-Shannon sampling frequency $F_s = 1/ T_s$ must exceed at least twice the maximum frequency in the spectrum of sampled signal. For discrete signals pair of integral Fourier transforms (5) and (6) takes the form:

$$\chi(\Omega) = \sum_{n=-\infty}^{+\infty} x(n) - e^{-j\Omega}; \quad x(n) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \chi(\Omega) - e^{j\Omega} d\Omega. \quad (9)$$

where $\Omega = \omega - T_s = 2\pi \frac{f}{F_s}$ - sampling rate normalized angular frequency.

Given the periodic spectrum $\chi(\Omega)$ discrete signal of period F_s (for normalized circular frequency 2π), in expression (10) limits of integration narrowed to one period, including zero frequency.

In practice, the spectral analysis of real signals is considered continuous signal $x(t)$, and a finite sequence of values of samples $x(n)$ in a given format. However, the use of the expression (9) to calculate the spectral function $\chi(\Omega)$ impossible through endless boundary value summation. That is why the technique of digital signal processing for spectral analysis used a discrete Fourier transform.

Discrete Fourier transform (DFT) is a pair of mutually single-valued transformation that connects the sampling discrete periodic signal $x(n)$ with complex coefficients $\chi(k)$ its discrete spectrum:

$$\chi(k) = \sum_{n=0}^{N-1} x(n) - e^{-j\frac{2\pi nk}{N}}, \quad k = 0, \dots, N-1, \quad x(n) = \frac{1}{N} \sum_{k=0}^{N-1} \chi(k) - e^{j\frac{2\pi nk}{N}}, \quad n = 0, \dots, N-1, \quad (10)$$

where N - number of samples in the period of the sampled signal.

The first transformation (10) is straightforward, and the second conversion - feedback DFT (FDPF). DFT is performed on finite periodic sequence of signal samples $x(n)$, in which the period consisting of N discrete values, $x(n) = x(n + N)$. Sequence $\chi(k)$ represents the spectrum of the sampled signal $x(n)$, which is also a periodic $\chi(k) = \chi(k + N)$.

Comparison of expressions (9) and (10) shows that the DFT represents a discrete sample $\chi(k)$ continuous spectral function $\chi(\Omega)$ discrete signal corresponding frequencies $\omega_k = 2\pi \frac{k}{NT_s}$:

$$\chi(k) = \chi(\omega_k) = \chi\left(2\pi \frac{k}{NT_s}\right). \quad (13)$$

Thus, the value of discrete frequencies ψ_k in the spectrum depends on the sampling period T_s , identification and resolution in frequency (spectral selectivity) is inversely proportional to the time the signal surveillance

$$T_0 = NT_s. \quad (14)$$

From relation (13) follows an important conclusion: if the current number of samples N to add a number of zeros, such as N , then the spectral function of a discrete signal changes, and DFT will double the number of spectral samples that correspond to frequencies, closely located in the range from zero to the sampling frequency.

3. Results and conclusions

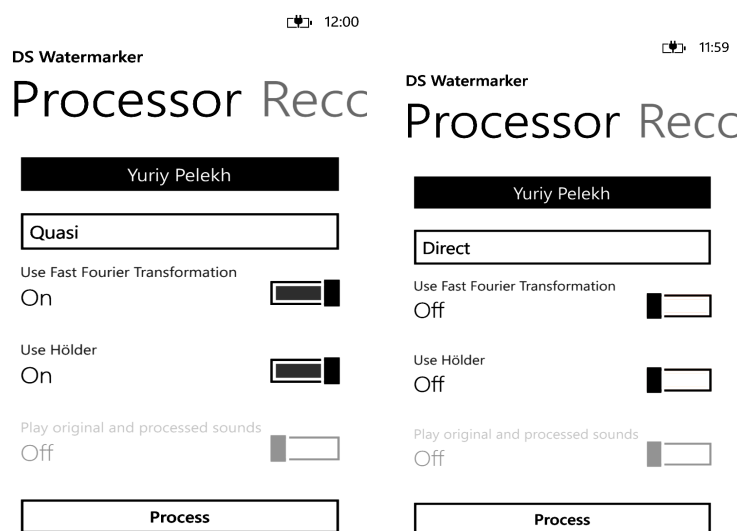
The results of this theoretical study have been verified experimentally. On the basis of the proposed method is implemented algorithm markeruvannya speech signal, which consists of the following steps:

- First, the speech signal is divided into uniform serial sections;
- Further, by combining, according to [2] built kvazistatsionary (primary segmentation of the speech signal);

- Using Fourier transform [7] obtained the spectra of quasi-stationary areas;
- Each section is calculated quasi Holder exponent;
- Using Holder exponent determined by counting signal which is encoded / decoded [5] watermark.

Software implementation of the algorithm was tested on the example label mono recordings words “test” and “mouse”. Characteristics of speech signals are as follows: length of signal 954 milliseconds and 1.09 seconds respectively. Sampling frequency 16 kHz. The length of the elementary section 512 samples. For the watermark is taken into “Yuriy Pelekh”. The test was carried out ten times for five different entries for each of the test words.

Note that this software implementation (Pict. 1) allows the user to choose the parameters encoding / decoding watermarks. For example, to build FFT spectrum can be used or not used. Then search for labeling samples will not run on spektri signal and on the signal. You can also split the signal at kvazistatsionary or elementarni uniform areas. Alternatively, the Hölder exponent can be calculated magnitude between two adjacent reference.



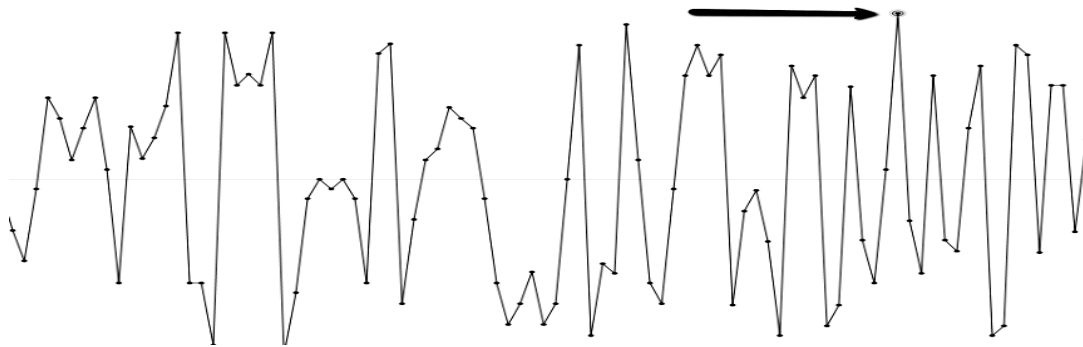
Pict. 1. Software implementation (choice of parameters labeling)

The table shows the results (presented in percentages) of the method of using the Hölder exponent and without. The first number represents the percentage of resemblance encoded and decoded watermark, and the second - the percentage of coincidence points of the speech signal, which happened encoding / decoding watermark.

Results of encoding / decoding a watermark in a digital signal

	Direct decomposition	Breakdown by quasi-stationary		Test word	
	Without FFT	With FFT	Without FFT		With FFT
Using Holder exponent	76 / 79	93 / 100	84 / 87	94 / 99	Test
	74 / 52	93 / 100	61 / 46	89 / 100	Mouse
Without the Holder exponent (calculation of magnitude between adjacent samples)	85 / 87	95 / 100	89 / 92	97 / 100	Test
	80 / 80	93 / 100	79 / 79	96 / 100	Mouse

In Pict. 2 shows the results point at which will be labeled signal with the help of Hölder exponent.



Pict. 2. Calculation smooth signal with Holder exponent

On the basis of the experiments proved that the improved method of marking digital voice signals using Hölder exponent and splitting the signal at quasi-stationary area, improves search on the maximum amplitude of the spectrum, which gives a more accurate result in the encoding and decoding of the watermark in the speech signal.

1. Peleshko D., Pelekh Y., Izonin I., Peleshko M. *Digital Watermarking of Speech Signals*. CSIT'2012, p.108, November 2012.
2. Пелешко Д.Д. Виділення псевдоінваріантів та квазістаціонарних ділянок мовних сигналів на основі сингулярних розкладів/ Д.Д.Пелешко, А.М.Ковальчук, Ю.М. Пелех, В.І.Киричук// Вісник НУ "Львівська політехніка" Комп'ютерні науки та інформаційні технології. – №732. – 2012. – С.58-66.
2. Пелешко Д. Д., Пелех Ю. М. Модифікація методу маркування мовних сигналів на основі моментів Зерніке. ISDMCI'2012, ст.392-393, травень 2012.
3. Peleshko D., Pelekh Y., Izonin I. *Digital Watermarking of Speech Signals Based on Quasistationary Areas*. CADSM'2013, p.283, February 2013.
4. Peleshko D.D., Pelekh Y.M., Izonin I.V., *Constructing Of Pseudoinvariants And Digital Watermarking Of Speech Signals Based On A Singular Value Decomposition*, Journal of Global Research in Computer Science. Volume 4, No. 2, February 2013. pp. 56-59.
5. Ручай А.Н. Текстозависимая верификация диктора: модель, статистические исследования, комплекс программ. Автореферат диссертация на соискание ученой степени к.ф.-м.н. Челябинский государственный университет, 2012.
6. Білінський Й.Й., Огородник К.В., Юкиши М.Й. Електронні системи. [Електронний ресурс]. Режим доступу: http://posibnyku.vntu.edu.ua/e_s/3.htm.
7. Хуан Пенцзюнь, Резнік Ю. Масштабоване кодування мови та аудіо з використанням комбінаторного кодування МДСТ-спектра. Патент України. №95185. Опубліковано: 11.07.2011. [Електронний ресурс]. Режим доступу: <http://uapatents.com/26-95185-masshtabovane-koduvannya-movi-ta-audio-z-vikoristannyam-kombinatornogo-koduvannya-mdct-spektra.html#formula>.
8. Стасюк В. І., Тимощук А. О., Старіков В. Д., Зайченко М. М. Спосіб захисту від несанкціонованого прослуховування та маскувальник мови для його здійснення Патент України. №: 44665. Опубліковано: 15.02.2002. [Електронний ресурс]. Режим доступу: <http://uapatents.com/3-44665-sposib-zakhistu-vid-nesankcionovanogo-proslukhovuvannya-ta-maskovalnik-movi-dlya-jjogo-zdijjsnennya.html>.

METHODOLOGICAL PRINCIPLES OF ELECTRONIC INFORMATION RESOURCES OF BRANCH

Ó Batyuk A., Viter M., Sendzyuk M., 2014

Здійснені аналіз і систематизація сучасної української філії електронних інформаційних ресурсів. Описані методологічні принципи їх організації та функціонування з точки зору їх інтеграції.

Ключові слова: Електронні інформаційні ресурси, класифікація, інтеграція, взаємодія, електронний уряд.

The analysis and systematization of modern Ukrainian branch electronic information resources are carried out. The methodological principles of their organization, functioning and terms of their integration are described.

Keywords: electronic information resources, classification, integration, interaction, e-government.

Топікаліті

According to regulatory documents [1], the national system of electronic information resources is a set of both national resources needed to solve the problems of socio-economic development and electronic information resources of the National Registry which are arranged with the help of the single technology. In the National Registry, it is also indicated which electronic information resources can be treated as national, state, public, and private ones.

However, as it follows from investigation, it is expedient to separate the branch electronic information resources (BEIR) as a particular type. The BEIR consists of information resources which are related to a particular branch, and are formed and administered by means of automated information systems of the public authority of a corresponding branch. As an example, we can mention the following BEIR: Unified State Register of legal entities and natural persons (entrepreneurs, AIS, "State Budget", AS "Treasury", State Register of Encumbrances of Movable Property, etc).

Each of these systems was formed separately, mainly according to the branch concept:

it functions in its subject area;

it is based on its programme platform;

it has its own databases, a system of classifications and reference books, which are formed particularly;

it ensures a particular customer's interface.

The importance of a separation of the consideration of the branch electronic information resources is related to the need of their integration into a single system of e-governance of Ukraine.

For effective organization of such electronic information coordination, it is necessary to coordinate interaction of proper software, hardware, as well as regulatory and organizational documents.

A thorough study of the branch electronic information resources can help to create a mechanism of their integration more effectively, to account the specific characteristics and functioning of each of them more accurately.

Overview of references

References review shows that in this area the research is usually conducted either in the direction of solving the general methodological problems of the formation and functioning of information resources of public authorities [2,3], or towards the analysis of the functioning of individual branch information and analytic systems [4.5].

An attempt to classify the information and analytic systems with the use of growing pyramidal system in terms of logical-linguistic information models is done in the work [6].

However, the task of detailed analysis and systematization of the branch electronic information resources as the basic structures of the e-government system of the state remains topical.

Objective

The purpose of this paper is to analyze the normative and methodological principles of the branch information resources formation as well as their classification based on their structural and functional features.

In connection with the specific of information that functions in automated information systems of Ukrainian banks, the electronic information resources of the banking sector in this study are not considered.

The main material

1. The history of formation

The process of formation of national branch electronic information resources started in the early years of the formation of the independent Ukrainian state, but not always it was systematic.

The tasks of construction of automated information systems in various areas of governance were formulated within the National informatization program's frame, but a lot of them have not been performed yet [7].

Among electronic information resources that were quite completely assembled in the course of accomplishment of National informatization program, we should name the automated information systems of the State Tax Administration, Ministry of Finance, State Treasury, State Customs Service, etc.

2. Requirements to the branch information resources

Despite the obvious functional differences between different BEIR, they are characterized by a unity of normative and methodological principles of organization and functioning. In particular, the process of creation and functioning of the BEIR is strictly regulated by the relevant regulatory acts.

Resolution of the Cabinet of Ministers of Ukraine on 04.02.1998 № 121 "On approval of the list of compulsory stages of the work in the design, implementation, and service of systems and means of automated data processing and data transfer" determines the list of required stages of the work in the design, implementation, and service of systems and the means of automated data processing and data transfer. There it is also said that the budgetary financing the design, implementation and operation of the information means for the purposes of the executive authorities activities informatization is carried out only under the condition of performance of the compulsory stages of the work in accordance with the state standards, and under the presence of a positive expertise of these stages' content, which is provided by the Science, Innovation and Information State Agency.

The processes of the supply, development and technical maintenance of software that is purchased or developed on the request of public authorities, are statutory stated [8].

Conditions of the software functioning in the executive authorities are also regulated by a respective Decree of the Cabinet of Ministers of Ukraine [9]. In particular, there it is said to purchase only licensed copies of the software or copies of freeware, which have to be provided with documentation confirming the legality of their use under license or their belonging to freeware.

In the case of absence of licenses, the procedure of software legalization in the executive authorities is used. It involves the bringing of the software use in accordance with the requirements of copyright law by replacing unlicensed copies of software with the licensed.

The interaction of public authorities in the sphere of information resources protection in the information and telecommunications systems are regulated by the Decree of the Cabinet of Ministers of Ukraine [10] and other documents.

Organization and functioning of the internal electronic information resources of public authorities is also governed by corresponding regulatory documents [11, 12].

3. BEIR classification

Among the large number of the branch electronic information resources being considered, it is possible to single out certain groups which are built on the same structural and functional bases, namely, accounting, and functional, special, and internal ones.

Accounting BEIR

Automated information systems whose main task is to systematize and account the relevant data are referred to the accounting BEIR. These are various registers, inventories, databases.

As examples of accounting BEIR, we can cite the following:

1) The register of information, telecommunication, and information and telecommunication systems of executive authorities, and enterprises, institutions and organizations, which are under their control. This information system is dedicated to accumulation, accounting, processing and storage of data on the composition, structure, location, functioning conditions, purpose, the state of information security in information, telecommunication, and information and telecommunication systems containing government electronic information resources or being planned for this, executive authorities as well as enterprises, institutions and organizations being under their control. Registry Administrator is an authorized structural subdivision of the State Special Communication Administration of the Administration of Civil Service for Special Communication and Information Protection of Ukraine.

2) Unified State Register of Legal Entities and Natural Persons (entrepreneurs) is an automated system of collection, accumulation, protection, accounting and providing with the information on legal entities and natural persons (entrepreneurs). The administrator of the system is the Ministry of Justice of Ukraine.

3) Unified Register of State Property is an automated system for collecting, accounting, storage, processing, protection and providing with the information on real estate, including leased one, leasing, concession or pledge of state enterprises and their associations, institutions and organizations, and also the state corporate rights and the state property not included in the share capital of business companies established during the privatization and corporatization. Administrator is the State Property Fund.

4) Unified State Register of Legal Acts is an automated system of collection, storage and processing of legislations. The administrator of the system is the Ministry of Justice of Ukraine

5) State Register of Encumbrances Over Movable Property is a unique computer database of the origin, change, termination of encumbrances as well as the repossession encumbrance. The administrator of the system is the Ministry of Justice of Ukraine .

6) State Land Cadastre is a unique state geographic information system of data on land located within the state border of Ukraine, its intended use, restrictions in its use, as well as the quantitative and qualitative characteristics of land, its estimation, the distribution of land between owners and users. The administrator of the system is the State Agency of Land Resources.

Functional BEIR

Systems providing automated implementation of the basic functions of the public authority, where they function, we refer to functional BEIR.

They include such automated information systems as “Treasury” of the State Treasury, “State Budget” of the Ministry of Finance, “Tax Unit” of the Revenue Service and others.

Their common feature is that the information that they have handled, is not static and the structure of the systems themselves is related to the functional structure of the bodies to which they belong.

Automated system of treasury budget execution AS “**Treasury**” is composed of subsystems “Treasury-Revenues” and “Treasury-Expenditures” [13].

Subsystem “Treasury-Revenues” with a special tool “Differentiation and accumulation” and with transactions that include all types of accounting transactions performs the following:

- distribute all revenues from taxpayers into two streams: government revenues and local revenues in terms of funds (general and special);
- allocate costs according to legislation between different budgets, budget funds, levels of budgets and recipients with the help of specific system modules;
- on time, every day, to transfer all revenues for their intended purpose;
- receive timely, prompt and reliable reporting on budget execution at all levels.

Implementation mechanism of treasury performance of budgets of all levels in terms of expenditure is realized by means of the automated system “Treasury-Expenditures”. It is designed for execution

accounting of state budget expenditures in the bodies of the State Treasury and based on its data compiling the financial and management reporting in accordance with applicable regulations and laws of Ukraine.

AS “Treasury-Expenditures” is based on Web-technologies and is provided with adequate protection against unauthorized access.

AIS “State Budget” of the Ministry of Finance of Ukraine is the base operating information system that provides the information and analytical support for all major stages of the budget process. [13]

This system consists of two software systems – software complex of the central level, and software for main administrators of funds.

Both software systems contain subsystems of exchange that provide the data exchange between the Ministry of Finance, State Treasury and main administrators of funds.

AIS “State Budget” provides drafting the state budget for the planned year, compiling a temporary list and the state budget list, analysis of budget execution, formation of state budget financing execution reporting.

The system of “Tax Unit” of the State Tax Service is designed to perform the tasks assigned to the State Tax Service Authorities [14]:

- informing taxpayers about their tax obligations and clarifying tax laws;
- tax accounting and registration of taxpayers;
- acceptance and processing of tax reporting and accounting fees;
- check of tax reports for compliance with requirements and standards established by the laws and regulatory acts of Ukraine;
- application of measures for repayment of tax debt in the cases of the violation of tax compliance by taxpayers;
- resolve disputes on the amount of tax additional charges and tax debt and like.

Special BEIR

We suggest the information resources that are designed to automate the specific functions of the state and which are neither economic nor business to refer to special BEIR.

A typical representative of such systems is the Unified Information Analytic System (UIAS) “Election”. This is a set of interrelated regulatory legal, organizational and administrative measures, and program-technical, telecommunication means to support the collection, processing, accumulation, analysis, storage and disclosure of information on elections and referendums. The main purpose of UIAS “Elections” is the computerization of processes undertaken in the course of the preparation and holding of elections and referendums in Ukraine, as well as the Central Election Commission activity, election commissions of presidential elections in Ukraine, people's deputies of Ukraine, deputies of local councils and village and settlement heads, mayors, commissions on national and local referendums [15].

UIAS “Elections” has a multilevel structure that corresponds to the organizational structure of the preparation and holding of elections and referendums in Ukraine, namely:

- a complex of UIAS “Elections” of the Central Election Commission level;
- a complex of UIAS “Elections” of election commissions and commissions on referendum of the lower level;
- corporate network of UIAS “Elections”.

At the level of the corresponding election commissions and commissions on referendum, UIAS “Elections” is a set of logically complete, functionally separate, but conceptually and information related subsystems.

It is possible to consider Central State Electronic Archive of Ukraine as special BEIR. It performs the tasks and functions of the state management of Archives and records management, provides accounting, storage of electronic documents of the National Archival Fund and electronic information resources and the use of their information [16].

Internal BEIR

Each central public authority has internal software complexes that provide automation of the corresponding internal processes of the corresponding authority. They include the electronic document management system, automated information systems of personnel services, finance accounting, internal and external websites, etc.

Conclusions

On the basis of done analysis, it is possible to draw the conclusion about the functional diversity and multi-sectoral aspect of the branch electronic information resources.

With this, for the effective functioning of the considered BEIR, it is necessary to solve the following two problems.

On the one hand, every branch resource within its functional purpose should completely *satisfy the needs of users*. For this purpose, the responsible central authorities form appropriate interfaces of interaction with users (sites, portals, system of electronic exchange, e-receptions, etc).

On the other hand, BEIR should be most effectively integrated into the state system of e-government. For this purpose, it is necessary to develop appropriate standards and technologies of such integration. One of the solutions to this problem can be a technology of UDDI with XML standard of data exchange [16].

The preliminary classification of BEIR, which is suggested in this article, allows us to develop the typical integration solutions for each type of electronic information resources.

1. Розпорядження Кабінету Міністрів України від 05.05.2003 № 259-р “Про затвердження Концепції формування системи національних електронних інформаційних ресурсів”. – [Електронний ресурс]. – Режим доступу: <http://zakon1.rada.gov.ua/laws/>. 2. Додонов О.Г. Архітектура автоматизованих інформаційно-аналітичних систем органів державної влади / О.Г. Додонов, О.В. Нестеренко, М.М. Будько // *Мат. машини і системи*. – 2003. – № 3–4. – С. 138–146. 3. Нестеренко О.В. Основи побудови автоматизованих інформаційно-аналітичних систем органів державної влади: моногр. / О.В. Нестеренко; Ін-т пробл. реєстрації інформації НАН України. – К.: Наук. думка, 2005. – 627 с. 4. Яремко С.А. Застосування інформаційних систем і технологій в контрольно-ревізійному процесі як запорука його удосконалення та ефективного функціонування / С.А. Яремко, Ю.М. Данильчук, Ю.О. Мухомора // *Міжвузівський збірник “Комп’ютерно-інтегровані технології: освіта, наука, виробництво”*. – Луцьк, 2011. Випуск № 4. – С. 291–293. 5. Гора О.Б. Інформаційно-аналітичні системи контролю за виконанням державних рішень / О.Б. Гора // *Реєстрація, зберігання і оброб. даних*. – 2009. – Т. 11, № 4. – С. 43–49. 6. Нестеренко А.В. Методология классификации и структурирования автоматизированных информационно-аналитических систем органов государственной власти / А.В. Нестеренко // *Искусств. интеллект*. – 2005. – № 3. – С. 504–520. 7. Вітер М.Б. Національна програма інформатизації та розвиток інформаційних технологій в органах виконавчої влади України / М.Б. Вітер // *Проблеми інформатизації та управління: Збірник праць*. – 2013. – Випуск 2 (42). – С. 19–23. 8. Постанова Кабінету Міністрів України від 12.08. 2009 р. № 869 “Про затвердження загальних вимог до програмних продуктів, які закуповуються та створюються на замовлення державних органів”. – [Електронний ресурс]. – Режим доступу: <http://zakon1.rada.gov.ua/laws/>. 9. Постанова Кабінету Міністрів України від 10.09.2003 № 1433 “Про затвердження Порядку використання комп’ютерних програм в органах виконавчої влади”. – [Електронний ресурс]. – Режим доступу: <http://zakon1.rada.gov.ua/laws/>. 10. Постанова Кабінету Міністрів України від 16.11.2002 № 1772 “Про затвердження Порядку взаємодії органів виконавчої влади з питань захисту державних інформаційних ресурсів в інформаційних та телекомунікаційних системах”. 11. Наказ Державного комітету інформаційної політики, телебачення і радіомовлення України, Державного комітету зв’язку та інформатизації України від 25.11.2002 № 327/225 “Про затвердження Порядку інформаційного наповнення та технічного забезпечення Єдиного веб-порталу органів виконавчої влади та Порядку функціонування веб-сайтів органів виконавчої влади”. – [Електронний ресурс]. – Режим доступу: <http://zakon1.rada.gov.ua/laws/>. 12. Постанова Кабінету

Міністрів України від 28.10.2004 № 1453 “Про затвердження Типового порядку здійснення електронного документообігу в органах виконавчої влади”. – [Електронний ресурс]. – Режим доступу: <http://zakon1.rada.gov.ua/laws/>. 13. Сендзюк М.А. Інформаційні системи і технології в економіці: навч. посібн. / М.А. Сендзюк, М.Б. Вітер. – К.: КНЕУ, 2011. – 422 с. 14. Сайт Міністерства доходів і зборів України – [Електронний ресурс]. – Режим доступу: <http://minrd.gov.ua/>. 15. Постанова Центральної виборчої комісії від 21.03.2003 № 16 “Про затвердження Концепції Єдиної інформаційно-аналітичної системи “Вибори”. – [Електронний ресурс]. – Режим доступу: <http://zakon1.rada.gov.ua/laws/>. 16. Наказ Міністерства юстиції України від 21.05.2012 № 759/5 “Про затвердження Положення про Центральний державний електронний архів України”. – [Електронний ресурс]. – Режим доступу: <http://zakon1.rada.gov.ua/laws/>. 17. Данилин А.В. Среда електронного взаємодія для електронного правительствa // Информационное общество. – 2003.– № 1. – С. 41–52.

УДК 004:351

Н. Кунанець, О. Малиновський
Національний університет “Львівська політехніка”,
кафедра інформаційних систем та мереж

ЕЛЕКТРОННІ БІБЛІОТЕКИ: ОБСЛУГОВУВАННЯ ОСІБ З ОСОБЛИВИМИ ПОТРЕБАМИ

© Кунанець Н., Малиновський О., 2014

Розглянуто шляхи підвищення ефективності подання інформації за допомогою мультимедійних технологій, забезпечуючи тим самим використання нового потужного інструменту для сприйняття інформації особами з різними формами нозології. Проаналізовані особливості формування мультимедійного інформаційного контенту для користувачів з особливими потребами. В роботі розглянуто електронну бібліотеку як інформаційну систему, в якій формується та нагромаджується контент з різних джерел, сортується, структурується та інтелектуально опрацьовується. Запропоновано набір інформаційно-технологічних послуг, що робить мультимедійний інформаційний контент доступним для користувачів з різними формами нозології.

Ключові слова: мультимедіа, користувачі з особливими потребами, цифрова бібліотека, інформаційний контент.

The ways of efficiency improving of information presenting with using of multimedia technology were examined in this article, thus ensuring the use of a new powerful tool for perception of information by disabled persons. The formation peculiarities of multimedia information content for disabled persons were analyzed. This paper considers a digital library as an information system, where the information is formed and collected from different sources, also, it is sorted, structured and intellectually processed. It's proposed a set of information-technology services, which make multimedia information content accessible for disabled persons.

Keywords: multimedia, disabled persons, digital library, information content

Постановка проблеми

У сучасному інформаційному суспільстві для ефективного бібліотечного обслуговування користувачів з обмеженими можливостями необхідне переосмислення традиційних форм роботи з читачами. Все більшої актуальності набуває потреба розширення можливостей зручного подання інформації для осіб з особливими потребами в різних доступних для їхнього сприйняття форматах. Розробляються та вдосконалюються інформаційні технології, методи та засоби, що реалізують комплектування та формування фондів електронних бібліотек спеціалізованими інформаційними

продуктами, що інтегрують різнотипову інформацію для забезпечення інформаційних потреб осіб з різними формами нозології.

Саме електронні бібліотеки покликані стати провідними бібліотечними інститутами, що забезпечують інформаційну підтримку та супровід освітніх та соціокомунікаційних процесів для осіб із специфічними потребами. Термін “електронна бібліотека” означає новий вид інформаційних систем, призначених для нагромадження, структурування й організації масиву електронних документів з відповідною системою доступу до них. Електронна бібліотека забезпечує доступ до віддалених, розподілених і різнорідних ресурсів [1].

Удосконалення діяльності бібліотеки насамперед стосується впровадження інформаційних технологій. Мається на увазі потреба удосконалення традиційних технологічних процесів із застосуванням нових інформаційних технологій, серед яких впровадження мультимедійних технологій формування і використання електронних ресурсів та розвиток взаємодії бібліотек на засадах корпоративних домовленостей. Це приводить до створення нових структурних підрозділів, змін функціональних обов’язків працівників, залучення спеціалістів з різних галузей знань.

Аналіз останніх досліджень та публікацій

Проблеми створення інформаційного освітнього контенту аналізують А.В. Крапивенко [4], В.В. Дунаєв [5], Е.В. Давидова [6], які досліджують проблеми формування ефективних мультимедійних інформаційних продуктів. Аспектам програмно-технологічного забезпечення процесів створення електронних бібліотек присвячені роботи Ф. С. Воройського [11], А. І. Земскова [12], Я. Л. Шрайберга [13], Е. С. Кожевникової [14]. Функціональні, інформаційні характеристики електронних бібліотек розглядалися Т. А. Калюжною [15], А. Н. Єлізаровим [16]; питання оцінки якості електронних ресурсів присутні у роботах Н.А. Лапкіної [17].

Однак, комплексного дослідження особливостей формування інформаційного контенту для осіб з особливими потребами ще не проведено. Водночас результати аналізу джерел показують, що досі немає єдиного загальноприйнятого визначення електронних бібліотек. У різних джерелах з проблем створення електронних бібліотек домінують роботи технічного плану, присвячені програмному забезпеченню електронних бібліотек, їх архітектурі, управлінню ресурсами.

Формулювання цілі статті

Метою статті є дослідження методологічних та методичних засад формування якісного мультимедійного інформаційного контенту для осіб з особливими потребами в електронних бібліотеках.

Виклад основного матеріалу

I. Функціональні завдання електронних бібліотек для осіб з особливими потребами

Електронною бібліотекою (digital library) для осіб з особливими потребами вважатимемо таку форму інтегрованої бібліотечно-інформаційної системи, що дає змогу накопичувати, зберігати та надавати доступ користувачам до різноманітних колекцій електронних повнотекстових та мультимедійних документів, поданих у зручному для них форматі з врахуванням комунікаційних каналів сприйняття ними інформації.

Характерною сукупністю базових ознак такої електронної бібліотеки є її комп’ютерне обладнання, цифровий спосіб запису інформації, технології онлайн-доступу до інформаційних ресурсів та ін.[2].

У цьому контексті бібліотечна підтримка освітнього процесу осіб з особливими потребами полягає в такому:

- використання специфічних форм подання бібліотечно-бібліографічної інформації для людей з особливими потребами, що, своєю чергою, впливає і на особливості відповідного інформаційного ресурсу;
- реалізація технологічних процесів на основі алгоритмів подання інформаційного бібліотечного ресурсу, створюваного для кожної категорії таких користувачів;
- розроблення методичних рекомендацій щодо реалізації бібліотечної підтримки і інформаційно-технологічного супроводу освітнього процесу осіб з обмеженими можливостями.

У формуванні електронних бібліотек важливим аспектом є організація і обслуговування спеціалізованих інформаційних сховищ, поданих у цифровій формі з використанням відповідних програмно-технічних систем їх формування. Також дуже важливо є створити умови для програмно-технічних комплексів локального та дистанційного використання цих систем [6]. Електронні бібліотеки є фактично електронними аналогами традиційних бібліотек, проте в них за допомогою програмно-алгоритмічних комплексів створюється інформаційне середовище, яке надає нових можливостей, зокрема для покращення процедур доступу до інформації особам з особливими потребами. Водночас зберігання та опрацювання фондів електронних бібліотек, надання доступу до них висуває ряд нових специфічних вимог до кваліфікації та діяльності бібліотекарів. Вони повинні забезпечувати перехід до нових технологій проблемно-орієнтованого спрямування та освоєння процесів роботи з електронними інформаційними ресурсами.

Електронна інформація, що присутня у системах соціальних комунікацій, подається у вигляді електронних документів з візуальною формою відображення даних. Мультимедійні електронні бібліотеки для осіб з особливими потребами – це електронні бібліотеки, які забезпечують формування контенту, зорієнтованого на задоволення інформаційних потреб користувачів з врахуванням різних форм нозології. Такі бібліотеки зберігають та опрацьовують інформацію, подану у текстовій формі, а також в аудіо- та відеоподанні. Мультимедійні бібліотеки можуть також містити типи контенту, які не були підтримані в традиційних бібліотеках, такі як 3D-об'єкти. Одним із головних завдань мультимедійної бібліотеки для осіб з особливими потребами є надання ефективного доступу до інформаційного контенту, поданого у зручному для використання форматі.

Електронна бібліотека є посередником між інформаційними потребами користувачів та адаптованим для їх забезпечення контентом.

Існують чотири області завдань, виконання яких необхідне для ефективного забезпечення цього посередництва:

- Відбір контенту: бібліотека аналізує та відбирає якісний контент, орієнтований на інформаційне забезпечення користувачів з різними нозологіями.
- Структурування контенту: бібліотека структурує контент відповідно до можливостей сприйняття інформації користувачів з особливими потребами.
- Збагачення контенту: бібліотека поповнює вміст електронної бібліотеки метаданими, екстрагованими бібліотекарями з наявних джерел.
- Бібліотечне обслуговування: надання доступу до контенту для його використання користувачами.

Ці завдання дають змогу цифровій бібліотеці скоротити розрив, який існує між великою кількістю наявних матеріалів і конкретними інформаційними потребами користувачів з особливими потребами.

А. Пошукові технології

Пошукові технології необхідні для ефективного і результативного аналізу та відбору відповідного контенту для поповнення фонду цифрової бібліотеки, а також пошуку в ньому релевантної інформації. Використовуються два підходи до пошуку в електронних бібліотеках, які можуть бути взаємопов'язаними: з одного боку, методи інформаційного та мультимедійного пошуку, з іншого боку, пошук на основі метаданих.

Інформаційно-пошукова система електронної бібліотеки заснована на аналізі змісту об'єктів електронної бібліотеки за ключовими словами або індексами. Для ефективного пошуку інформації пошукові індекси автоматично вибудовуються і оновлюються. Метаданими вважаємо дані другого ступеня тобто дані про масиви документів цифрової бібліотеки. Метадані існують для опису інформаційних об'єктів і для підтримки інформаційних процесів у межах відповідної області. За допомогою метаданих у сховищі даних формується масив інформації. Їх можна використати для пошуку релевантної інформації, при виборі потрібних об'єктів у списку результатів пошуку, при формуванні вмісту електронних бібліотек (наприклад, використовуючи як пошукові образи прізвище автора роботи, отримуються відомості про опублікування документа). Управління метаданими можна порівняти з процесом формування карткових каталогів в традиційних бібліотеках.

В. Інші послуги електронної бібліотеки

Окрім пошукових послуг, електронна бібліотека забезпечує інші форми бібліотечно-інформаційного обслуговування. Вони передбачають співпрацю з користувачами в процесі збирання контенту, даючи їм змогу включати свій власний контент у фонди електронних бібліотек.

Анотаційні послуги дозволяють користувачам, додавати анотації у вигляді коментарів, рейтингів і т.д. Таку анотацію можна використати в процесі пошуку, оскільки коментарі про документ дають додаткову інформацію та зміст і його цінність для користувачів.

Окрім обслуговування користувачів, електронна бібліотека також підтримує адміністративні процеси управління її діяльністю. Основною метою цього процесу є цілеспрямоване формування фонду електронних документів для кожної категорії користувачів з особливими потребами. При цьому застосовується технологія ухвалення рішень щодо включення нових документів або видалення контенту з фонду. Періодично відбувається реструктуризація фонду та його кластеризація. Адміністративні процеси електронної бібліотеки передбачають управління формуванням підсистеми користувачів, оцифровкою контенту, створенням та формуванням метаданих. До кола управлінських функцій входить з'ясування коректності дотримання прав доступу, авторських прав та прав інтелектуальної власності, що сприяє підвищенню довіри користувачів до послуг електронної бібліотеки [7].

С. Сучасні тенденції формування електронної бібліотеки для осіб з особливими потребами

Перше покоління електронних бібліотек було побудовано як експеримент. Після того, як стали зрозумілими усі переваги такої бібліотеки, було встановлено базову функціональність типової електронної бібліотеки. Розроблені системи управління електронними бібліотеками на платформах DSpace та Greenstone містять модулі функціональності управління такою бібліотекою. Такі системи тепер доступні і використовуються в різних електронних бібліотеках. Останні тенденції в області технологій електронної бібліотеки є більш децентралізованими та із сервіс-орієнтованим підходом до її архітектури. Загальна мета тут полягає в систематичному доопрацюванні функцій електронної бібліотеки для підвищення ефективності доступу до інформації користувачів з особливими потребами, зниження витрат на розроблення цієї технології, поліпшення гнучкості, адаптованості. У цьому контексті з сучасних технологічних розробок використовуються такі, як Grid Computing, Web, Grid Services, Peer-to-Peer. Загалом електронні бібліотеки переходять на режим динамічного управління послугами, що забезпечує ефективне обслуговування користувачів з особливими потребами.

Друга тенденція полягає у інноваційному підході до формування та управління контентом, який ґрунтується на оновленому розумінні ролі електронної бібліотеки в контексті сучасних вимог до інформаційного забезпечення користувачів з особливими потребами. Загалом ідея електронної бібліотеки полягає в підтримці інформаційних процесів суспільства, забезпеченні швидкого доступу до інформації осіб з особливими потребами.

Третя тенденція, притаманна електронній бібліотеці, полягає у використанні Web- технологій для інтелектуального пошуку релевантних документів за запитом користувачів з особливими потребами. Ця технологія передбачає анотування документів із застосуванням методу онтологій і кластеризації результатів запиту. Методом онтологій також можна вибудовувати моделі потенційних користувачів та їх інформаційних потреб. Такий підхід не використовується у бібліотечній практиці персоналізації користувачів. У середовищі електронної бібліотеки обов'язково задіюється контекст документів для створення їх пошукових образів, наприклад, окремі відомості з вмісту документів пов'язуються через гіперпосилання із зовнішніми інформаційними ресурсами або метаданими.

II. Вимоги до інформаційних продуктів для осіб з особливими потребами

У сучасних умовах вже немає потреби доводити, що парадигма інформаційного обслуговування, яка ґрунтується лише на основі паперових носіїв, застаріла, і на зміну їй приходить нова – заснована на електронному представленні найрізноманітнішої документної інформації, що тиражується в необмеженій кількості і є доступною через глобальні комп'ютерні мережі незалежно від часу і місцезнаходження користувача. Прискорення темпів створення документної інформації, удосконалення засобів її збереження і передавання, зміна принципів взаємодії людей (бібліотекар –

користувач) в процесі обміну цією інформацією призвели до виникнення нових форм подання інформаційних ресурсів.

Інформатизація бібліотек на практиці – це нові можливості в обслуговуванні читачів і абонентів, забезпечення доступності фонду, збереження каталогів, вивільнення співробітників від ряду рутинних робіт з підготовки картотек, видань, списків, замовлень, листів, звітної документації тощо. Ці завдання можливо вирішити лише за умови активного використання прогресивних інформаційних технологій на основі мережевої інтеграції інформаційних ресурсів.

У інформаційному суспільстві сформовані передумови для переходу до нової парадигми обслуговування користувачів. Це, насамперед, перетворення бібліотек на бібліотечно-інформаційні центри, що надають широкий спектр послуг. Для ефективного виконання своєї місії такі центри потребують оснащення необхідною технікою, адаптивними технологіями і пристроями. Це дає змогу сформувати віртуальне середовище бібліотеки, впроваджувати автоматизовані інформаційно-бібліотечні системи, формувати власні електронні ресурси та надавати доступ до передплачених ресурсів. І тут важливо усвідомити, що доступ до інформації в бібліотеках повинні мати всі категорії громадян, зокрема користувачі з особливими потребами.

Утвердження соціальної місії бібліотек як інституту, що забезпечує основні права людей з особливими потребами при доступі до різних видів інформації, інтелектуальної і творчої діяльності, спонукає до модернізації їх роботи, набуття функціональних засад інформаційно-ресурсного центру, який покликаний забезпечувати доступність і поширення інформаційних ресурсів для осіб з фізичними обмеженнями, а також поєднувати в собі інформаційні, освітні і загальнокультурні функції.

Мультимедійний контент електронних бібліотек для осіб з особливими потребами потребує диференційованого подання навчального матеріалу відповідно до нозологій користувачів. Впровадження інноваційних технологій формування електронних бібліотек та обслуговування користувачів з особливими потребами неминуче спричиняє потребу в навчанні персоналу, а також передбачає витрати на постійний розвиток персоналу.

Ефективне бібліотечне обслуговування користувачів з обмеженими можливостями передбачає створення бібліотечних фондів для задоволення специфічних потреб цієї кожної категорії користувачів із специфічними інформаційними потребами.

Для обслуговування користувачів з вадами зору бібліотечні фонди повинні формуватися з книг, набраних шрифтом Брайля, а також книг, що “розмовляють”, виготовлених з використанням сучасних цифрових форматів, прийнятих міжнародним консорціумом Дейзі.

Для задоволення інформаційних потреб користувачів із проблемами слуху бібліотеки повинні володіти документами, які подають інформацію у формі відеорядів жестових повідомлень.

Для читачів, які не мають можливості безпосередньо відвідувати бібліотеку, повинні передбачатися послуги з надання інформації у формі відсканованих або оцифрованих документів з бібліотечного фонду.

Кожній категорії цих користувачів надається можливість отримання інформації у найзручнішій для них формі. Бібліотечно-інформаційне обслуговування користувачів з обмеженими можливостями висуває певні вимоги щодо облаштування і технічно-технологічного забезпечення читачів окремими автоматизованими робочими місцями, проблемно-орієнтованими читальними залами, швидкісними телекомунікаційними інтернет-каналами.

Важливим технологічним кроком створення електронних бібліотек стало формування інформаційного контенту, що ґрунтується на принципово новій цифровій інформаційній технології зберігання інформації на мікрочипах та компакт-дисках, які отримали назву – “цифрова книга, що розмовляє” (Digital Talking Books). Саме цей інформаційний контент насамперед накопичується у фондах електронних бібліотек, що обслуговують користувачів з особливими потребами.

Використовуючи технологію DAISY, бібліотеки і бібліотекарі мають змогу розширювати та впроваджувати на практиці досягнення у галузі інформаційних технологій, що надають змогу незрячим людям “бачити” достатні обсяги інформації, “йти в ногу” з розвитком суспільства, досягти результатів, які під силу здоровим людям. Синхронізована мультимедіа – це ключова інформаційна технологія для бібліотечного обслуговування користувачів з особливими потребами у найближчому майбутньому.

Книга у форматі DAISY може складатися з аудіоконтенту, з тексту і зображень або їх комбінації. Всі книги цього формату використовують спільний набір файлових типів. Практично всі типи файлів ґрунтуються на XML. Щоб книга у DAISY-форматі була зручною для використання, вона повинна містити такі типи файлів: пакетний; текстового контенту; зображення; аудіофайл; файл синхронізації; файл навігаційного контролю; ресурсний файл; файл презентації стилю; файл перетворень.

Цифрова книга, що “розмовляє” (DTB), – це набір електронних файлів відповідної тематики для подання інформації за допомогою альтернативних засобів відображення та відтворення інформації, а саме синтезованих мовних текстів, рельєфно-крапкового алфавіту Брайля, візуальних зображень і зокрема широкоформатного друку. Файли, сформовані і подані в DTB відповідно до зазначеного стандарту, створюють широкий спектр можливостей, серед яких швидка, гнучка навігація; маркування і використання закладок; пошук за ключовими словами; пошук слів за звуковим запитом; користувацький контроль відібраного за запитом контенту (наприклад, зноски, номери сторінок тощо). Такі способи подання інформації допомагають читачам з вадами зору. Контент книг, що “розмовляють” DTB, подається за допомогою аудіо-сигналів, комбінації аудіо, тексту і зображень або лише за допомогою тексту. Навігація, яка використовується при створенні книг, що “розмовляють”, дає змогу незрячим читачам “гортати” книгу, тобто пропускати зазначені розділи і сторінки, абзаци, відзначати голосовою закладкою необхідну інформацію в цифровому аудіозаписі.

Для читачів з проблемами слуху слід забезпечити можливість роботи з документами, представленими у відеоформаті.

Для цього бібліотека повинна мати у своєму розпорядженні декілька складових:

- тренажер жестової мови;
- фонд документів, що містять інформацію у вигляді відеорядів жестового перекладу;
- програмно-апаратні засоби для відтворення інформації з відеодокументів (книги, журнали, газети).

Специфіка бібліотечного обслуговування користувачів з обмеженими можливостями пересування вимагає вирішення питань, пов'язаних з розробленням методичних рекомендацій організації надання послуг дистанційно:

- забезпечення on-line доступу до бібліотечних фондів;
- пересилання документів у цифрових форматах або відсканованому вигляді.

Вимоги, які ставляться до зберігання та опрацювання інформації, достатньо складні. Масиви, які необхідно зберігати, дуже великі, більшість процесів потребують запаралелення, масиви записів потребують редагування та внесення змін, а також постійного контролювання в реальному масштабі часу. Найважливіші процеси опрацювання інформації, зокрема аналіз змісту, класифікація, пошук і отримання необхідних документів повинні виконуватися інтерактивно під управлінням користувача з одночасним коригуванням і підтриманням в робочому стані інформаційних масивів із врахуванням змін, що виникають у процесі опрацювання інформації. Все це ставить до концепції створення електронної бібліотеки надзвичайно високі вимоги та потребує чіткості при плануванні всіх складових, застосування нових методів організації масивів, що спрощують виконання операцій в динамічно змінних умовах опрацювання інформації.

При обслуговуванні користувачів зі особливими потребами особливої ваги набуває надання доступу до інтернет ресурсів через Wi-Fi. Можливість автономної роботи користувача на території бібліотеки створює додаткові зручності для людини, яка приходить у бібліотеку зі своїм комп'ютером, iPhone чи Daisy-плеєром. Це надає можливість поєднувати роботу з документами, що зберігаються на цих пристроях, з інформацією, отриманою зі світової інформаційної павутини, розташувавшись у найзручнішому місці, почувавши себе комфортно.

Достатньо використовуваною в бібліотеках, що обслуговують користувачів з особливими потребами, є послуга, що передбачає використання інтернет-телефонії. Найпоширенішим є використання вільного програмного забезпечення Skype, що дає змогу здійснювати телефонні дзвінки безпосередньо з комп'ютера. Програма забезпечує безкоштовний зв'язок абонентам, на комп'ютерах яких встановлено Skype. До переваг цього програмного забезпечення належать якість голосового зв'язку, простота використання програми і захист безпеки з'єднання. Інформаційна

технологія, що базується на використанні Skype, дає змогу проводити професійні конференції та консультування користувачів. При наданні консультацій можна використовувати наочні матеріали, посібники та підручники, відеоматеріали.

Формування електронних бібліотек пов'язане з радикальними змінами наявних методів, забезпечення користувачів можливістю доступу до баз даних в оперативному режимі за допомогою віддалених терміналів. Технології обслуговування осіб з особливими потребами вимагають формування вартісних файлів із забезпеченням швидкого доступу до них.

Електронним бібліотекам вдається вийти на якісно новий рівень задоволення інформаційних потреб науковців завдяки використанню новітніх інформаційних технологій. Неабияку роль відіграють кількість доступних інформаційних джерел, ступінь їх релевантності, актуальність, повнота й оперативність отримання інформації.

Створення електронної бібліотеки породжує цілком нове коло проблем, серед яких визначення обсягів формування фондів та побудови стратегії її формування, об'єднання ресурсів декількох мереж, створення віддалених сховищ та просторів даних, організація доступу до них в межах дотримання авторського права та розроблення системи захисту інформації та санкціонованого доступу до неї.

Побудова електронної бібліотеки повинна ґрунтуватися на таких основних принципах:

- формування системи, в якій за однократного введення інформації забезпечується послідовне її опрацювання, тобто введена інформація змінюється;
- залучення на корпоративних засадах розпорощених ресурсів і надання до них доступу;
- створення адаптивного середовища, що забезпечує вплив користувачів на основні інтелектуальні процеси формування інформаційних ресурсів.

Такий підхід забезпечує електронну бібліотеку можливістю швидко реагувати на потреби користувачів з метою забезпечення їх релевантною інформацією.

Висновки

Застосування всіх видів інформації в одному сховищі даних підвищує якість інформаційно-бібліотечного обслуговування. Поєднання відео, тексту, зображень та звуку сприяє наданню ефективної допомоги для забезпечення доступу до релевантної інформації. Але повною мірою потенціал використання інформаційних технологій для бібліотечного обслуговування цієї категорії користувачів може бути реалізований тільки після того, як відбудеться деяка реорганізація матеріально-технічної бази бібліотеки.

Оскільки вільне використання інформації з інтернету має низку недоліків, саме бібліотекарі електронних бібліотек повинні аналітично опрацьовувати інформацію, фільтруючи неважливу та недостовірну.

Призначення електронних бібліотек полягає не тільки у ефективному забезпеченні інформацією та швидкому доступі до інформаційного вмісту, а також організації сховищ даних, які забезпечують довгостроковий доступ до об'єктів контенту. При цьому слід враховувати тенденції та швидку зміну в технологіях зберігання даних.

Запропонована технологія мультимедійного бібліотечно-інформаційного обслуговування осіб з особливими потребами дає змогу диференційовано надавати бібліотечно-інформаційні послуги користувачам та удосконалити існуючі форми їх обслуговування у всіх бібліотеках на засадах реального включення осіб з інвалідністю в активне суспільне життя. Це, своєю чергою, зумовлює необхідність формування спеціального інформаційно-комунікаційного бібліотечного середовища для обслуговування усіх категорій користувачів, зокрема і з обмеженнями життєдіяльності (з особливими потребами). За такої системи обслуговування користувачам надаються інформаційні та сервісні бібліотечні послуги, які враховують особливості інформаційно-бібліотечного забезпечення конкретного споживача з урахуванням особливостей комунікаційних каналів сприймання інформації на засадах реального включення осіб з інвалідністю в суспільні процеси.

Використання можливостей мультимедіа в бібліотеках викликає потребу подальшого теоретичного осмислення феномену “е-освіта” як в аспекті технологічних змін, так і дидактико-навчальних трансформацій; визначення і розроблення інформаційно-аналітичної складової діяльності електронних бібліотек в плані максимально-повного опрацювання і представлення

інформаційних ресурсів. Перспективним напрямом дослідження слід вважати розроблення засобів доступу до наукової інформації в електронних бібліотеках. Крім цього, потребує подальшого вивчення питання створення модуля мета-даних як інструменту доступу та пошуку інформації в електронному оточенні, а також проектування соціотехнічних систем у сфері бібліотечного обслуговування користувачів з особливими потребами.

1. Шрайберг Я.Л. *Современные тенденции развития библиотечно-информационных технологий* / Я.Л. Шрайберг // *Ежегодный пленарный доклад международных конференций "Крым"*, 2001. – М.: Изд-во ГПНТБ России, 2002. – С. 44.
2. Сукиасян Э.Р. *Дискуссионный клуб "Термин"* / Э.Р. Сукиасян // *Научные и технические библиотеки*. – 2000. – № 6. – С. 113–119.
3. Watstein S.B. *Digital library: keywords* / S.B. Watstein, P.V. Calarco, F.S. Ghaphery // *Reference Services Review*. – 1999. – Vol. 27, № 4. – P. 344–352.
4. Крапивенко А.В. *Технологии мультимедиа и восприятие ощущений: учеб. пособие* / А.В. Крапивенко. – М.: БИНОМ. Лаборатория знаний, 2009. – 271 с.
5. Дунаев В.В. *Macromedia Flash MX 2004* / В.В. Дунаев. – С.-Петербург: "Питер", 2004. – 368 с.
6. Давидова Е.В. *Создания графики для Web-страниц* / Е.В. Давидова // *Информатика та освіта*. – 2001. – № 6. – С. 72–80.
7. Шлыкова О.В. *Культура мультимедиа: учебное пособие для студентов* / О.В. Шлыкова. – М.: ФАИР-ПРЕСС, 2004. – 416 с.
8. *The Cambridge Handbook of Multimedia Learning* / [ed. by Richard E. Mayer]. – New York: Cambridge University Press, 2005. – 680 p.
9. Arms William Y. *Digital Libraries and Electronic Publishing* / William Y. Arms. – Cambridge: MIT Press, 2000. – 287 p.
10. Reese Terry *Building digital libraries* / Terry Reese, Kyle Vanerjee. – New York: Neal-Schuman Publishers, 2008. – 277 p.
11. Воройский Ф.С. *Основы проектирования автоматизированных библиотечно-информационных систем* / Ф.С. Воройский. – М.: Физматлит, 2002. – 384 с.
12. Земсков А.И. *Электронные библиотеки* / А.И. Земсков, Я.Л. Шрайберг. – М.: Либерия, 2003. – 352 с.
13. Шрайберг Я.Л. *Современные тенденции развития библиотечно-информационных технологий: ежегодный доклад* / Я.Л. Шрайберг // *Науч. и техн. б-ки*. – 2002. – № 1. – С. 25–47.
14. Кожевникова Е.С. *Современные информационно-библиотечные технологии в библиотеке вуза: проблемы создания электронной библиотеки* / Е. С. Кожевникова // *Науч. и техн. б-ки*. – 2002. – № 7. – С. 73–78.
15. Калюжная Т.А. *Электронная библиотека ГПНТБ СО РАН* / Калюжная Т.А., Шнурик В.В. // *Библиосфера*. – 2007. – № 2. – Режим доступа: <http://cyberleninka.ru/article/n/elektronnaya-biblioteka-gpntb-so-ra>.
16. Елизаров А.М. *О концепции проекта "Электронная библиотека Республики Татарстан"* / М.Р. Биктимиров, Э.Е. Шабаишвили, А.М. Елизаров // *Вестник РФФИ*. – 2001. – № 3. – С.37–46.
17. Лапкина Н.А. *Критерии качества оцифрованной информации* / Н.А. Лапкина. – Режим доступа: <http://www.ifapcom.ru>.

ПОБУДОВА ДИФУЗІЙНОПОДІБНОЇ МОДЕЛІ ІНФОРМАЦІЙНОГО ПРОЦЕСУ ПОШИРЕННЯ ЗНАННЄВОГО ПОТЕНЦІАЛУ

© Бомба А., Назарук М., Пасічник В., 2014

Описано інформаційну модель освітнього середовища крупного міста, проаналізовано процеси освітньо-кваліфікаційного зростання особистості, починаючи від перших освітніх процесів до повної професійної зрілості та подальшого підвищення відповідного знаннєвого потенціалу. Запропоновано процеси поширення знаннєвого потенціалу у соціокомунікаційному середовищі подавати у вигляді дифузійноподібної моделі.

Ключові слова: освітнє середовище міста, агент, клік, знаннєвий потенціал, дифузія.

An information model of the city educational environment have been described, the processes of educational qualification growth of the individual from the birth up to full profession readiness and further increment of the knowledge potential have been analyzed. There have been suggested processes of the knowledge potential propagation within socio-communicative environment in the form of the diffusion-like model.

Keywords: educational environment of the city, agent, click, knowledge potential, diffusion.

Постановка проблеми у загальному вигляді

У сучасному світі відбуваються глобальні зміни в підходах до отримання знань – трансформуються як освітні системи, так і процеси отримання та опрацювання інформації школярами, студентами, людьми похилого віку, особами з особливими потребами та ін. Закономірним наслідком означених вище явищ є комп'ютеризація та інформатизація освітніх процесів, яка передбачає використання комп'ютерних та телекомунікаційних технологічних розроблень, інфокомунікаційних соціальних сервісів; пришвидшення процесів здобуття нових і закріплення раніше набутих знань, обміну інформацією. Відбувається просторове наближення та соціопсихологічна адаптація інформативних та пізнавальних освітніх матеріалів до кінцевого споживача, який використовує їх для власного самонавчання та професійного становлення в процесі неперервного навчання впродовж життя. Водночас із розвитком інформаційного суспільства суттєво зростають потоки інформації, швидкість її опрацювання та поширення.

Це, своєю чергою, зумовлює необхідність розроблення моделей інформаційних процесів поширення знаннєвого потенціалу, що уможливить проведення комплексного системного дослідження соціально-освітніх зв'язків, які в крупних соціополісах притаманні таким освітнім структурам, як: дитячі садки, школи, технікуми, коледжі, університети – з одного боку та бізнесу і суспільній громаді – з іншого, з метою реалізації вираженої спільної освітньої, наукової та культурної діяльності.

Базовою соціокомунікаційною структурною одиницею при дослідженні зазначених процесів обрано міський соціополіс, яскравим представником якого є крупне місто, яке за означенням має такі концептуальні ознаки:

- економічну основу (виробничо-функціональний та адаптивний стрижень соціополісу як самодостатнього соціоутворення);
- розвинений сектор соціального та гуманітарного профілю;
- розвинений сектор освітнього та наукового профілю;
- функціональний механізм запровадження інноваційних технологій тощо.

Такий вибір об'єкта дослідження, яким є сучасне крупне місто з розлогими та добре розвиненими системами освітньої, наукової та бізнесової галузей, дає змогу комплексно розробляти та опрацьовувати відповідні моделі з мінімальними обмежувальними факторами. Якщо йдеться зокрема про сферу освіти, то гіпотетично “модельний прототип крупного міста” містить дошкільні, шкільні, професійно-технічні, вищі навчальні заклади, наукові та науково-дослідні установи, які “активно продукують та споживають знання”. Водночас отримані результати моделювання можуть без особливих перешкод масштабуватись та переноситись на рівень суттєво менших за розмірами міст, населених пунктів та інших віртуальних мережеских соціоутворень. Загалом моделювання передбачається проводити для систем з кількістю суб'єктів від декількох одиниць до десятків мільйонів.

Аналіз сучасних досліджень та публікацій

Моделювання розвитку міста в сучасних умовах є важливою складовою державного управління. Все більше країн використовують моделі розвитку міст для прийняття управлінських рішень, використовуючи в своїй роботі спеціальні програмні засоби та інформаційні технології. Міська система належить до класу складних динамічних систем, один із напрямків моделювання її розвитку заснований на використанні загальних фізичних закономірностей, пропонується широкий спектр різноманітних математичних моделей розвитку території крупного міста, зокрема значну увагу звертають на моделювання транспортної мережі [1,2].

З'являється щораз більше наукових праць, присвячених застосуванню фізичних аналогій та відповідних закономірностей і законів для моделювання інформаційних процесів в економічних і соціальних системах. Зокрема в роботі [3] розглянуто формалізовані методи прийняття рішень щодо керування соціальними системами, які ґрунтуються на застосуванні спінових моделей Ізінга. Використання методів фізико-математичного моделювання дало змогу означити ентропію й оцінити ризики функціонування соціально-економічної системи [4]. Зазначимо, що використання термодинамічних законів показало високий рівень їх адекватності і в моделях управління містами [5].

Результати досліджень поширення інформаційних потоків в соціумі подано у роботах професора Д.В. Ланде, зокрема в [5] описано логістичну модель взаємодії інформаційних потоків; запропоновано підходи до створення інструментів моніторингу, адаптивного агрегування та опрацювання потоків інформації із глобальних комп'ютерних мереж для забезпечення інформаційно-аналітичної діяльності [6]; запропоновано оригінальний метод вордлет-діаграм для аналізу та візуалізації розповсюдження інформаційних масивів [7]; з використанням математичного апарату клітинних автоматів розроблено дифузійні моделі поширення інформації в соціальних та соціокомунікаційних сферах [8].

Наукові роботи Д.В. Ланде роблять значний внесок у розвиток теоретичних основ і практичних рішень для створення методів і засобів дослідження процесів поширення інформаційних потоків. Водночас у згаданих вище роботах не враховано особливостей обробки інформації у формі знань.

Метою статті є побудова дифузійноподібної моделі інформаційних процесів поширення знанневого потенціалу в освітньому соціокомунікаційному середовищі крупного міста.

Аналіз отриманих наукових результатів

Не претендуючи на вирішення всіх проблемних питань та усвідомлюючи поліаспектність предмета дослідження, в межах цієї статті сфокусуємо увагу на інформаційних процесах поширення знанневого потенціалу в освітньому середовищі крупного міста.

Інформаційна модель освітнього соціокомунікаційного середовища крупного міста

Місто як суспільний соціотвірний конструктивний елемент запускає механізми відтворення населення та є продуцентом і основним споживачем освітнього інформаційного продукту, який людина “споживає” протягом всього життя.

Як базовий прототип об'єкта досліджень у статті розглядається крупне місто з розгалуженою мережею навчальних закладів, наукових і науково-методичних установ, закладів післядипломної освіти з різними формами навчання, науково-виробничих підприємств, державних і місцевих органів управління освітою та самоврядних осередків в галузі освіти, яке наділене властивостями

ієрархічної структури, що має необхідні та достатні засоби для відтворення свого освітньо-виробничого потенціалу.

Соціокомунікаційне середовище у великих містах має систему міжособистісних відношень між суб'єктами навчально-виховного процесу та широкий спектр різноманітних видів діяльності, необхідних для соціалізації осіб, що навчаються, відповідно до вікових та індивідуальних соціокультурних потреб та можливостей.

Освітнє соціокомунікаційне середовище міста визначаємо як багатогранне та поліфункціональне утворення, що функціонує й розвивається в реальному часі й визначеному територіальному просторі (реальному чи віртуальному), діяльністю якого можна управляти та координувати, через яке здійснюється відповідний педагогічний вплив зовнішніх об'єктивних і суб'єктивних факторів на суб'єкти навчального процесу.

Загалом широкий спектр освітніх процесів, що проходять у соціокомунікаційному середовищі крупного міста, концептуально вкладаються в узагальнене методологічне подання понятійно-термінологічної тріади “ІНФОРМАЦІЯ – ДАНІ – ЗНАННЯ”. В статті використовуються загальноприйняті трактування цих термінів. Для поняття “знання” запропоновано подавати додаткову характеристику, за допомогою якої фіксується певний рівень знань особи, що подається як характеристика певної сукупності, суми знань того чи іншого індивіда, накопиченої впродовж відповідного життєвого періоду. Пропонується означувати цю характеристику як “знанневий потенціал” – j . Знанневим потенціалом може бути наділений як реальний суб'єкт, так і деякий віртуальний – бібліотека, інформаційний інтернет-ресурс і т.п.

Освітнє середовище крупного сучасного міста формується і функціонує в системі інформаційних потоків (перетікання знаннєвого потенціалу), де зокрема з використанням комп'ютерних мереж та комплексів інформаційно-комунікаційних технологій відбуваються процеси створення, засвоєння та передавання знань від груп одних суб'єктів іншим. Як логічний наслідок такого трактування концепту поширення знаннєвого потенціалу і є спроба опису цих процесів у вигляді відповідної дифузійноподібної моделі.

Відомо, що дифузія (лат. *diffusion* – поширення, розсіювання, взаємодія) в класичному традиційному розумінні терміна є процесом взаємного проникнення молекул або атомів однієї речовини між молекулами або атомами іншої, що зазвичай приводить до вирівнювання їхніх концентрацій по всьому займаному об'єму. У деяких ситуаціях одна з речовин уже має вирівняну концентрацію, і говорять про дифузію однієї речовини в іншій [10]. Традиційно процеси дифузії за умов конвекції та масообміну (наявності внутрішніх джерел) описують за допомогою такого диференційного рівняння:

$$\frac{\partial j}{\partial t} = D \left(\frac{\partial^2 j}{\partial x^2} + \frac{\partial^2 j}{\partial y^2} + \frac{\partial^2 j}{\partial z^2} \right) - \left(v_x \frac{\partial j}{\partial x} + v_y \frac{\partial j}{\partial y} + v_z \frac{\partial j}{\partial z} \right) + f(x, y, z, t, j, a_1, a_1, \dots, b_1, b_2, \dots), \quad (1)$$

де j – концентрація речовини, яка бере участь у моделюванні; D – коефіцієнт дифузії; t – час; x, y, z – просторові координати; v_x, v_y, v_z – швидкості конвективного перенесення; f – задана функція, що характеризує інтенсивність внутрішніх джерел (забруднень, тепла, радіації тощо залежно від предметної області); a, b – параметри, що характеризують відповідно “внутрішні індивідуальності” та зовнішні фактори. При цьому підкреслимо, що, наприклад, при виведенні простішого одновимірного дифузійного рівняння $j_t = D j_{xx}$ відштовхуються від трьох нескінченно близьких точок (див. рис.1 а)), де в силу “переливання” із \bar{x} до x значення j повинно зростати, а відносно \underline{x} і x спадати і, якщо процес відбувається на деякому відрізку $[0, l]$, то таких “трійок”, як відомо, розглядається континууміальна кількість.

У запропонованій нижче математичній моделі процесу поширення знаннєвого потенціалу при встановленні залежності зміни його значення для конкретного члена міського освітнього соціокомунікаційного середовища використовуватимемо традиційні дифузійні залежності (співвідношення), тобто в моделі фігуруватиме відповідно інформація стосовно сусідніх членів деякої групи (колективу) соціальної спільноти.

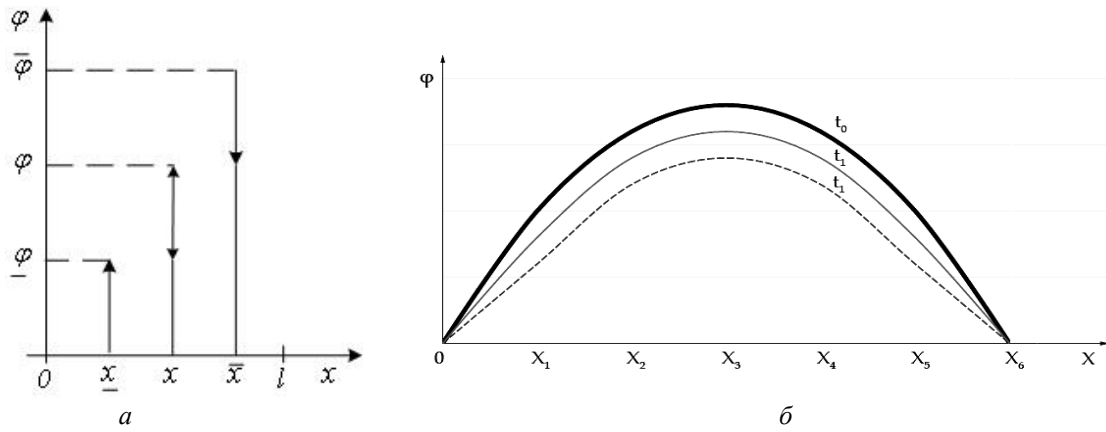


Рис. 1. Традиційна схема “дифузійного переливання” (а) і результати розрахунку знаннєвого потенціалу в початковий момент часу $t_0 = 0$ (жирним) та в момент часу $t_1 = 1$ за традиційною дифузійною схемою (суцільна лінія) та дифузійноподібною (пунктир) (б)

Міське освітнє соціокомунікаційне середовище, в якому знаннєвий потенціал поширюється дифузійно, в роботі [11] запропоновано подавати у вигляді мережевого графу $G = (A, R)$, де A – не порожня скінченна множина вузлів (агентів), R – множина неупорядкованих пар різних елементів з A (відношення між агентами).

У нашому випадку агенти – це вихованці, школярі, студенти, аспіранти та інші особи, які навчаються, а також вчителі, вихователі, науково-педагогічні працівники, батьки та представники бізнесу, установ, фірм, корпорацій, громадських організацій крупного міста, які беруть участь у навчально-виховних та освітніх процесах.

Відношення між агентами можуть зокрема інтерпретуватись як “дружба”, “співпраця”, “комунікація”, “навчання” та інші форми взаємодії, які притаманні суб’єктам освітнього соціокомунікаційного середовища міста.

Формування кліків у межах освітнього середовища на рівні шкіл міста

Агенти в межах певного освітнього рівня можуть об’єднуватися за спільними ознаками та властивостями (наприклад, вік, рівень освіти, навчання в одному закладі), тобто утворюють так звані кліки (соціальні спільноти) $K_j, (j = \overline{1, n})$ – підграфи чи класи, для яких зв’язки між вузлами всередині такого класу чи групи є міцніші та чисельніші, ніж між вузлами інших класів чи груп, а відповідно підграфів (рис. 2).

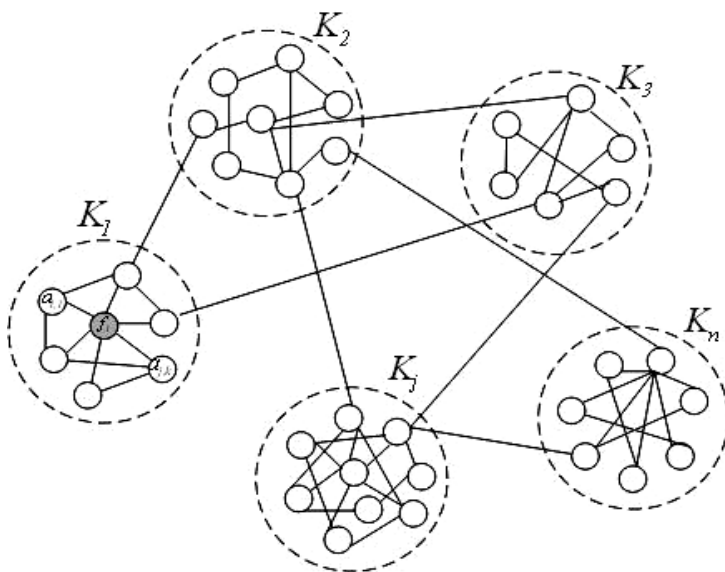


Рис.2. Схематичне зображення кліків освітнього соціокомунікаційного середовища міста в межах заданого освітнього рівня

Відповідні кліки (групи, асоціації, класи) в освітньому соціокомунікаційному середовищі крупного міста в багатьох випадках можуть формуватися з активним залученням сучасних мобільних телекомунікаційних засобів. Це, своєю чергою, технічно забезпечує процеси структуризації віртуального освітнього соціокомунікаційного середовища з мінімізацією впливу такого фактора, як фізична відстань між учасниками освітнього процесу, ознаками віку та специфіки освітніх потреб конкретної особистості. В одну освітню соціальну групу (кліку) можуть входити, на перший погляд, доволі різномірні за характеристиками індивідууми. Основним об'єднуючим фактором при формуванні та функціонуванні такої освітньої соціальної групи є необхідність підвищення знаннєвого потенціалу j її учасників за певним профілем та рівнем.

Знання в певний клік (групу) може надходити як від декількох фізичних суб'єктів, так і "віртуальних суб'єктів", з якими можуть бути певні мобільні або стаціонарні пристрої – "гаджети", що з'єднані іншими джерелами знаннєвого потенціалу.

Наприклад, агенти $a_{j,k} (j = \overline{1, k}, k = \overline{1, k_j})$, навчаючись в одному класі, отримуватимуть знаннєвий потенціал від вчителя, разом з тим агенти $a_{j,k}$ можуть входити в підгрупу із розроблення проекту з деякого навчального предмета, виконуючи який отримують також знаннєвий потенціал, але вже не від вчителя, а читаючи книги в бібліотеці чи інформаційні інтернет-ресурси за заданою темою (рис. 3).

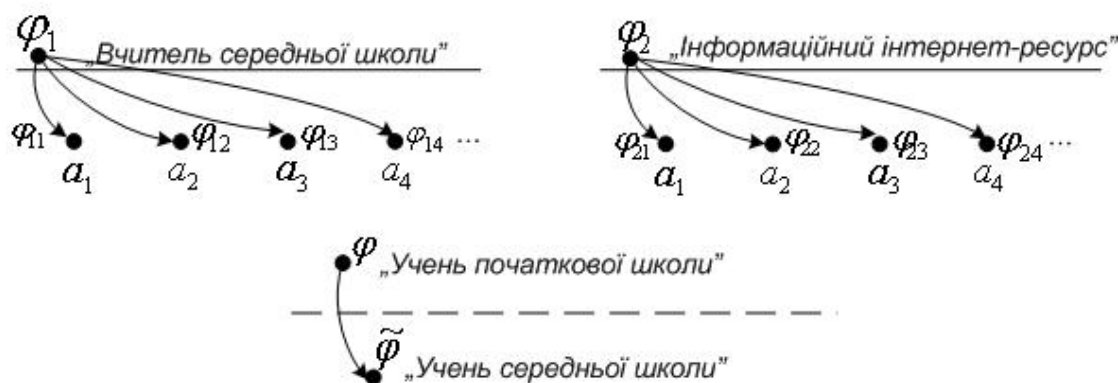


Рис. 3. Схематичне зображення процесу перетікання знаннєвого потенціалу

Зміна знаннєвого потенціалу j дозволяє суб'єктам освітнього соціокомунікаційного середовища міста (вихованцям ДНЗ, учням початкової школи, учням середньої школи, студентам бакалаврського рівня, студентам магістерського рівня та ін.) перейти на наступний освітній або освітньо-кваліфікаційний рівень.

Математична дифузійноподібна модель процесу поширення знаннєвого потенціалу

У роботі акцентовано увагу на описі (модельованні) процесів перерозподілу знаннєвого потенціалу в межах шкільного освітнього рівня міста, при цьому сформовано відповідні зовнішні та внутрішні залежності між учасниками (агентами) навчально-виховних та освітніх процесів.

Спочатку сформуємо залежності між агентами $a_{j,k} (j = \overline{1, k}, k = \overline{1, k_j})$. Через $j_{j,k,m}$ позначимо знаннєвий потенціал агента $a_{j,k}$ у деякий момент часу $t = t_m (m = 0, 1, 2, \dots; t_m = Dtm, \text{ де } Dt - \text{ деякий часовий інтервал})$. Для зручності викладу покладемо $Dt = 1$.

Агенти в межах певного освітнього рівня можуть об'єднуватися за спільними ознаками та властивостями (наприклад, вік, рівень освіти, навчання в одному закладі) у соціокомунікаційні спільноти (кліки), тому дифузійний процес поширення знаннєвого потенціалу між агентами в межах деякого кліка K_j , за формулою (1) запишемо:

$$f_{j,k,m} + D_{j,k,m} \sum_{I \leq \underline{k} < k < \bar{k} \leq k_j} s_{k,\underline{k},\bar{k}} \left((j_{j,\bar{k},m} - j_{j,k,m}) - (j_{j,k,m} - j_{j,\underline{k},m}) \right) = j_{j,k,m+1} - j_{j,k,m} \text{ або}$$

$$j_{j,k,m+1} = j_{j,k,m} + f_{j,k,m} + D_{j,k,m} \sum_{I \leq \underline{k} < k < \bar{k} \leq k_j} s_{k,\underline{k},\bar{k}} \left(j_{j,\bar{k},m} - 2j_{j,k,m} + j_{j,\underline{k},m} \right) \quad (2)$$

$$I \leq \underline{k} < k < \bar{k} \leq k_j, \bar{k} \neq \underline{k},$$

де $D_{j,k,m}$ – коефіцієнт, що характеризує здатність k -го агента j -ї освітньої групи перерозподіляти інформацію (знання) в момент часу m (аналог коефіцієнта дифузії); $f_{j,k,m}$ – числова характеристика основного джерела інформації (знань); $s_{k,\underline{k},\bar{k}}$ – деякі вагові коефіцієнти. Зауважимо, що джерелом інформації може бути один або декілька із виділених агентів певної соціокомунікаційної спільноти, наприклад, $f_{j,k,m} = g_m j_{j,k,m}$, де $k = \bar{k}_1, \bar{k}_2, \dots, \bar{k}_j, I < \bar{k}_1 < \bar{k}_2 < \dots < \bar{k}_j < k_j$ (якщо клік в межах шкільного освітнього рівня, це об'єднання учнів певного профільного класу, то роль джерела “знань” виконуватиме їх вчитель).

Безумовно, що окремої уваги та дослідження потребує побудова інформаційного (знаннєвого) аналогу процесу конвективного перенесення (перенесення знань “зі сторони” іншими шляхами). У цій роботі такого роду шляхи поширення знань не досліджувались.

Можливими є різні варіанти встановлення шляхів (способів) перерозподілу інформації (знань) між агентами, що належать до однієї соціальної спільноти (кліку), групами агентів різних кліків, а також між кліками в межах деякого освітнього рівня та в освітньому соціокомунікаційному середовищі міста загалом (при цьому слід вводити додатковий індекс, що відповідав би тому чи іншому освітньому рівню).

Одним із варіантів введення узагальненого потенціалу K_j -ї освітньої соціокомунікаційної спільноти $j_{j,m}$ є представлення його у вигляді деякої функції від $j_{j,k,m}$, зокрема у вигляді узагальненого середнього арифметичного через потенціали агентів певної спільноти:

$$j_{j,m} = \frac{1}{k_j} \sum_{k=1}^{k_j} a_{j,k} j_{j,k,m}, \quad (3)$$

де $a_{j,k}$ – деякий ваговий коефіцієнт.

Тоді закон перерозподілу відповідного “усередненого” знаннєвого потенціалу у часі між кліками можемо подати у вигляді:

$$j_{j,m+1} - j_{j,m} = f_{j,m} + D_{j,m} \sum_{I \leq \underline{j} < j < \bar{j} \leq j_k} w_{j,\underline{j},\bar{j}} \left((j_{j,\bar{j},m} - j_{j,m}) - (j_{j,m} - j_{j,\underline{j},m}) \right) \text{ або}$$

$$j_{j,m+1} = j_{j,m} + f_{j,m} + D_{j,m} \sum_{I \leq \underline{j} < j < \bar{j} \leq j_k} w_{j,\underline{j},\bar{j}} \left(j_{j,\bar{j},m} - 2j_{j,m} + j_{j,\underline{j},m} \right) \quad (4)$$

$$I \leq \underline{j} < j < \bar{j} \leq j_k, \bar{j} \neq \underline{j},$$

де $D_{j,m}$ – коефіцієнт, що характеризує здатність j -ї освітньої соціокомунікаційної спільноти перерозподіляти знаннєвий потенціал в момент часу m (аналог коефіцієнта дифузії); $f_{j,m}$ – числова характеристика основного джерела знань (відповідного освітнього рівня загалом); $w_{j,\underline{j},\bar{j}}$ – деякі вагові коефіцієнти. Зауважимо, що джерелом поширення інформації (знань) може бути одна або декілька із вибраних освітніх соціокомунікаційних спільнот певного освітнього рівня, наприклад $f_{j,m} = q_m j_{j,m}$, де $j = \bar{j}_1, \bar{j}_2, \dots, \bar{j}_k, I < \bar{j}_1 < \bar{j}_2 < \dots < \bar{j}_k < j_k$.

При цьому зворотний вплив значень знаннєвих потенціалів $j_{j,m}$ на $f_{j,k,m}$ у простішому випадку можемо моделювати шляхом введення залежності $f_{j,k,m}$ від $j_{j,m}$ ($f_{j,k,m} = g_j(j_{j,m})$, $j = \overline{1, n}$), наприклад, лінійно:

$$\left\{ \begin{array}{l} f_{j,k,m} = \sum_{i=1}^n b_{j,k,i} j_{i,m} = \sum_{i=1}^n b_{j,k,i} \left(\frac{1}{k_j} \sum_{k=1}^{k_j} a_{i,k} j_{j,k,m} \right), \\ k = \overline{1, k_j}, j = \overline{1, n} \end{array} \right. \quad (5)$$

де $b_{j,k,i}$ – деякі коефіцієнти, що визначаються (ідентифікуються) на основі накопичення, формування, аналізу та попереднього досвіду, зафіксованого у базах даних.

Отже, згідно з (2)–(5) для перерозподілу в часі знаннєвих потенціалів маємо:

$$\left\{ \begin{array}{l} j_{j,k,m+1} = j_{j,k,m} + f_{j,k,m} + D_{j,k,m} \sum_{1 \leq \underline{k} < k < \bar{k} \leq k_j} s_{k,\underline{k},\bar{k}} (j_{j,\underline{k},m} - 2j_{j,k,m} + j_{j,\bar{k},m}) \\ j_{j,m} = \frac{1}{k_j} \sum_{k=1}^{k_j} a_{j,k} j_{j,k,m} \\ j_{j,m+1} = j_{j,m} + f_{j,m} + D_{j,m} \sum_{1 \leq \underline{j} < j < \bar{j} \leq j_k} w_{j,\underline{j},\bar{j}} (j_{j,\underline{j},m} - 2j_{j,m} + j_{j,\bar{j},m}) \\ f_{j,k,m} = \sum_{i=1}^n b_{j,k,i} j_{i,m} = \sum_{i=1}^n b_{j,k,i} \left(\frac{1}{k_j} \sum_{k=1}^{k_j} a_{i,k} j_{j,k,m} \right) \end{array} \right. \quad (6)$$

Очевидно, що цю систему потрібно ще доповнити певними співвідношеннями (даними) – аналогами початкових та граничних умов для дифузійних рівнянь (робитимемо це під час побудови відповідного алгоритму розв’язку).

Алгоритму розрахунку перерозподілу знаннєвого потенціалу в часі:

1. Задаємо характерні модельні параметри, що входять в систему (6).
2. Задаємо початковий стан системи (розподіл знаннєвого потенціалу при $m = 0$):

$$j_{j,k,0} = \tilde{f}_{j,k}, \quad j = \overline{1, n}, \quad k = \overline{1, k_j} \quad \left(\sum_{j=1}^n k_j - \text{значень} \right).$$

3. За формулами (3) обчислимо узагальнений потенціал K_j -ї освітньої соціокомунікаційної спільноти в початковий момент часу:

$$j_{j,0} = \frac{1}{k_j} \sum_{k=1}^{k_j} a_{j,k} j_{j,k,0}.$$

4. Згідно з (5) знаходимо значення основного джерела інформації (знань) (наприклад, деякої вибраної освітньої соціокомунікаційної спільноти загалом) у початковий момент часу:

$$f_{j,k,0} = \sum_{i=1}^n b_{j,k,i} j_{i,0}.$$

5. За формулами (2) та (4) (при $m = 0$) обчислюємо $j_{j,k,1}$ та $j_{j,1}$ і переходимо до наступного часового інтервалу.

При цьому, як і в класичних дифузійних задачах, нам не вистачатиме, так би мовити, “крайніх значень” знаннєвого потенціалу j (тобто при $k = 1$ та $k = j$), що приводить до необхідності встановлення ієрархії агентів у межах кожного кліку заданого освітнього рівня (наприклад,

“вчитель–учень”), а також “крайніх умов” для характеристики знанневих потенціалів кожного кліку. В цьому випадку маємо 2 варіанти їх задання:

1) аналогічно до внутрішніх (агентів) характеристик вводимо ієрархію кліків і при цьому задаємо потенціали “крайніх”, наприклад, встановлюємо ієрархію шкіл у певному районі міста, задаючи потенціали “найслабшої” та “найсильнішої”;

2) вважатимемо, що кліки є рівноправні в межах освітнього рівня, при цьому упорядкування все ж вводимо “суто формально” (не за пріоритетами), вважаючи, що $j_{1,\dots} = j_{n,\dots}$, тобто задаємо умови, аналогічні умовам періодичності (дифузії).

Вважаємо за необхідне зробити деякі зауваження, які відображають певні особливості застосування запропонованого підходу.

Зауваження 1. Можливим є врахування в межах моделі взаємодії між агентами, що належать до різних освітніх соціокомунікаційних спільнот, наприклад, перерозподіл у часі знанневого потенціалу 3-го за списком в класному журналі учня деякого 10 класу з урахуванням впливу на нього 5-го за списком в журналі учня 11 класу деякої гіпотетичної школи, Використовуючи формулу (2), запишемо:

$$j_{10,3,m+1} = j_{10,3,m} + f_{11,5,m} + D_{11,5,m} \sum_{1 \leq \bar{k} < 3 < 5 \leq k_j} s_{3,\bar{k},5} (j_{11,5,m} - 2j_{10,3,m} + j_{10,\bar{k},m}).$$

Зауваження 2. Ще раз підкреслимо, що класичні дифузійні моделі для опису інформаційного процесу поширення знанневого потенціалу між агентами в межах деякого кліка K_j суттєво відрізняються від описаних вище дифузійноподібних. Зокрема на рис.1б) жирною лінією зображено розподіл знанневого потенціалу в початковий момент часу $t_0 = 0$, суцільною лінією – розподіл знанневого потенціалу, розрахований за традиційною дифузійною схемою в момент часу $t_1 = 1$, а пунктирною лінією зображено дифузійноподібні перерозподіли згідно із формулами (2–5) при $f_{j,k} = 1$, $D = 1$, при взаємодії навколишніх сусідніх агентів із заданими, $D = 0,1$ – в інших випадках, $s = 1$, коли знанневі потенціали знань “крайніх” агентів у межах деякого соціального кільця дорівнюють нулю, а початковий їх розподіл (за умов ієрархії) задано симетрично: $j_{0,0} = 0$, $j_{1,0} = 5$, $j_{2,0} = 8$, $j_{3,0} = 9$, $j_{4,0} = 8$, $j_{5,0} = 5$, $j_{6,0} = 0$.

Зауваження 3. Якщо джерелом знань є один із агентів у межах освітньої соціокомунікаційної спільноти, то, розв’язуючи обернені задачі типу ідентифікації параметрів моделей (див., напр., [12, 13]), можемо знаходити значення відповідного параметра $f_{j,k,m}$ для забезпечення необхідного інформаційного рівня знань цього агента-вчителя (наприклад, згідно із рис. 4, а), щоби вершина вихідної кривої залишалась незмінною в часі або знаходилась на певному інтервалі за умови малого початкового знанневого потенціалу (рис. 4, б)).

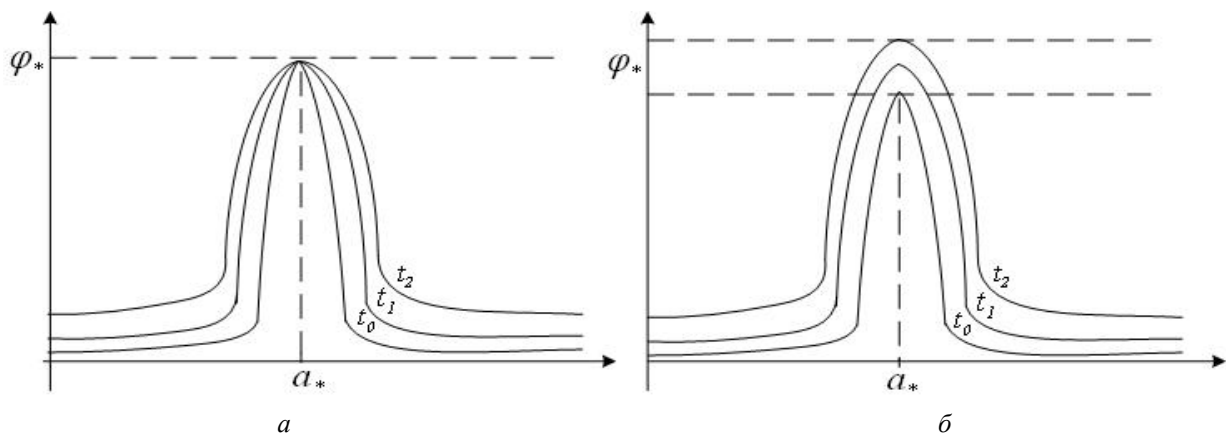
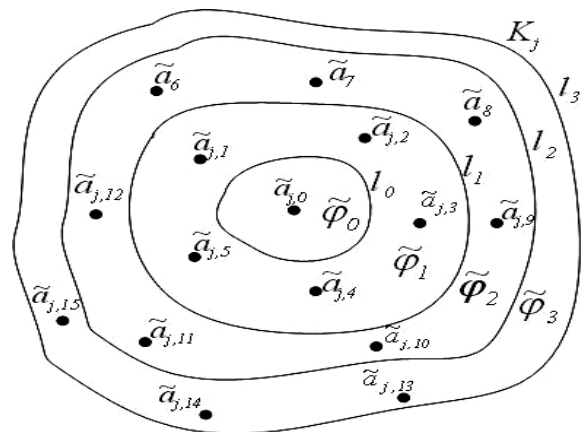


Рис.4. Схематичний розподіл передачі знань агентом-вчителем a_* за умови стабільності в часі його рівня знань

Зауваження 4. Маючи розрахований раніше знанневий потенціал j та вводячи певні характерні значення потенціалів, наприклад, вчителів \tilde{f}_0 , “відмінників” \tilde{f}_1 , учнів, які мають “добрі” \tilde{f}_2 та “задовільні” \tilde{f}_3 знання, в результаті їх порівняння із одержаними значеннями $j = j_{j,k,m}$ можемо побудувати деякі еквіпотенціали знань l_0, l_1, l_2, l_3 відповідно, що є лініями розділу освітніх соціогруп (прошарків) у межах заданого (j -го) кліку (див. рис.5, де $\tilde{a}_{j,k}$ – характеристики агентів згідно із відповідною перенумерацією).

Рис.5. Схематичне зображення формування еквіпотенційних ліній розділу освітніх соціогруп



Висновки

Освітнє середовище соціополісу формується і функціонує в системі інформаційних потоків, де на основі сучасних мереж зв'язку та комплексу інформаційно-комунікаційних технологій відбуваються процеси створення, засвоєння та передавання знань від одних суб'єктів іншим. Саме тому побудова моделей та моделювання інформаційного процесу поширення знанневого потенціалу уможливить проведення комплексного системного дослідження в межах крупного соціополісу соціально-освітніх зв'язків, які притаманні таким освітнім структурам, як дитячі садки, школи, технікуми, коледжі, університети – з одного боку та бізнесом і суспільною громадою міста – з іншого з метою реалізації вираженої стратегії спільної освітньої, наукової, виробничої та культурної діяльності.

Автори подали інформаційну модель освітнього середовища міста, проаналізували освітньо-кваліфікаційне зростання особистості, починаючи від народження до повної фахової зрілості. Запропоновано оригінальний модельний підхід до інформаційних процесів поширення знанневого потенціалу, який оснований на відповідній дифузійно-подібній моделі. Описано процеси перерозподілу знанневого потенціалу між агентами, що належать до однієї освітньої соціокомунікаційної спільноти (кліку), а також між агентами різних кліків у межах шкільного освітнього рівня міста. Автори розробили алгоритм розрахунку перерозподілу знанневого потенціалу в часі.

Подальші дослідження будуть спрямовані на розроблення математичної моделі процесів перерозподілу знанневого потенціалу між спільнотами різних освітніх рівнів в умовах соціокомунікаційного середовища крупного міста.

1. Mandelas A. Fuzzy Cellular Automata Based Shell for Modeling Urban Growth / A. Mandelas, T Hatzichristos, P. Prastacos // 10th AGILE International Conference on Geographic Information Science. – 2007.
2. Parish Y. Procedural Modeling of Cities / Y. Parish, P. Müller // Proceedings of the 28th annual conference on Computer graphics and interactive techniques. – 2001. – P. 301–308.
3. Згуровский М.З. Методы принятия решений в социальных системах на основе спиновых моделей Изинга / М.З. Згуровский, Т.Н. Померанцева // Проблемы управления и информатики. – 1995. – № 1. – С. 89–97.
4. Сявакво М.С. Нечітко-інтервальні методи оцінки ризику та ентропії поведінки соціально-економічної системи / М.С. Сявакво, О.М. Третяк // Економічна кібернетика. – 2006. – № 3–4(39–40). – С. 53–61.
5. Яцишин Ю.В. Модель управління містом на основі термодинамічних законів / Ю.В. Яцишин,

Н.Б. Шаховська // Вісник Національного університету “Львівська політехніка”. Серія Інформаційні системи та мережі. – 2000. – № 406. – С. 239–247. 6. Ландэ Д.В. Моделирование динамики информационных потоков / Д.В. Ландэ // Фундаментальные исследования. – 2012. – № 6. – С. 652–654. 7. Додонов А.Г. Сетевые информационные потоки как содержательная составляющая информационно-аналитических систем / А.Г. Додонов, Д.В. Ландэ, В.В. Жигало // Реєстрація, зберігання і обробка даних. – 2010. – N 1. – С. 39-48. 8. Ландэ Д.В. Объектно-статистический анализ информационных потоков / Д.В. Ландэ, В.Н. Фурашев // Сб. научн. трудов: Открытые информационные и компьютерные интегрированные технологии. – Х.: Нац. аэрокосмический ун-т “ХАИ”, 2007. – № 35. – С. 133–137. 9. Ландэ Д.В. Модель диффузии информации / Д.В. Ландэ // Информационные технологии и безопасность. Менеджмент информационной безопасности. Сборник научных трудов Института проблем регистрации информации. – 2007. – Вып. 10. – С. 51–67. 10. Будник А.Ф. Тепломасоперенос у процесах і матеріалах дизайну матеріалів: Навчальний посібник / А.Ф. Будник – Суми: Вид-во СумДУ, 2008. – 158 с. 11. Назарук М.В. Моделювання міського освітнього середовища як профільної соціальної мережі / М.В. Назарук, В.В. Пасічник // Міжнародний науково-технічний журнал “Інформаційні технології та комп’ютерна інженерія”. – Вінниця: ВНТУ, 2013. – № 3 (28). – С. 42–47. 12. Сергиенко І.В. Ідентифікація параметрів системи конвективно-диффузійного переносу / І.В. Сергиенко, В.С. Дейнека // Кибернетика и системный анализ. – 2009. – № 1. – С. 42–63. 13. Бомба А.Я. Ідентифікація параметрів сингулярно збурених задач типу “конвекція-дифузія-масообмін” в різнопористих середовищах / А.Я. Бомба, І.М. Присяжнюк, О.В. Присяжнюк // Сучасні проблеми математичного моделювання та обчислювальних методів. Матеріали Всеукраїнської наукової конференції, 22–23 лютого 2013 р., м. Рівне. – Рівне, 2013. – С. 34.

І. Огородник, *Т. Огородник
Національний університет “Львівська політехніка”,
кафедра автоматизованих систем управління,
*Центр інформаційного забезпечення

КОНЦЕПЦІЯ УПРАВЛІННЯ ЯКІСТЮ МІСЬКОГО СЕРЕДОВИЩА ВЕЛИКОГО МІСТА

© Огородник І., Огородник Т., 2014

Розглянуто основні поняття якості міського середовища, підхід до управління якістю міського середовища, сучасний стан міського середовища, недоліки управління, задачі управління якістю міського середовища. Запропоновано систему цілей управління якістю міського середовища, вибору показників якості на основі цієї системи та концептуальну модель управління якістю. Розроблено технологію управління якістю міського середовища, запропоновано бальну систему оцінювання показників якості, розглянуто приклад експертного оцінювання показників, показано, як розширювати діапазон оцінок показників якості, окреслено шляхи розвитку системи управління якістю міського середовища.

Ключові слова: місто, середовище, якість, управління, модель, цілі, аналіз, експертні групи, оцінки, розвиток, інформація, технологія, концепція.

The basic concepts of quality of city environment are considered, going near the management of quality of city environment, state of city environment for today, lacks of management, raising of task of quality management of city environment. The system of aims of quality management of city environment, choice of indexes of quality on the basis of this system and conceptual model of quality management, is offered. Technology of quality management of city environment is developed, the ball system of estimation of indexes of quality is offered, the example of expert estimation of indexes is considered, it is rotined, how to extend the range of estimations of indexes of quality, the ways of development of control the system by quality city environment are set.

Key words: city, environment, quality, management, model, aims, analysis, expert groups, estimations, development, information, technology, conception.

Вступ

Якість міського середовища визначається якістю його складових: антропогенного, природного та соціального середовищ. Під якістю міського середовища розуміють рівень задоволення фізіологічних, матеріальних та духовних потреб населення міста завдяки організації та функціонуванню цього середовища. Середовище міста повинно бути таким, щоб якнайкраще задовольняти потреби мешканця міста. Однак сьогодні управляють містом не як соціально-виробничим, а як суто виробничим комплексом. Систему управління містом мало цікавить, наскільки середовище міста задовольняє потреби мешканця. Більше цікавить, як всі сфери діяльності міста виробляють продукцію чи надають послуги і вимірюють це в квадратних метрах, тоннах, штуках, обслугованих мешканцях, посаджених деревах, кілометрах відремонтованих доріг і т.д. Менш за все систему управління містом цікавить, наскільки ця продукція задовольняє потреби мешканців міста. Тобто систему управління містом переважно цікавить кількість, а не якість продукції чи послуг.

В умовах ринкової економіки, тотальної конкуренції продукції та послуг, в умовах, коли споживач стає вимогливішим до умов проживання, такий стан речей, коли якість цих умов не знаходить належної

оцінки з боку системи управління, не може вважатися нормальним. Тому все більше і більше наукових розробок відносно міського середовища стосуються якості цього середовища.

Розглянемо концепцію побудови управління якістю міського середовища великого міста, застосовуючи системні методи та експертні оцінки показників якості міського середовища.

Аналіз існуючого стану управління якістю міського середовища

Система управління великим містом переважно працює за галузевим принципом: управління комунальним господарством, охорони здоров'я, будівництва, освіти, транспорту тощо [1]. Система управління містом планує і контролює діяльність міста у вказаних напрямках та сама оцінює результати управління. Населення міста в оцінці результатів діяльності міста участі не бере, хоча все, що робиться в місті, робиться для його мешканців. Результати діяльності системи управління містом здебільшого оцінюється за принципом – “план – факт”. Про яку якість роботи тут можна говорити? Система управління містом сама себе контролює, незважаючи на думку населення міста, для якого цю роботу виконано. Такий метод роботи системи управління містом є неприйнятним для мешканців. Мешканці міста беруть участь в управлінні містом тільки тоді, коли скаржаться на цю систему. Це по-перше.

По-друге. Недоліком роботи системи управління містом є небачення цілісної картини існуючої ситуації – вона контролює окремі фрагменти ситуації і діє некомплексно, отже, система управління не може визначити, який мікрорайон міста кращий від іншого і внаслідок чого.

По-третє. Система управління містом є доволі непрозорою для мешканців міста. Вона не інформує населення про результати своєї діяльності, мовчить про свої промахи, слабо у своїй діяльності використовує інтернет.

По-четверте. Відсутні чітко сформульовані цілі управління містом, і тому складно, а іноді неможливо визначити, чи повною мірою володіє проблемою система управління містом. Не визначено повноту функцій системи управління містом.

По-п'яте. Не визначено показників якості міського середовища – це не проста задача. До цієї роботи мають бути залучені спеціалісти різних напрямів. Це повинна бути консолідована інформація на всіх рівнях управління містом і всіх суміжних підрозділів системи управління містом [2].

Постановка задачі створення системи управління якістю міського середовища

Для створення системи управління якістю міського середовища необхідно розробити інформаційну модель управління, визначити систему цілей управління якістю міського середовища, показники якості міського середовища, концептуальну модель управління якістю, напрямки розвитку системи управління якістю міського середовища.

Аналіз системи цілей управління якістю міського середовища

На рис. 1 зображено систему цілей управління якістю міського середовища. Головна ціль управління якістю міського середовища – створення умов для задоволення фізіологічних, матеріальних і духовних потреб мешканців міста. Іншими словами – досягнення заданого рівня якості міського середовища. Як видно з рисунка, головної цілі буде досягнуто, якщо буде досягнуто цілей 1-го рівня: досягти якості антропогенного, природного та соціального середовища та цілей 2-го рівня: 1.1. – 1.4., 2.1. – 2.6., 3.1. – 3.8. Для того, щоб на рівні міста, району чи мікрорайону експлуатувати запропоновану систему управління, цих показників якості, на нашу думку, цілком достатньо [3]. У разі необхідності подальшої деталізації показників якості можна ввести 3-й рівень “дерева цілей”. Запропонована система показників охоплює показники організації об'єктів матеріального середовища і показники, які характеризують особистість. Цього достатньо, щоб належно оцінювати міське середовище і використовувати ці оцінки в широкому діапазоні потреб управління якістю міського середовища.

Як видно з дерева цілей, міське середовище повинно формувати високодуховного мешканця міста з громадянською позицією, соціально активного, з належним рівнем життя.

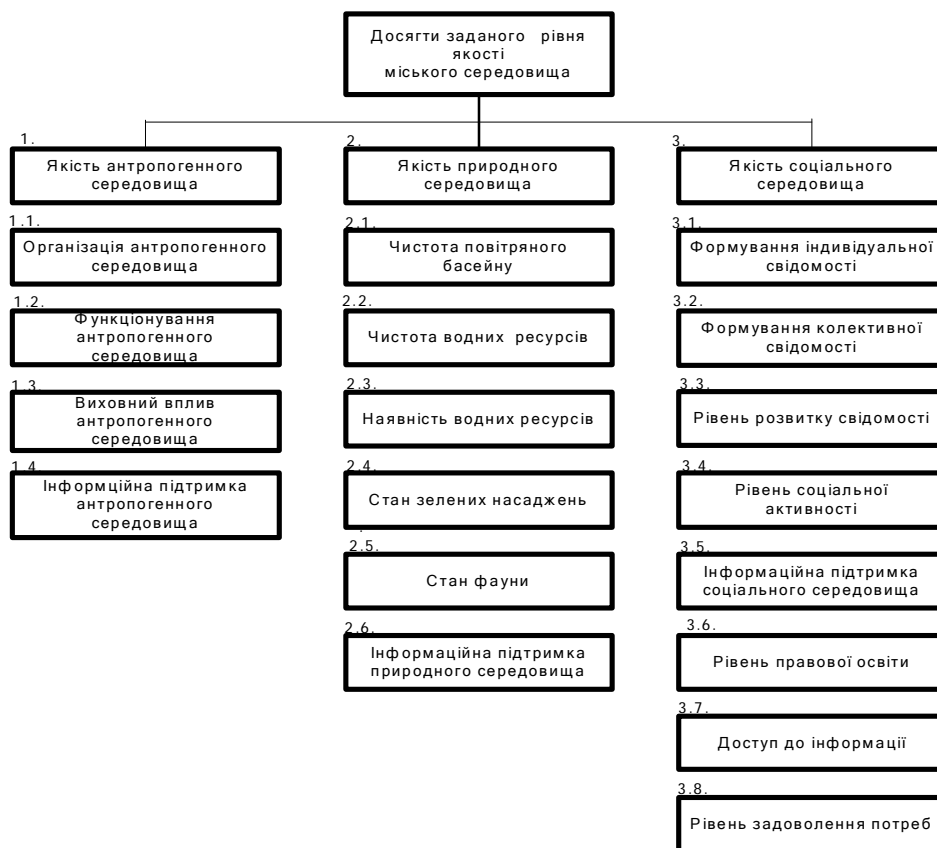


Рис. 1. Система цілей управління якістю міського середовища

Концепція управління якістю міського середовища

Міське середовище повонно бути так організоване і так функціонувати, щоби задовольняти не тільки фізіологічні та матеріальні, але й духовні потреби мешканців міста.

Кількість і якість таких послуг сьогодні здебільшого визначається нормативними документами. Проте міське середовище, створене на основі нормативів, не може враховувати зміни вимог мешканців міста до тих чи інших елементів міського середовища. Тому потрібно розробити таку систему управління якістю міського середовища, яка буде, крім нормативних вимог, орієнтуватися на вимоги мешканців міста щодо умов життя та духовного розвитку особи [4].

Крім цього, доцільно включити в систему управління населення міста як важливе джерело інформації. Дуже цінним було би регулярне опитування населення про те, чим задоволені і чим не задоволені мешканці міста і що вони пропонують покращити в організації та функціонуванні міського середовища [5]. Потрібне широке ознайомлення населення міста з результатами роботи системи управління та результатами обговорення тих чи інших пропозицій громади міста.

Система управління якістю міського середовища, яка пропонується в цій роботі, дасть можливість аналізувати стан міського середовища за різними показниками не тільки на рівні міста загалом, а й на рівні районів і мікрорайонів міста. А це стане можливим, якщо інформація буде прив'язана до просторових координат міста [6]. Таку прив'язку в роботі передбачено. Для цього територію міста розбивають на умовні квадрати – мікрорайони – із спорудами, житловими будинками, підприємствами, школами, дитячими садками, майданчиками, магазинами, лікарнями, об'єктами природного середовища, населення зі своєю демографічною структурою.

Потім розробляються показники міського середовища, які характеризують окремі об'єкти мікрорайону. На основі цих показників створюють відповідні бази даних та інформаційну систему "Якість середовища" [7, 8], яка:

- накопичує результати контролю показників якості міського середовища;

- актуалізує БД;
- аналізує якість міського середовища;
- видає результати аналізу органам управління містом.

Для забезпечення функціонування інформаційної системи необхідно створити відповідну структуру, яка здійснюватиме збирання і експертне оцінювання показників якості. Це міг би бути відповідний відділ міськвиконкому.

Визначати показники якості міського середовища спочатку пропонується експертно за 100-бальною шкалою [9–11]. Показники якості міського середовища виберемо, користуючись “деревом цілей” (рис. 1). Досягнення головної цілі оцінюємо за нормативом у 100 балів. Показники якості середовища оцінюємо так. Якість антропогенного середовища, що відповідає цілям 1-го рівня, оцінюємо в 40 балів. Якість природного середовища, що відповідає цілям 1-го рівня, оцінюємо в 25 балів. Якість соціального середовища, що відповідає цілям 1-го рівня, оцінюємо в 35 балів.

Приймаємо цей поділ як нормативний. На другому рівні здійснюємо аналогічний поділ. Отримуємо експертні оцінки фактичних показників якості, які зведено у таблицю.

На рис. 2 наведено концептуальну схему управління якістю міського середовища.

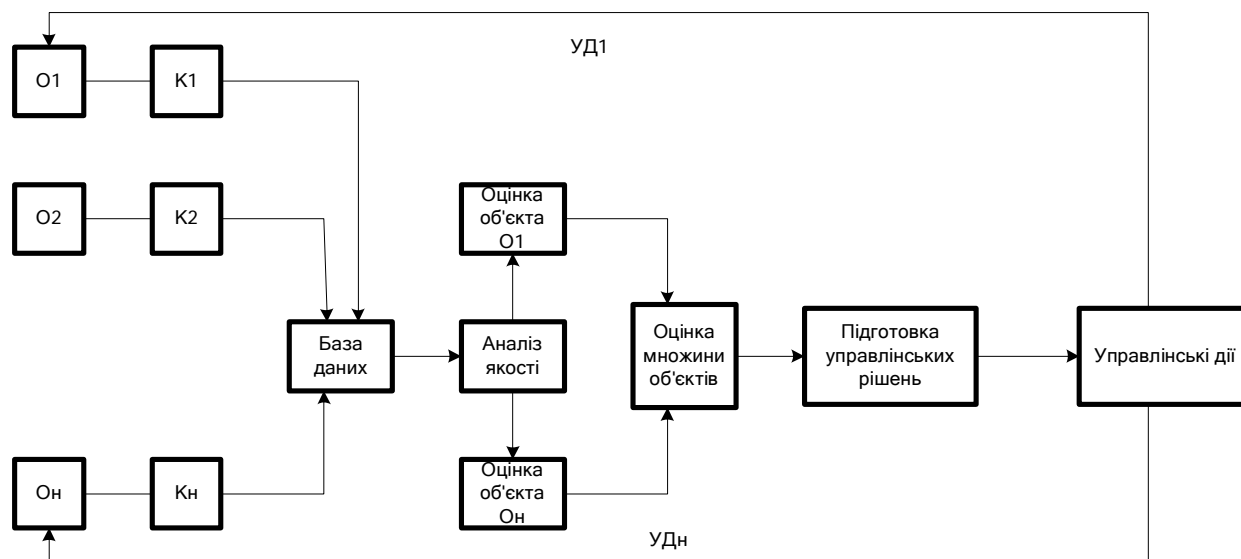


Рис. 2. Концептуальна схема управління якістю міського середовища

Розглянемо цю схему уважніше. Об'єкти міського середовища – $O_1 \dots O_n$ контролюються системою управління, результати контролю заносять до бази даних (БД) і відповідним чином групуються і аналізуються. Об'єднуємо об'єкти в множини, які відповідають крупнішим об'єктам чи системам. Значення відповідних показників усереднюються і заносяться в таблицю. Ця процедура повторюється для трьох складових міського середовища. Експертно визначаємо відносну вагу нормативного значення інтегрованого показника якості міського середовища. Це становить у нашому випадку: 40, 25, 35 балів для кожної складової відповідно, в сумі 100 балів. Результат контролю фактичних значень показників якості заноситься в базу даних, як це зображено в таблиці. Назву показників ототожнюємо з назвою цілей і підцілей. Це виправдано, оскільки показник якості повинен означати ступінь досягнення цілей і підцілей, і цей ступінь визначає експертна група. В результаті аналізу та оцінювання показників отримуємо таблицю, яку наведено нижче. На основі таблиці приймаємо відповідні рішення, скеровані на покращення показників міського середовища. Тобто, фактична експертна оцінка якості міського середовища становить: $29 + 21 + 25 = 75$ балів проти нормативної у 100 балів.

Результати оцінювання показників якості

Ціль	Норматив	Факт	Ціль	Норматив	Факт	Ціль	Норматив	Факт
1	40 балів	29	2	25 балів	21	3	35 балів	25
1.1.	15	10	2.1	05	04	3.1	04	04
1.2	15	12	2.2	05	04	3.2	04	03
1.3	05	03	2.3	04	03	3.3	05	03
1.4	05	04	2.4	04	3,5	3.4	05	04
			2.5	03	2,5	3.5	05	03
			2.6	04	04	3.6	04	03
						3.7	05	02
						3.8	03	03

Як бачимо з таблиці, фактичні значення ряду показників якості нижчі від нормативних значень. Це означає, що ці показники потрібно “підтягнути”, тобто виявити певні напрями діяльності системи управління містом щодо покращення стану міського середовища.

Запропонована система управління якістю міського середовища дає змогу скорочувати чи розширювати перелік показників якості залежно від можливостей або потреб управління. За одними показниками контроль може бути частіший, за іншими – рідший. Інформаційна система дає змогу накопичувати результати контролю за певний період часу і здійснювати ретроспективний аналіз якості міського середовища. Результати планування і контролю якості міського середовища можна оформити у вигляді паспортів якості і аналізувати стан міського середовища за різними критеріями.

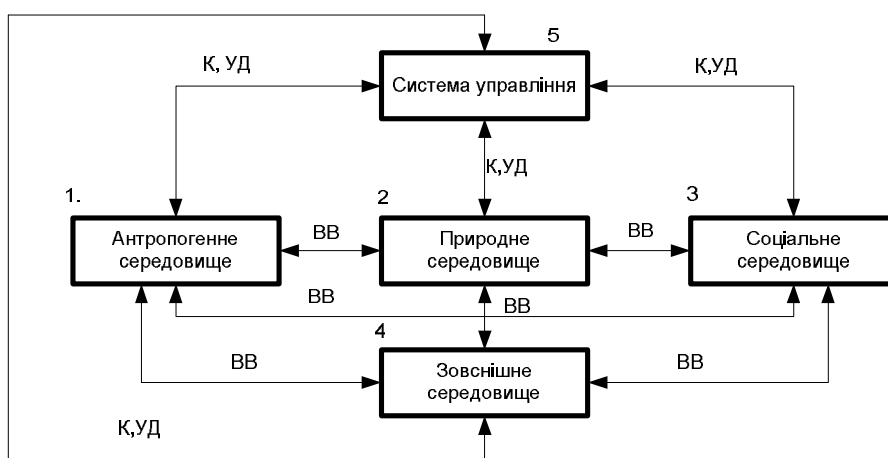


Рис. 3. Інформаційна модель системи управління якістю міського середовища

На рис. 3. наведено інформаційну модель управління якістю міського середовища, яка відображає всі інформаційні зв'язки між всіма складовими міського середовища та їх зв'язки з системою управління. На цій схемі позначки означають:

- ВВ – інформаційні взаємовпливи, які приводять до порушення рівноваги;
- К,УД – контроль, управлінські дії.

Система управління якістю міського середовища контролює всі складові цього середовища, а також контролює і певною мірою впливає на зовнішнє середовище. Зовнішнє середовище також впливає як на всі складові міського середовища, так і на систему управління. Це все система управління повинна відслідковувати і враховувати, готуючи управлінські рішення.

Отже, запропонована інформаційна система і система управління якістю міського середовища на її основі дадуть змогу систематично контролювати стан міського середовища, оцінювати його якість та якість системи управління містом і регулярно доводити інформацію до громади міста, зробить систему управління містом більш прозорою та ефективною.

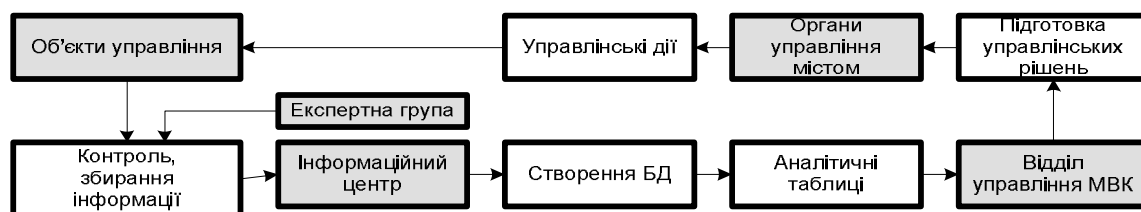


Рис. 4. Технологічна схема управління якістю міського середовища

На рис. 4. наведено технологічну схему управління якістю міського середовища. Об'єктами управління тут є об'єкти міського середовища. Контроль і збирання інформації з об'єктів здійснює експертна група, після чого вона надходить на обробку в інформаційний центр. Тут створюються відповідні бази даних і аналітичні таблиці. Таблиці передаються у відділ управління якістю міського середовища, де готують варіанти управлінських рішень і передають їх органам управління містом. Органи управління вибирають з набору управлінських рішень якнайкращі для цієї ситуації і формують відповідні управлінські дії щодо об'єктів управління. Потім технологічний процес управління повторюється.

Шляхи розвитку системи управління якістю міського середовища

Запропонований підхід до побудови системи управління якістю міського середовища дасть можливість кількісно оцінювати якість організації та функціонування міського середовища та якість управління ним. Для побудови такої системи необхідно створити:

- бази даних показників якості міського середовища ;
- групу експертів, яка систематично контролює стан міського середовища, розробити структуру і методику роботи експертної групи;
- відділ управління якістю міського середовища та відповідну організаційну і методичну документацію для цього відділу;
- динамічний сайт міськвиконкому, який допоможе залучати мешканців міста до управління містом та інформувати населення про розроблення та впровадження системи управління якістю міського середовища;
- інформаційний центр для збирання, обробки та зберігання інформації про міське середовище та забезпечення цією інформацією органи управління містом.

Для успішного функціонування системи управління якістю міського середовища потрібно багато різномірної інформації, значна частина якої сьогодні відсутня. Цю інформацію необхідно виявити і ввести в дію, щоб можна було розв'язувати задачі управління, які сьогодні не вирішуються. Потрібно зробити інформацію про результати діяльності доступною для всіх органів управління містом і для всіх мешканців міста.

Дуже важливо здійснити прив'язку інформації про міське середовище до просторових координат міста. Системі управління містом потрібна не загальна інформація, а конкретна: що діється або який стан міського середовища в конкретному місці. Проблеми, що виникають, розосереджені на території міста нерівномірно, а концентруються в певних місцях міста. Саме запропонована інформаційна система здатна швидко визначати ці проблемні місця.

Висновки

1. Здійснено аналіз літературних джерел та існуючого стану з розроблення та використання інформаційних систем для підтримки управління якістю міського середовища.

2. Розроблено концепцію управління якістю міського середовища з застосуванням системотехнічних методів проектування. Виявлені цілі створення системи управління якістю міського середовища, здійснено системний аналіз проблеми.

3. Розроблені інформаційна модель та технологічна схема управління якістю міського середовища.

4. На прикладі розглянуто процес оцінювання показників якості міського середовища.

5. Для вибору показників якості міського середовища запропоновано використання “дерева цілей” системи управління.

1. Огородник І. Особливості управління великим містом // Вісник Національного університету “Львівська політехніка”. – 2008. – № 629. – С. 137–145. 2. Шибakov В.Г. Город как сложная эколого – социально-экономическая система / В.Г. Шибakov, Л.В. Котляр, И.А. Шибакoва / Фундаментальные исследования. – 2004. – № 5. – С. 71–72. 3. Деякі аспекти вдосконалення управління великим містом: Мат. III Міжнар. конф. “Комп’ютерні науки та інформаційні технології” CSIT 2008. – 158–160 с. 4. Катренко А.В. Системний аналіз об’єктів і процесів комп’ютеризації: Навч. посібник. – Львів: Новий Світ, 2000. – 2000. – 424 с. 5. Бабаєв В.М. Підвищення ефективності управління великим містом у структурі державного управління: автореф. дис. ... канд. наук. Нац. академія державного управління при президентіві України. – К., 2004. 6. Новаковский Б.А., Прасолова А.И., Каргашин П.Е., Садов А.П. Принципы создания баз данных в медико-экологическом геоинформационном картографировании // Геоинформатика. – 2006. – № 1. – С. 16. 7. Огородник И.М., Евсеев А.С., Кесельман В.О. Модель планирования развития планировочной структуры города // Весник Львовского политехнического института “Резервы прогресса в архитектуре и строительстве”. – Львов, 1962. 8. Рудницкий А.М. Пути практической реализации программы комплексного управления городской средой // Весник Львовского политехнического института “Резервы прогресса в архитектуре и строительстве”. – Львов, 1982. 9. Перегудов Ф.И., Тарасенко Ф.П. Введение в системный анализ. – М.: Высшая шк., 1989. 10. Жариков О.Н. и др. Системный подход к управлению. – М.: ЮНИТИ, 2000. 11. Карий О. Комплексний розвиток міст: теорія та методологія стратегічного планування: монографія / О. Карий. – Львів: Видавництво Львівської політехніки, 2011. – 308 с.

МОДЕЛЬ ПРОЦЕСУ АДАПТАЦІЇ ЛЮДИНО-МАШИНИХ СИСТЕМ ДО КОРИСТУВАЧІВ

© Скринникова Г., 2014

Засоби управління курсором типу миша, клавіатура економічно вигідні для побудови адаптивних до психоемоційних станів користувачів людино-машинних систем. У роботі виявлено джерело розбіжності результатів при ідентифікації індивідуальних особливостей користувачів цими засобами – розмаїття поєднання факторів впливу. Розроблена модель процесу генерації динаміки маніпуляцій засобами управління курсором сприяла побудові моделі процесу адаптації людино-машинної системи до психоемоційних станів користувачів за динамікою та виявленню її властивостей.

Ключові слова: миша, клавіатура, людино-машинна система, адаптація.

Means of cursor control, such as a mouse, a keyboard, are the most economic advantageous for build human-computer systems adaptive to users. In this paper the cause of incorrect results in identify individual characteristics of users found: variety combination of affecting factors. The model of generating the dynamics of manipulations by means the cursor control developed. Its simplified variants helped build the model of the adaptation process of human-machine systems to the psycho-emotional states of users and detect its properties.

Key words: mouse, keyboard, human-computer systems, adaptive system.

Постановка проблеми

Сучасні людино-машинні системи (ЛМС) розробляють дружніми, але вони не передбачають реагування навіть на істотні зміни станів користувачів. Впровадження адаптивних до психоемоційних станів користувачів систем сприятиме розвитку емоційного інтелекту та підвищенню продуктивності праці.

Засоби управління курсором (ЗУК) є найдешевшими для побудови таких систем. Але основана виключно на динаміці маніпуляцій ЗУК (ДМЗУК) ідентифікація користувачів або їх індивідуальних особливостей не завжди дає високі результати, а адаптивних систем на основі ДМЗУК не розроблено.

Аналіз наукових досліджень і публікацій

У роботі [1] показано, що достовірність різниці середніх значень деяких ознак ДМЗУК у нейтральному стані та стані пониженої уваги і реакції на прикладі алкогольного сп'яніння становить 99 %. Та психоемоційні стани самі залежать від великої кількості факторів: втомленості, фізіологічного стану стресу, схильності до переживання певних емоцій, самоконтролю, ідеаторності емоцій, вживання напоїв, що містять спирт, кофеїн або ліків, гіперактивності щитовидної залози [2], близького оточення, основних подій, що відбуваються в світі тощо.

Графологи відрізняють почерки за “чоловічим” та “жіночим” типами, хоча Р. Пікард [3] не виявила гендерних або вікових суттєвих відмінностей. Графологи довели, що порушення функцій дрібної моторики при деяких захворюваннях спричиняє специфічні зміни в почерку. Однак не досліджувався вплив на ДМЗУК довжини пальців, тону м'язів й інших особливостей будови верхніх кінцівок тіла.

Імовірно, професія людини також впливає на ДМЗУК: конструктори, геймери роблять точніші та швидші рухи мишею, а секретарки швидше друкують. Зі зміною діяльності зміниться й

ДМЗУК. З часом вправність та індивідуальність маніпуляцій пристроями управління курсором зростає, а тривала відсутність щоденної практики спотворить ДМЗУК

Отже, серед факторів впливу на ДМЗУК виділено індивідуальні особливості користувачів.

У роботі [4] досліджено залежність індивідуальної динаміки управління мишею від технічних характеристик зовнішніх пристроїв ПК. Гіпотезу про випадкову залежність відкинута або її достовірність нижча за 90 % за акселерацією, швидкістю рухів, вагою, частотою опитування порту підключення миші, роздільною здатністю екрана монітора. О.М. Лепьошкін та А.В. Скубицький у [5] показали, що найшвидше 5 випробуваних працювали в середині робочого дня, трохи повільніше – вранці та набагато повільніше – увечері, проте це не доведено для інших користувачів. Не досліджувався вплив на ДМЗУК сезонних і тижневих коливань, травм рук, умов роботи: ергономічність робочого місця, відволікаючі фактори, пози тіла.

Унаслідок цього, серед факторів впливу на динаміку виділено фактори зовнішнього середовища.

Динаміку маніпуляцій мишею отримують за тривалістю часу від півсекунди до півгодини (з інтервалом від 10 мс та більше) або спираються на безперервну перевірку [6], не враховуючи пауз в роботі. П. Зіммерман обробив 67 ознак динаміки, але ані всі вони, ані зменшена їх кількість не дали значущих результатів [7]. Аналіз лише швидкості в [8], але у 8 напрямках, дав до 6 % помилок. Форми збирання даних (від ігор до анкет) можуть спотворювати ознаки внаслідок внесення зайвих емоцій, розгубленості, послаблення уваги. К. Епп для перевірки оцінювання емоцій використав необ'єктивний метод збирання даних – самозвіт [9]. Методи обробки даних: статистичні чи штучні нейронні мережі дають до 20% помилок. Головне обмеження другої групи методів – складність створення правильної навчальної вибірки.

Для ідентифікації використовують еталонний зразок. Але не вирішено питання про кількість часу, необхідну для отримання корисного еталонного зразку. Автор впевнений, що півгодини недостатньо, зважаючи на виявлені фактори, а місяць – більш імовірно. Крім того, дослідники ДМЗУК говорять про один еталонний зразок, а можливо, їх повинно бути декілька чи кожного дня новий. На думку автора, “свіжий” зразок, отриманий за короткий час, – кращий варіант.

Отже, виділено обмеження, встановлені розробниками систем, в яких враховують ДМЗУК.

На ДМЗУК впливають три групи факторів: 1) індивідуальні $f(t)$: e – психоемоційний стан; p – фізіологічний стан; l – рівень навичок; c – професійна спрямованість; 2) зовнішні $z(t)$: b – технічні характеристики маніпулятора, монітора, w – умови роботи, h – час доби, сезонні, тижневі коливання, 3) обмеження, встановлені розробниками $r(t)$: k – тривалість часу/довжина фрази, необхідні для ідентифікації; i – інформативні ознаки; n – часова/просторова сегментація; m – форма, методика збирання; q – обробки даних.

Мета статті – розробити модель процесу адаптації ЛМС до психоемоційних станів користувачів за ДМЗУК. Для досягнення поставленої мети слід вирішити такі завдання: 1) виявити фактори впливу на ДМЗУК; 2) розробити модель генерації ДМЗУК; 3) здійснити моделювання ДМЗУК.

Основні матеріали дослідження

Принципова модель процесу утворення ДМЗУК

Нехай $X = (x_1, \dots, x_N)$ – сукупність значень ознак ДМЗУК: x_1 – тривалість утримання клавіш, x_2 – паузи між натисненнями клавіш клавіатури; x_3 – поправка на відстань при досягненні вказівником миші цілі; x_4 – швидкість руху; x_5 – прискорення вказівника миші; x_6 – кривизна кривої, що описує вказівник при переміщенні миші. Тоді ДМЗУК – це множина $I: \{ X \cup MX_i \}$, де MX_i – статистичні характеристики x_i , $i=1, \dots, 6$.

Принципову модель утворення ДМЗУК наведено на рис. 1. Безпосередньо на ДМЗУК впливають перші дві групи факторів, а третя впливає при отриманні та обробці ознак. Вплив може бути систематичним, наприклад, добові коливання, або випадковим – відволікаючі фактори. Оскільки ці фактори можуть змінюватись раз за разом, то й ДМЗУК змінюється.

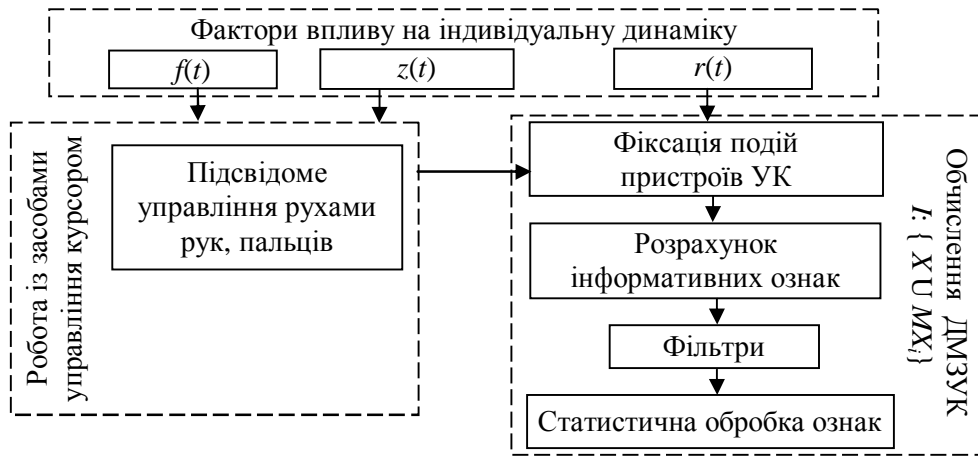


Рис. 1. Принципова модель процесу утворення ДМЗУК

Нехай в $f(t)$ $e=\{e1,e2,e3\}$ – психоемоційні стани, де $e1$ – стан зниженої реакції, афект, психоз і т.п., $e2$ – інтенсивна емоція, $e3$ – нейтральний стан; $l=\{l1,l2,l3\}$ – стан рівня навичок ($l1$ – високий, $l2$ – середній, $l3$ – низький); $c=\{c1,c2,c3\}$ – множина, що утворюється за професійною спрямованістю ($c1$ – конструктори, $c2$ – секретарі, програмісти й т.д.); $p=\{p1,p2\}$ – фізичний стан ($p1$ – користувач здоровий, $p2$ – навпаки); в $z(t)$ $h=\{h1,h2,\dots,h9\}$ – множина часу доби та тижня (сполучення: ранок, день, вечір, понеділок, середина тижня, днів перед вихідним); $b=\{b1,b2\}$ – стан технічних характеристик ЗУК ($b1$ – змінились або незручні користувачеві, $b2$ – навпаки); $w=\{w1, w2\}$ – множина умов роботи ($w1$ – відволікаючі фактори відсутні, $w2$ – навпаки). Стани, їх переходи у разі впливу на ознаки можна описати функціями, а потім скомпонувати за групами факторів. Наприклад,

$$w = \begin{cases} 1, & w1; \\ \frac{1}{b} e^{-|t-t_0|}, & w2, \end{cases} \quad (1)$$

де β – сила фактора; t_0 – час настання відволікаючого фактора. Та доведення впливу будь-якого фактору, його форми – скрупульозна робота. Імовірно, вплив $f(t)$ та $z(t)$ на різні ознаки різний, оскільки, наприклад, ознака x_2 більше залежить від навичок роботи, а x_1 – від типу нервової системи.

Отже, виявлено фактори впливу на ДМЗУК. Та на кожного користувача цей вплив може бути індивідуальним, як і фільтри для факторів впливу (рис.1). Розроблена принципова модель процесу утворення ДМЗУК з урахуванням цих факторів показує, на якому етапі відбувається вплив. Фільтри або поправкові коефіцієнти на виявлені фактори підвищують точність систем ідентифікації ЗУК.

Модель процесу генерації ДМЗУК в термінах теорії автоматичного керування

На рис.2 представлено розроблену модель процесу генерації користувачем ДМЗУК в термінах теорії автоматичного керування. СУ – суб'єкт управління (людина); ПОУ – проміжний об'єкт управління (миша, клавіатура); ОУ – курсор ЗУК на моніторі ПК чи ін. цифрового пристрою, який здійснює процес квантування та дає $x(t)$; $f(t)$ – задаючий вплив; $z(t)$ – збурюючий вплив, в тому числі сигнальна інформація з робочого місця; $g(t)$ – управляючий вплив: аналогові сигнали, що посилає підсвідомість людини руці щодо маніпуляцій ПОУ; $x_c(t)$ – події миші/клавіатури, які оброблено ПК з ОУ; $x(t)$ – дискретний сигнал з виходу ОУ – значення ознак динаміки маніпуляцій ЗУК.

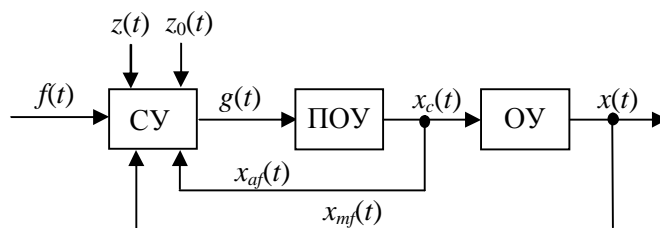


Рис. 2. Модель процесу генерації користувачем ДМЗУК

Опис моделі керування користувачем своєю ДМЗУК: внутрішній посил доброї індивідуальної волі $f(t)$ користувача (СУ), який знаходиться під впливом зовнішніх факторів $z(t)$, реалізується управлінням низкою підсвідомих рухів, що сформувались підсвідомістю СУ, за сигналами $g(t)$ по досягненню поставлених цілей за допомогою ЗУК (ПОУ). Відчуваючи мишу/клавіатуру, дивлячись на набрані символи, керуючи курсором, людина може здійснювати комбіноване управління: регулювати технічні характеристики миші, пити збуджуючі напої, змінювати позу тощо (управління за збурюванням) або виправляти помилки, проявляти більше внутрішнього старання (управління за відхиленням). На рис.2 $z_0(t)$ – контрольований користувачем збурюючий вплив, $x_{mf}(t)$ – головний зворотний зв'язок, $x_{af}(t)$ – додатковий зворотний зв'язок.

Для клавіатурного почерку $x_c(t)$ – це значення часу настання подій клавіатури: ступінчата неперіодична функція; для динаміки управління мишею – це значення часу положення миші або настання подій: ступінчата функція в системі (X, Y, t) . Вирази переходу $x_c(t)$ в $x(t)$ відомі [1].

Представимо $g(t)$ сигналом управління “природними” синергіями, які визначаються власними векторами лінеаризованих рівнянь Лагранжа за моделлю динаміки руки людини. Для управління надлишковим апаратом м'язів і суглобів у нервовій системі формуються рухові синергії керуючих команд [10], які проявляються в узгоджених змінах суглобових кутів (кінематичні синергії) і суглобових моментів (динамічні синергії). Зв'язок між кінематичними і динамічними синергіями визначається складною динамічною взаємодією ланок біомеханічного ланцюга. Рух кожного сегмента впливає на рух всіх інших, а активація будь-якого м'яза втягує в рух всі сегменти, а не тільки ті, до яких цей м'яз прикріплений.

У [11] показано, що суглоб можна моделювати в'язко-еластичною пружиною з часовою затримкою 50–100мс. У моделі з затримкою діючі моменти сил в суглобах описуються рівнянням лінійного “ПД-збурення”:

$$\dot{I}_l(t) = \dot{I}_l(t-\tau) + S_{lm}[(h_d^m(t) - h^m(t-\tau))] + V_{lm}[(\dot{h}_d^m(t) - \dot{h}^m(t-\tau))], \quad (2)$$

де S_{lm} – матриця жорсткості; V_{lm} – матриця в'язкості; $h^m(t)$ і $\dot{h}^m(t)$ – суглобові кути і кутові швидкості в момент t ; $h_d^m(t)$ і $\dot{h}_d^m(t)$ – бажані значення суглобового кута і кутової швидкості; τ – затримка за часом. Суглобові кути $h^m(t)$, кутові швидкості $\dot{h}^m(t)$ та кутові прискорення $\ddot{h}^m(t)$ обчислюються за даними реєстрації руху. Зовнішні сили вважаються заданими. Маса і моменти інерції ланок, які входять у вирази для метричного тензора і символу Крістоффеля, визначаються для кожної людини індивідуально за допомогою антропометричних таблиць. В'язко-пружні властивості суглобів – матриці жорсткості та в'язкості, S_{lm} і V_{lm} , визначаються за допомогою лінійної регресійної моделі, застосованої для рівняння (2). Цим способом на основі зареєстрованої кінематики рухів розраховувалися в'язко-пружні властивості ліктьового суглоба при несподіваному та керованому розвантаженні передпліччя, суглобів дволанкової моделі руки при її цілеспрямованих рухах [12]. Однак, крім індивідуального підходу, складно знайти оператора, який такі складні управляючі команди перетворить на вихідний сигнал $x_c(t)$.

Важко навіть за запропонованим спрощеним (рис.3) поданням входу $f(t)$ знайти оператор перетворення сигналів на СУ. Моделі динаміки сприйняття зовнішньої сигнальної інформації через слух, зір, нюх, дотик і виникнення емоцій від таких впливів, впливу емоцій на кінематику рухів тіла, опису системних механізмів мозку (механізмів свідомості, мислення) в основному представлені у виді словесних або схемних рішень. Інші рішення базуються на даних реєстрації біоелектричної активності ділянок кори мозку. Наприклад, моделювання реакцій головного мозку людини на зовнішні подразники вперше описано в [13] за допомогою системи стохастичних диференціальних рівнянь, однак модель побудовано за даними з ЕЕГ. І штучні нейронні мережі, що використовуються в сфері штучного інтелекту як спрощені моделі нейронних мереж в головному мозку, не відображають реальної його структури.

Імітаційне моделювання процесу генерації ДМЗУК

Оскільки характер процесів, що проходять у системі (рис. 2), не дає змоги описати ці процеси в аналітичній формі, застосуємо імітаційне моделювання. Спростимо модель з рис. 2, подаючи на вхід вибірку з стандартного безперервного рівномірного розподілу (рис.3). ОУ – це курсор ЗУК.

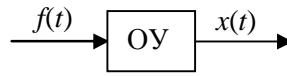


Рис. 3. Спрощена модель процесу генерації ДМЗУК

Для застосування імітаційного моделювання встановимо закони розподілу x_1, \dots, x_6 . Вживши метод кількісного групування до кожної вибірки x_j з 10 сукупностей значень ознак X , узятих у трьох користувачів, а для перевірки згоди дослідного розподілу з теоретичним непараметричний критерій Колмогорова–Смирнова, отримаємо: 1) при обсязі вибірки $n=42$ і при підборі параметрів розподіл x_1 з імовірністю не меншою за 95 % тотожний розподілу Вейбулла; 2) розподіл x_2 вже при $n=21$ з імовірністю не меншою за 95 % тотожний розподілу Вейбулла; 3) розподіл x_4 при $n=63$ з імовірністю не меншою за 90 % тотожний експоненціальному розподілу; 4) розподіл x_5 при $n=63$ з імовірністю не меншою за 90 % тотожний логістичному розподілу; 5) розподіл x_6 при $n=63$ з імовірністю не меншою за 95 % тотожний нормальному розподілу.

Однак, імітаційне моделювання дає задовільний результат лише “за середнім”. Наприклад, при генерації значень x_4 методом зворотного перетворення функція експоненціального розподілу $F(x) = 1 - e^{-I(x-x_0)}$ строго зростає, її зворотна функція набуває вигляду $F^{-1}(x) = x_0 - \frac{1}{I} \ln(1-x)$. Якщо U_1, \dots, U_n – вибірка зі стандартного безперервного рівномірного розподілу, то $\hat{x}_1, \dots, \hat{x}_n$, де $\hat{x}_i = x_0 - \frac{1}{I} \ln(1 - U_i)$, $i = 1, \dots, n$ – шукана вибірка з експоненціального розподілу. Та перевірка згоди дослідного розподілу з розподілом, отриманим числовим моделюванням, за Колмогоровим–Смирновим на 70 сукупностях значень ознак X за x_4 в нейтральному стані дала: імовірність тотожності не менша за 90 % – у 18 % випадків, імовірність тотожності від 90 до 5 % – у 65 % випадків, імовірність відхилити згоду – у 17 % випадків.

Моделювання ДМЗУК методом аналізу електричних ланцюгів

Розглянемо інший варіант. Подаватимемо на вхід (рис. 3) $f(t)$ вольовий посил, розкладаючи його на складові за допомогою одиничної функції Хевісайда $1(t)$. ОУ – це рука+ЗУК+курсор.

Оскільки воля – це психічний процес, а основними елементами нервової системи є нейрони, які проводять біоелектричні імпульси, для моделювання скористаємося методом аналізу електричних ланцюгів. Основне припущення – ОУ має властивість лінійності, як в теорії електричних ланцюгів.

У [14] показано, що сила F кожного мотиву визначається виразом:

$$F = pRDGu^{1-\frac{t}{\tau_1}}, \quad (3)$$

де p – постійний коефіцієнт; R – значущість цілі для людини; D – сила бажання досягти успіхів у різних видах діяльності; $Gu = I_B / I_C$ – число, що характеризує інформаційний стан мотиву в момент його появи (I_C – інформація про засоби, необхідні для досягнення мети (прогностична інформація); I_B – інформація про засоби, які суб’єкт має в певний момент часу (прагматична інформація)), $0 < Gu < 1$; τ – час; τ_1 – час, необхідний для досягнення мети.

Взявши в (3) за $a = pRD$, $b = Gu$, отримаємо:

$$F = ab^{1-\frac{t}{\tau_1}}. \quad (4)$$

Якщо на ОУ в момент $t = 0$ надати вплив F , то на ОУ F дорівнюватиме $F \cdot x(t)$. Якщо ту ж F прикласти не в момент $t = 0$, а в момент τ , то на ОУ F дорівнюватиме $F \cdot x(t-\tau)$. Задане безперервною функцією F можна подати у вигляді:

$$f(t) = f(0) \cdot 1(t) + \Delta f_1 \cdot 1(t - \Delta t) + \Delta f_2 \cdot 1(t - 2\Delta t) + \dots + \Delta f_n \cdot 1(t - n\Delta t). \quad (5)$$

Якщо тестуючий вплив подати у вигляді послідовності функцій $1(t)$, то на ОУ він визначається як послідовність реакцій ОУ на послідовність функцій Хевісайда та набуде вигляду:

$$u(t) = f(0) \cdot x(t) + \Delta f_1 \cdot x(t - \Delta t) + \Delta f_2 \cdot x(t - 2\Delta t) + \dots + \Delta f_n \cdot x(t - n\Delta t). \quad (6)$$

Якщо безмежно збільшувати число n , то границею ламаної (5) буде крива (4). Помноживши і розділивши обидві частини рівнянь (5) і (6) на прирощення за часом Δt , отримаємо:

$$f(t) = f(0) \cdot 1(t) + \sum_{i=1}^n \frac{\Delta f_i}{\Delta t} 1(t - i\Delta t) \cdot \Delta t, \quad (7)$$

$$u(t) = f(0) \cdot x(t) + \sum_{i=1}^n \frac{\Delta f_i}{\Delta t} x(t - i\Delta t) \cdot \Delta t \quad (8)$$

При $\Delta t \rightarrow 0$ перейдемо від кінцевого кроку відліку Δt до $d\tau \rightarrow 0$. Позначивши $n \cdot \Delta t = \tau$, отримаємо:

$$u(t) = f(0) \cdot x(t) + \int_0^t x(t - \tau) \cdot f'(\tau) d\tau. \quad (9)$$

Оскільки вид функції $f(t)$ відомий, щоб знайти $u(t)$, як перехідну функцію $x(t)$ візьмемо закон розподілу відповідної ознаки ДМЗУК. Підінтегральний вираз (9) при заміні $\tau = t - \theta$ має вигляд:

$$-x(q) \cdot f'(t - q) dq. \quad (10)$$

За зворотної заміни отримуємо рівність, еквівалентну (9):

$$u(t) = f(0) \cdot x(t) + \int_0^t x(t) \cdot f'(t - \tau) d\tau. \quad (11)$$

Інтегруючи за частинами (9), отримаємо:

$$f(0) \cdot x(t) + \int_0^t x(t - \tau) \cdot f'(\tau) d\tau = f(t) \cdot x(0) + \int_0^t f(\tau) \cdot x'(t - \tau) d\tau, \text{ отже}$$

$$u(t) = f(t) \cdot x(0) + \int_0^t f(\tau) \cdot x'(t - \tau) d\tau. \quad (12)$$

Інтеграл виразу (12) при заміні $\tau = t - \theta$ має вигляд

$$- \int_t^0 f(t - q) \cdot x'(q) dq. \quad (13)$$

Застосовуючи до (13) формулу диференціювання під знаком інтеграла для випадку, коли межі інтегрування змінні:

$$u(t) = \frac{d}{dt} \int_0^t f(t - q) \cdot x(q) dq \quad (14)$$

та підстановку $\tau = t - \theta$, маємо: $u(t) = \frac{d}{dt} \int_0^t f(\tau) \cdot x(t - \tau) d\tau.$

Пропускаючи через ОУ збурення F вигляду (4), згідно з (14) отримаємо:

$$u(t) = \frac{d}{dt} \left(ab^{\frac{1-t}{t_1}} \cdot \int_0^t ab^{\frac{1+\tau}{t_1}} \cdot x(\tau) d\tau \right). \quad (15)$$

Знаходимо похідну добутку (15):

$$u(t) = -\frac{\ln a}{t_1} \cdot ab^{1-\frac{t}{t_1}} \cdot \int_0^t ab^{\frac{t}{t_1}} \cdot x(t) dt + ab^{1-\frac{t}{t_1}} \cdot \frac{d}{dt} \left(\int_0^t ab^{\frac{t}{t_1}} \cdot x(t) dt \right). \quad (16)$$

Застосовуючи до (16) диференціювання для випадку, коли межі інтегрування змінні, отримуємо $\frac{d}{dt} \int_0^t ab^{\frac{t}{t_1}} \cdot x(t) dt = ab^{\frac{t}{t_1}} x(t)$ і

$$u(t) = \frac{-1}{t_1} \ln a \cdot a^2 b^{1-\frac{t}{t_1}} \cdot \int_0^t b^{\frac{t}{t_1}} \cdot x(t) dt + a^2 b^2 x(t). \quad (17)$$

Функція $u(t)$, що визначає стан F на ОУ, має ту саму форму, що й задаючий вплив $f(t)$. Наприклад, якщо задати в (2) $p = 1, R = 1, D = 1, Gu = 0.2, t_1 = 1$, отримаємо $u(t) = 0.04x(t)$.

Чим менше Gu , тим швидше на виході досягається стає значення. Чим більше R і D , тим повільніше на виході досягається стає значення. Тобто згідно з (3) занадто висока значимість мети для людини і сила бажання досягти успіху, велика кількість прагматичної інформації заважають стабільності керування ЗУК. А психологи наполягають, що чим вищий емоційний інтелект людини, тим швидше вона вирішить завдання, пов'язані з мотивацією. Отже, якщо підтримувати або регулювати певні емоції на робочому місці, то можна домогтися швидшого виконання поставлених перед працівником задач.

Зауважимо, що при управлінні користувачем ЗУК цілей завжди досягають (якщо тільки не діяли суттєві відволікаючі фактори або користувач у стані зниженої реакції), хоча, наприклад, курсор миші рухається нелінійно, безперервно і за різною траєкторією при виконанні одних і тих самих завдань. Розглядаючи управління курсором як завдання управління для динамічної системи, можна стверджувати, що така система стійка, оскільки цілі досягнуто.

Модель процесу адаптації людино-машинної системи до психоемоційних станів

Нехай є деякий еталонний зразок ДМЗУК $x_e(t)$. Свавілья обмежень, встановлених дослідником, важко уникнути. Та вважатимемо, що такі похибки зведені до мінімуму і можна їх не враховувати.

У [16] виявлено залежності в середньому значень ознак ДМЗУК від впливу емоції радості. При створенні ЛМС, адаптивних до психоемоційних станів користувачів на основі ДМЗУК (рис.4), за Dx можна визначати стан користувача та подавати відповідний аудіовізуальний вплив $z_p(t)$ для зниження рівня високоінтенсивних емоцій, наприклад, впливати на технічні характеристики ЗУК $z_2(t)$. Однак невідомо, коли очікувати працюючу модель мозку СУ. Якщо ж спростити модель рис.4 (див. рис. 5): за Woc взяти (17), на вхід якої подається задаючий вплив $x_e(t)$, – то модель буде доречна лише у випадку автоматичного керування ЗУК.

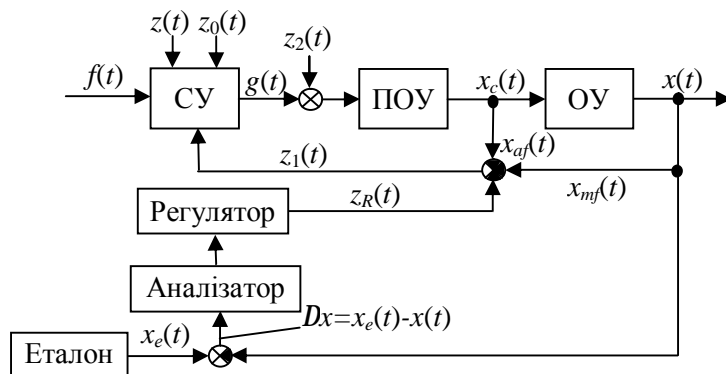


Рис. 4. Модель автоматизованого процесу адаптації ЛМС до психоемоційних особливостей СУ

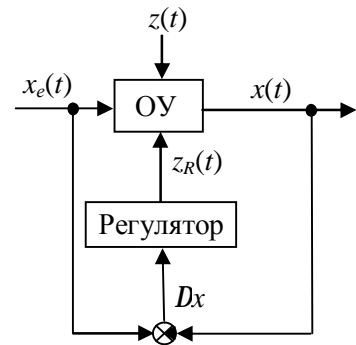


Рис.5. Модель автоматичного управління ЗУК

Нехай відомі Woc – передатна функція ОУ; $Wioc$ – передатна функція ПОУ; Wsc – передатна функція СУ; W_R – передатна функція регулятора, тоді стан системи на рис. 4 описується такими рівняннями: $g = Wsc z f z_0 z_1$, $x_c = Wioc (g+z_2)$, $x = Woc x_c$, $z_R = W_R Dx$, $Dx = x_e - x$, $z_1 = -z_R - x_{mf} - x_{af}$. Вирішивши їх відносно x , маємо:

$$x = \frac{Woc \cdot Wioc \cdot z_2 - Woc \cdot Wioc \cdot Wsc \cdot z_0 \cdot z \cdot f \cdot (W_R \cdot x_e + x_{mf} + x_{af})}{1 - W_R \cdot Woc \cdot Wioc \cdot Wsc \cdot z_0 \cdot z \cdot f}. \quad (18)$$

Якщо при $z_2=0$ будь-яке з Wsc , Woc , $Wioc$, z , f чи $z_0 \rightarrow \infty$, то властивості ланки визначаються тільки властивостями ланцюга зворотного зв'язку та $x = x_e + \frac{x_{mf} + x_{af}}{W_R}$. Якщо при $z_2=0$ СУ не забезпечує $z_0 (z_0=0)$, то $x = 0$; якщо $W_R = 0$, то $x = -x_{mf} - x_{af}$; якщо $x_{mf} = 0$ та $x_{af} = 0$ чи $x_{mf} + x_{af} = 0$, а $z_0 \neq 0$, то $x = x_e - \frac{x_e}{1 - W_R \cdot Woc \cdot Wioc \cdot Wsc \cdot z_0 \cdot z \cdot f}$; якщо забезпечити високе посилення в ланцюгу зворотного зв'язку $W_R \rightarrow \infty$, то $x = x_e$. Однак зробити $W_R \rightarrow \infty$ неможливо. Провівши ряд аналітичних перетворень при $z_2 = 0$, знаходимо, що $x = x_e$, коли

$$x = -(x_{mf} + x_{af}) Woc Wioc Wsc z f z_0. \quad (19)$$

Та при отриманні x і x_e вплив $f(t)$, x_{mf} та ін. (рис. 4) буде різним. Отже, кожен зразок ДМЗУК унікальний.

Підібрати зворотний аудіовізуальний зв'язок, загальний для всіх користувачів за змістом та силою, складно, вважаючи на різні культурні, релігійні, соціальні, вікові характеристики. Крім того, необхідно забезпечити його різноманітність, ненав'язливість, якість. У результаті це буде багатогодинний вміст з багатьох короткометражних сюжетів, що потребуватиме ресурсів ПК.

Оскільки ідентифікація станів за ДМЗУК дає невисокі результати [1,7,9] внаслідок як складності формування емоцій, так і різноманітного сполучення факторів впливу на ДМЗУК, а повнота відповідальності за можливий моральний та непрямий матеріальний збиток, зокрема збиток для соматичного та нервово-психічного здоров'я при неправильній адаптації до можливо невірної ідентифікованих психоемоційних станів лежить на розробникові – вживання додаткових сенсорів тиску на клавіші, відеокамер й т.п. підвищить результативність ідентифікації станів, але й обмежить масовість застосування запропонованої моделі процесу адаптації.

Висновки

Поставлені завдання дослідження виконано, мети досягнуто.

1. На основі виявлених факторів впливу на ДМЗУК розроблено модель процесу генерації ДМЗУК – аналогово-дискретну систему. Результати ранніх експериментів з ідентифікації користувачів за ДМЗУК було отримано без урахування цих факторів, і тому вони неточні. Поправкові коефіцієнти або фільтри для виявлених факторів підвищують точність систем ідентифікації ЗУК.

2. Моделювання процесу генерації ДМЗУК методами теорії електричних ланцюгів показало, що надто висока значимість мети і сила бажання досягти успіху заважають стабільності керування ЗУК. Впровадження адаптивних до певних станів користувачів ЛМС підвищить продуктивність праці.

3. Аналіз розробленої моделі процесу адаптації ЛМС до користувачів на основі ДМЗУК показав: стабільність ДМЗУК забезпечить виконання виразу (19). Однак ДМЗУК залежить від впливу багатьох факторів, і складові (19) не є сталими, а зворотний зв'язок потребує індивідуального підходу, додаткових досліджень та витрат. Критерії ефективності моделі можна сформулювати за результатами аналізу ДМЗУК.

1. Скринникова Г.В. Дослідження впливу станів зниженої реакції та уваги на динаміку маніпуляцій пристроями управління курсором / Г.В. Скринникова // *Кіберагресія: психолого-педагогічні та кібернетичні проблеми безпеки: Матер. Всеукр. НПК. Луганськ, 12-14 лист. 2012* – С. 46–54. 2. *Mouse Movements Biometric Identification: A Feasibility Study/ Weiss A., Ramapanicker A., Shah P.[at al]// Proceedings of CSIS, Pace University, May 4th, 2007.* – pp. 21–28. 3. *Picard R.W. Emotion research by the people, for the people. 2010* – URL: <http://affect.media.mit.edu/pdfs/10.Picard-ER->

revised.pdf (дата звернення 17.11.2011). 4. Скринникова Г.В. Дослідження впливу технічних характеристик зовнішніх пристроїв персональних комп'ютерів на індивідуальну динаміку маніпуляцій мишею / Г.В. Скринникова // Інформаційна безпека. – Луганськ: СХУ ім. В. Даля, 2012. – № 2(8). – С. 144–150. 5. Лепёшкин О.М. Разработка подхода к распознаванию биометрического портрета пользователя по клавиатурному почерку на основе методов нелинейной динамики / О.М. Лепёшкин, А.В. Скубицкий // Информационное противодействие угрозам терроризма. – М.: ФГПУ НТЦ, 2008. – № 11. – С.102–112. 6. Nazar A., Traore I., Ahmed A.A.E. Inverse biometrics for mouse dynamics// *International Journal of Pattern Recognition and Artificial Intelligence*. – Vol. 22, № 3. – 2008. – pp. 461–495. 7. Zimmermann P.G. Beyond usability – measuring aspects of user experience. *Dr. Sciences. Swiss federal institute of technology, Zurich, 2008.* – 112 p. 8. Singh S. Mouse interaction based authentication system by classifying the distance travelled by the mouse/S.Singh, Dr.K.V.Arya// *Int. Journ. of Computer Applications Vol.17, No.1, 2011.* – pp.45-48. 9. Epp C. Identifying emotional states using keystroke dynamics/ C. Epp, M. Lippold, R.L. Mandryk // *CHI 2011, May 7-12, 2011, Vancouver, BC, Canada.* – pp. 715-724. 10. Бернштейн Н. А. Физиология движений и активность [Текст]/ Под ред. О. Г. Газенко. – М. : Наука, 1990. – 496 с. 11. Frolov A.A. On the possibility of linear modeling of the human arm neuromuscular apparatus / Frolov A.A., Dufosse M., Rizek S., Kaladjan A. // *Biological Cybernetics Vol 82 (6), 2000.* – pp. 499-515. 12. Frolov A. A. Adjustment of the human arm viscoelastic properties to the direction of reaching/ Frolov A. A., Prokopenko R. A., Dufosse M., Ouezdou F. B. // *Biological Cybernetics. Vol 94, 2006.* – pp. 97-109. 13. Золкин С.Г. Уравнения динамического стереотипа в прогнозе реакции неокортекса человека/ С.Г. Золкин // “Штучний інтелект” 2005. – № 2. – С. 39–44. 14. Глазунов Ю.Т. Роль и значение воли в процессах целеполагания/ Ю.Т. Глазунов // *Вестник МГТУ, 2013.* – Т. 16, № 2. – С. 279–287.

АНАЛІЗ СУЧАСНИХ МЕТОДІВ ТА ВИДІВ ГРАФІЧНОГО ЗАХИСТУ ДРУКОВАНИХ ДОКУМЕНТІВ

Ї Назаркевич М., Троян О., 2014

Проаналізовано методи захисту друкованих документів та показано, що латентні зображення залишаються перспективними для захисту. Розроблено мікрографічний спосіб захисту з товщиною ліній від 60 мкм. Побудовано гільйошні елементи на основі збурення ліній за технологією pdf. Сформовано графічні пастки, які містять фрагменти з повторюваних штрихів із заданою товщиною і періодом лінії. Створено латентні елементи шляхом побудови ліній різної товщини.

Ключові слова: спосіб захисту, цінні папери, захисні елементи.

The methods of protection printed documents and it is shown that latent images are promising in defense. A micrographic method of protection with a thickness of 60 micron lines. Powered guilloche items based on perturbation line technology pdf. Graphic traps are formed that contain fragments of repetitive strokes with a given thickness and period line. A latent elements are building lines of varying thickness.

Key words: method of protection, securities, security printing

Вступ

Сучасні інформаційні технології, які формують захищені документи, розвиваються швидкими темпами. Є потреба створювати нові види захисту друкованих документів, оскільки засоби та методи фальсифікації стають все поширенішими. Сьогодні фальсифіковані документи створюються новими більш технологічними методами, які максимально близькі до методів виготовлення оригіналу. З кожним роком технічні характеристики копіювальних апаратів стають досконалішими, тому потрібно розробляти нові методи захисту друкованих документів. Одним з ефективних, економічних та надійних способів захисту є графічний захист документів.

Обстежено об'єкт інформаційної діяльності друкованого документа (див. рис. 1) щодо видів захисту. Загальноприйнято, що захищений документ буде створений при:

- 1) використанні нових технологій створення паперу (водяні знаки, захисні волокна, конфетті, полімерної стрічки, хімічного захисту);
- 2) використанні особливих технології друку (високий друк, глибокий друк, ірисовий друк, орловський друк та ін.);
- 3) захисті на основі особливих фарб (фарби “овіай”, флюоресцентні фарби та інші);
- 4) захисті на етапі додрукарської підготовки (гільйошні елементи, гравюри, мікротекст, латентні зображення, мікроплекс, мультиплекс, графічні кодовані зображення);
- 5) післядрукарських захистах (перфорація, введення мікрочіпів, біометричні елементи).

Графічний захист друкованого документу можна виконати позитивним та негативним способом. Товщина лінії для позитивного виконання ліній 40–80 мкм, а для негативного – 60–100 мкм. Колір та відтінок ліній підбирають так, щоб при копіюванні та скануванні лінії не відтворювалися.

Точність відтворення елементів захисту задовольняє технологічні умови друкування документа за роздільної здатності 2400 x 2400 dpi (точок на дюйм) та вищої. Ще одним показником виведення є lpi (кількість ліній на дюйм), який може сягати 100–300 lpi, що відповідає найсучаснішим вимогам.

Особливої уваги заслуговують латентні зображення, які належать до графічних видів захисту. Латентні зображення мають властивість приховувати зображення при зміні умов спостереження. Латентні зображення можна створювати різними способами: за допомогою засобів голографії, з використанням явища поляризації, застосовуючи спеціальні фарби і покриття, за певними методами формування елементів зображення.

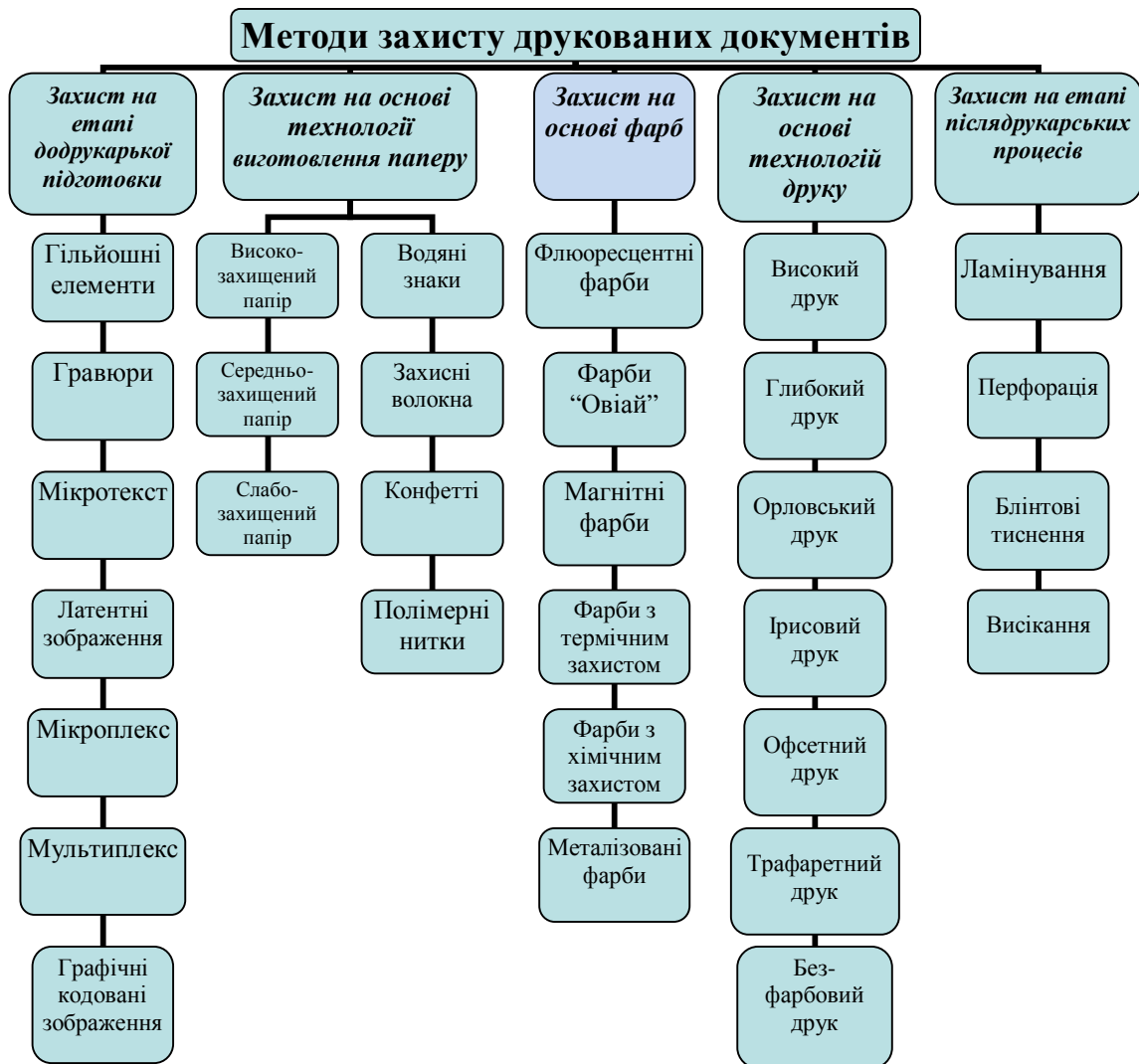


Рис. 1. Аналіз методів захисту друкованих документів

Високий ступінь захисту поліграфічної продукції визначається трьома складовими:

- складністю технологічних процесів;
- обмеженням доступу до матеріалів та обладнання;
- новизною і закритістю методів, що застосовуються.

Розробляти і застосовувати складні технології, наприклад, металографічного, глибокого друку, використовувати спеціальні фарби і покриття, захисні ламінати, голограми не завжди доцільно з економічного погляду. Застосування матеріалів зі спеціальними хімічними, фізичними властивостями створює необхідність застосування спеціального обладнання для визначення автентичності продукції.

Найпростішими, зручними і економічно доцільними є засоби захисту, що реалізуються на додрукарському етапі. Їх застосування не вимагає спеціального устаткування і матеріалів. При цьому залишається можливість їх використання для створення комбінованих засобів захисту. Розроблення нових та вдосконалення існуючих методів захисту друкованих документів має велике значення для безпеки та захисту друкованих документів.

Формування захисту на основі мікрографіки

Захист з використанням мікрографіки ґрунтується на створенні тонких графічних елементів: гільйошів, сіток, розеток, віньєток, прихованих об'єктів і мікрографіки. Відтворення тонких ліній можливе лише при використанні технології друкарських процесів. Поліграфічний захист вважається ефективним, якщо на документі не менше 70 % площі займає мікрографіка. Складність репродукції пов'язана зі складною геометричною структурою та мінімально можливою товщиною ліній елементів, яку неможливо відтворити репрографією.

Захист друкованих та електронних документів оснований на високому ступені надійності та здатності захистити від пошкодження та фальсифікації друковані та електронні документи. За цим способом захисту утворюють графічні елементи, які формуються як масив точок перетину вхідного зображення з лініями графіки, до утворення у зоні заповнення кольором.

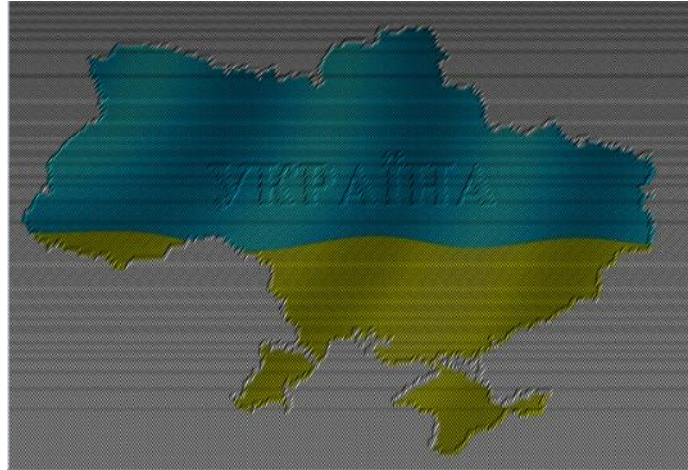


Рис. 2. Захисне зображення карти України методом заповнення кольору

Формування захисту на основі гільйошних елементів

При виготовленні друкованого документа з захистом використовують гільйошні елементи, які дають змогу підвищити ступінь захисту. За цим способом встановити достовірність документа можна, створюючи тонкі лінії, які є неперервними та не перетинаються між собою на всій площині документа. За спроби фальсифікації документ сканують та оцифрують за допомогою комп'ютерної техніки, перетворюють на масив пікселів, що не дає можливість відтворити гільйош. Зображення на фальсифікованому документі буде сірим, розмитим, переривчастим. Візуально легко відрізнити фальсифікацію від оригіналу [1]. На рис. 3 показано формування захисту на основі гільйошів.



Рис. 3. Формування зображення на основі збурення ліній

Побудовано векторне зображення, у якому контур формується на основі збурення ліній. Алгоритм реалізовано так. У пікселях, де накладається контур зображення з кривою, створюється збурення лінії шляхом зміщення по контуру. Одинична лінія створюється графічними примітивами. Збурення лінії можливе при використанні будь-якого з графічних примітивів [3].

Метод побудови гільйошних елементів має унікальний вид кривих та побудову ліній за заданими математичним апаратом. При виведенні документа реалізовано формат PDF, що забезпечує високу якість друку. Запропонований метод можна використати для захисту етикеток, бланків суворої звітності, акцизних марок тощо.

Формування захисту на основі графічних пасток

Одним з способів захисту документів є створення графічних пасток, у яких порушується існуюче зображення або текст документа. Створюють графічні пастки, умисно спотворюючи лінії. Для побудови графічної пастки використовують різні хитрощі: малопомітні розриви в графічних орнаментах; навмисне порушення локальної симетрії при відтворенні одного з декількох повторюваних елементів орнаменту; застосування в текстових реквізитах одиничних знаків, що відрізняються від інших за розміром, шрифтом або нахилом і т. п. Крім того, можна вводити секретні орнаменти та інші складні малюнки, фрагменти з повторюваних штрихів з заданими товщиною і періодом (рис. 4).

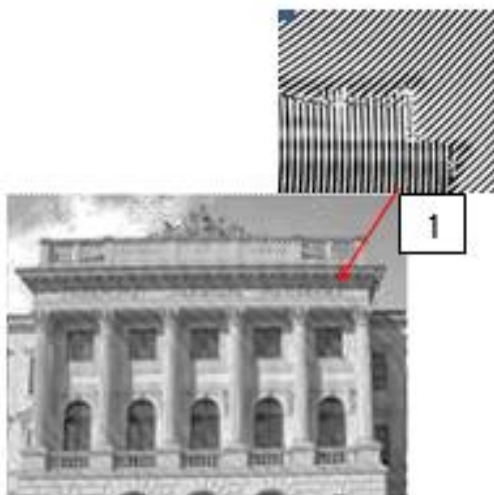


Рис. 4. Формування зображення на основі графічних пасток

Формування захисту на основі прихованих елементів

Латентні елементи, які відображаються за допомогою ефекту прихованого зображення на відбитку – це текст, нанесений шрифтом з висотою знаків не більшою за 0,2 мм, практично не помітний для людського ока. При використанні копіювального обладнання з допомогою цього ефекту легко відрізнити оригінал від підробки [3] (рис. 5).



Рис. 5. Сканований вигляд документа, захищеного на основі прихованих елементів

За цим методом створюють нейтральний фон, який буде надруковано на оригіналі. При копіюванні оригіналу приховане зображення стає видимим, а також створюється фонові сітка, яка

містить унікальний візерунок, який з'являється під час копіювання. Фонова сітка базується на побудові ліній різної товщини [2]. Це дає змогу підтвердити справжність на рівні експертизи і відрізнити оригінал від підробки. Генерацією унікального коду забезпечують надійність, навіть якщо буде використано те саме програмне забезпечення.

Висновки

Проаналізовано графічні способи захисту, які дають змогу ефективно боротись з підробками та фальсифікацією документів. Проаналізовано можливості сучасних технологій захисту і встановлено, що надійний захист можна забезпечити, розробляючи нові інформаційні технології.

Розглянуто різні способи реалізації латентних зображень на основі гільйойшних елементів, мікрографіки, графічних пасток та прихованих елементів. Роботу проілюстровано прикладами.

1. Maria Nazarkevych Analysis of Software Protection and Development of Methods of Latency in Printed Documents / Maria Nazarkevych, Oksana Troyan // In Proc. of the VIIIth International Scientific and Technical Conference CSIT 2013, 16-18 November, Lviv 2013, p.120-121. 2. Дронюк І. Розробка методу захисту цінних паперів на стадії додрукарської підготовки / І. Дронюк, М. Назаркевич, О. Миронюк // Вісник Національного університету "Львівська політехніка". Комп'ютерні науки та інформаційні технології. - 2011. - № 694. - С. 352-358. 3. Назаркевич М. Аналіз сучасних методів та програмних ужитків з графічним захистом друкованих документів / Марія Назаркевич, Оксана Троян // Технічні вісті. – 2013. № 1 (37). – С. 42 – 44.

OPTIMAL PARAMETERS OF FOUR-LAYER OPTICAL STRUCTURES WHEN THE LIGHT IS INCIDENT AT AN ANGLE

© Petsko V., Mitsa A., Geche F., Kotsovsky V., Batyuk A., 2014

Наведено результати теоретичних розрахунків можливостей просвітлення чотиришаровою оптичною структурою низькозаломлювальної підкладки під час падіння світла під кутом. Алгоритм розрахунків заснований на використанні методу негладкої оптимізації (*r*-алгоритму).

Ключові слова: оптичне багатошарове покриття, *r*-алгоритм, матричний метод Абе́ле, закон Брюстера.

The results of theoretical calculations of the optimal four-layer structure for optical bleaching of the substrate when the light is incident at an angle are given in the paper. Computational algorithm is based on the nonsmooth optimization methods (*r*-algorithm).

Key words: optical multi-layer coating, *r*-algorithm, Abeles matrix method, law of Brewster.

Interference bleaching multilayer coatings are commonly used in optical systems, space systems, optical instrument, integrated optics, X-ray and neutron spectroscopy, electrodynamics of open systems, generators constructing, and other optical instruments. The aim of this study was to find the optimal parameters of four-layer homogeneous structures in enlightenment of dioptric substrate, dynamics of change with an increase in the spectral range when the light is incident at an angle. In this paper as the substrate was chosen the most popular low dioptric substrate with a refractive index $n_s = 1.51$ (glass).

We use Abeles matrix method [1-2] to calculate the spectral characteristics of optical multilayer coatings using. The objective function is the mean square error of the transmittance on spectral band under consideration [3-4]:

$$\Omega(I_1, I_2) = \max_{\bar{n}, \bar{d}} F(\bar{n}, \bar{d}) = \max_{\bar{n}, \bar{d}} \left(\frac{1}{L} \sum_{i=1}^L T^2(\bar{n}, \bar{d}, I_{(i)}) \right)^{1/2}, \quad (1)$$

where $\bar{n} = (n_1, n_2, \dots, n_{k-1}, n_k)$ – is a vector of the values of refraction indices of layers; $\bar{d} = (d_1, d_2, \dots, d_{k-1}, d_k)$ – is a vector of geometrical thicknesses of layers; L – is the number of points of the spectral band from I_1 to I_2 in case of its uniform division with an step ΔI : $\Delta I = (I_2 - I_1) / L$. Bounding values of the parameters were following $1.35 \leq n_j \leq 2.6$, $50 \text{ nm} \leq d_j \leq 750 \text{ nm}$ ($j = \overline{1, 4}$), $I_1 = 200 \text{ nm}$, $300 \text{ nm} \leq I_2 \leq 1200 \text{ nm}$.

The objective function has many local maxima. Therefore, we chose for our optimization algorithm searching optimal four-layer structures 512 initial approximations. They were chosen as follows: the whole area of all possible values of parameters was divided into 512 subregions. Then starting points of the algorithm was selected in each area. To find the optimal parameters we used *r*-algorithm, described in [5-6]. The software is written in Delphi environment. The study used special techniques of [1], which provided that the observable parameters lie in mentioned limits.

We studied the optimal value of the functional (1) for four-layer structures in the case of different angles of incidence. First consider the case when the beam falls perpendicularly to the layered surface (see Fig. 1a). It is well known that for a four-layer surfaces, as well as for two-layer and three-layer the optimal

refraction parameter values of the upper layer for all spectral bands is constant and equal to the lower limit for the refractive index — 1.35 [1]. By increasing the right boundary λ_2 from 400 to 600 nm values of the functional decreases from 0.997464 about 0.8%, with an increase of λ_2 from 600 to 650 nm — increased by 0.1%, while further enlargement of right boundary from 650 to 1200 nm causes fall by about 0.6%. Thus, the functional $\Omega(I_1, I_2)$ is not strictly decreasing with increasing right border, but has a tendency to decrease and this decrease is not significant.

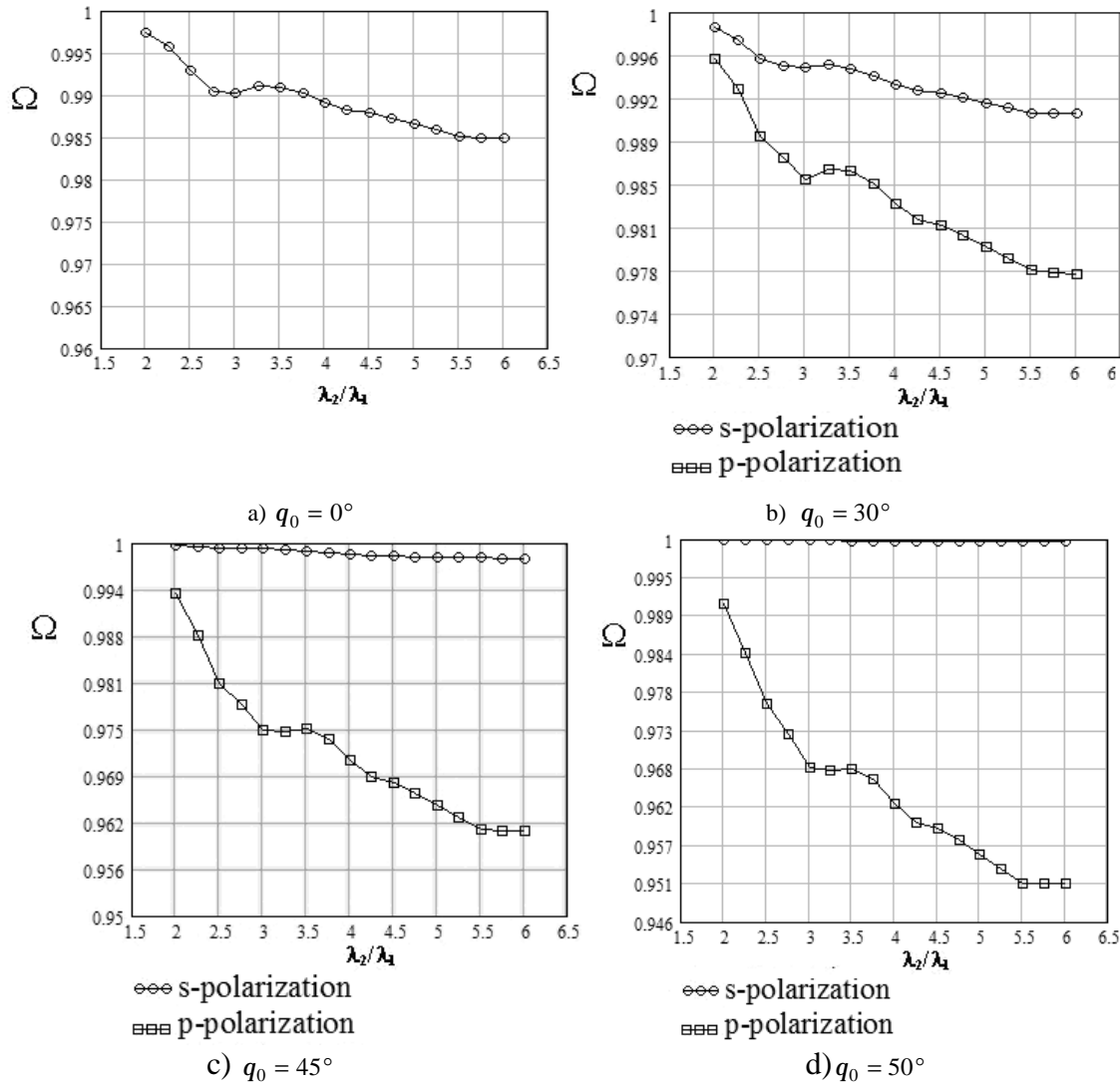


Fig. 1. The dynamics of the change of functional $\Omega(I_1, I_2)$ for four-layer homogeneous structure in cases $q_0 = 0^\circ$, $q_0 = 30^\circ$, $q_0 = 45^\circ$, $q_0 = 50^\circ$

When the angle is not equal to $q_0 = 0^\circ$ light beam can be decomposed into TE-wave (*s*-polarization) and TM-wave (*p*-polarization). These two cases are considered separately.

Then we studied the case when the angle of incidence is equal to $q_0 = 30^\circ$. The Fig. 1b shows that the functional $\Omega(I_1, I_2)$ in the *s*-polarization case decreases from 0.995880 by about 1%, while the right boundary increases from 400 to 600 nm, and in the case of *p*-polarization decreases from 0.998646 to about 0.3 %. By increasing the right boundary I_2 from 600 to 650 nm the functional value increases by 0.1 % for *s*-polarization and by 0.03% for *p*-polarization. The further decreasing of the right border up to 1200 nm causes the growth of the functional (1) by 0.9% for *s*-polarization and 0.4% for *p*-polarization. As

in case $q_0 = 0^\circ$ the functional Ω in both cases is not strictly decreasing, but has a tendency to decrease. We note that in case of s -polarization the decreasing is more rapid than in the case of p -polarization.

When incidence angle is equal to $q_0 = 45^\circ$ (Fig. 1c) or $q_0 = 50^\circ$ (Fig. 1d) the functional $\Omega(I_1, I_2)$ in p -polarization case slowly decreases from 0.999 to 0.996.

In the case that incidence angles equal to $q_0 = 55^\circ$ (Fig. 2a) or $q_0 = 60^\circ$ (Fig. 2b) the value of functional $\Omega(I_1, I_2)$ in p -polarization case decreases to 0.995 and 0.918 respectively.

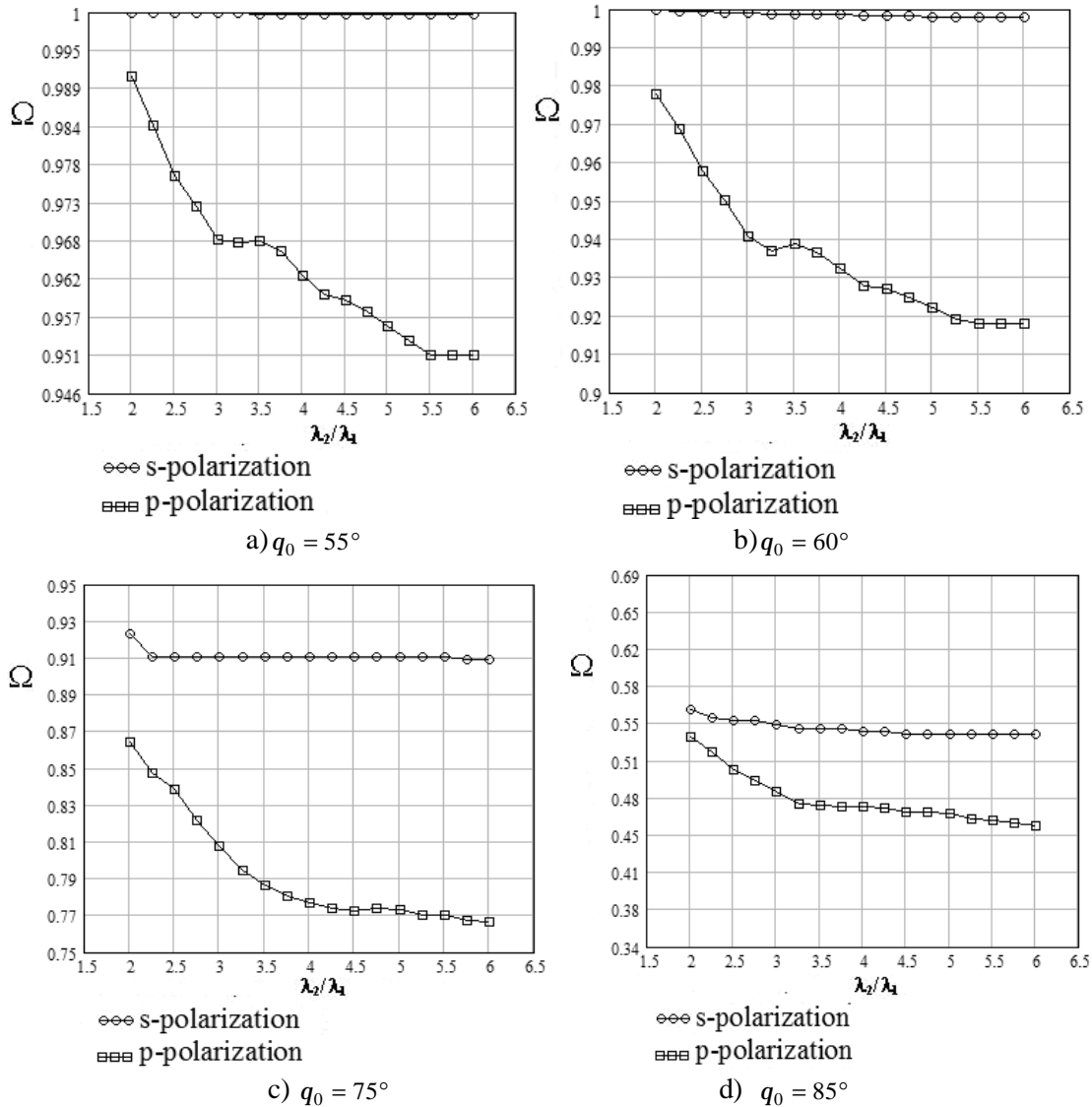
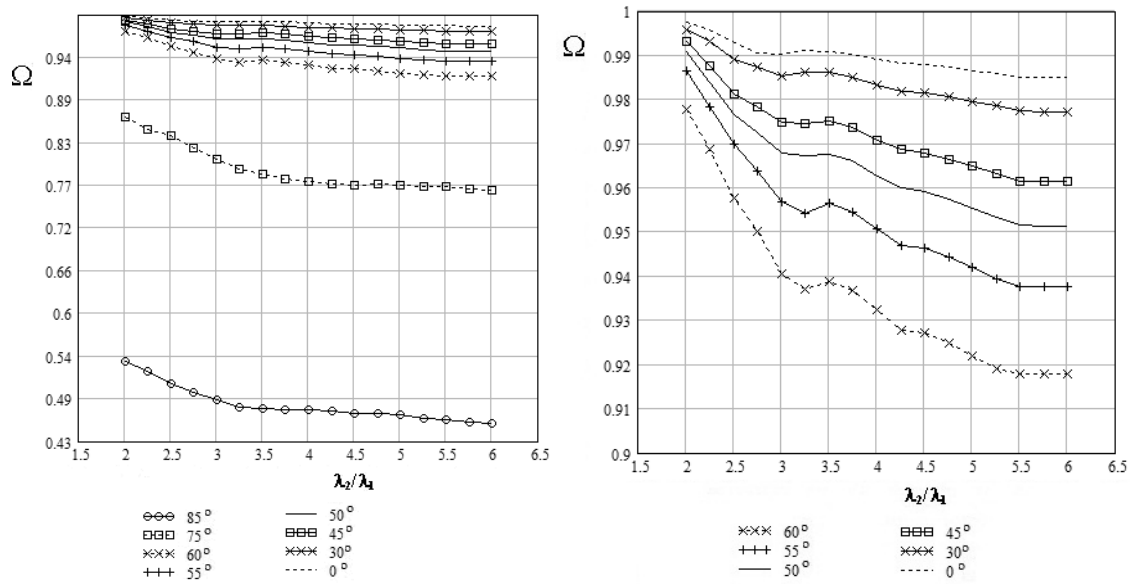


Fig. 2. The dynamics of the change of functional $\Omega(I_1, I_2)$ for four-layer homogeneous structure in cases $q_0 = 55^\circ, q_0 = 60^\circ, q_0 = 75^\circ, q_0 = 85^\circ$

In s -polarization case the decreases of the functional $\Omega(I_1, I_2)$ is significantly greater. Thus, when the incidence angle equals to $q_0 = 45^\circ$ the functional $\Omega(I_1, I_2)$ decreases approximately by 3.5 % starting from 0.993200 (Fig. 1c); in the case $q_0 = 50^\circ$ — from the value 0.990823 decreasing by about 4 % (Fig. 1d); in the case $q_0 = 55^\circ$ — from 0.986419 decreasing by about 4.5% (Fig. 2a); in the case $q_0 = 60^\circ$ — from the value 0.977784 decreasing by about 6 % (Fig. 2b).

If the angle of incidence equals to $q_0 = 75^\circ$ (Fig. 2c) the functional $\Omega(I_1, I_2)$ in s -polarization case decreases from 0.864027 to 0.772539 with increasing right border I_2 from 400 to 900 nm. If the right border I_2 increases from 900 to 950 nm the functional (1) increases up to 0.773879. Further increase of the right border I_2 from 950 to 1200 nm causes that the value of our objective functional slowly decreases and comes to 0.766204, i.e. in s -polarization case $\Omega(I_1, I_2)$ has the tendency to decrease. In the case of p -polarization the functional (1) decreases from 0.923152 to 0.910704, while I_2 grows from 400 to 450 nm. Further increase of the right border I_2 to 1200 nm causes the slowly decrease of objective functional to 0.910499, that is the functional $\Omega(I_1, I_2)$ in p -polarization case is decreasing.

Fig. 2d provides an illustration of behavior of objective functional in the case of s -polarization for the angle of incidence $q_0 = 85^\circ$. If the right border I_2 increases from 400 to 650 nm, the value of $\Omega(I_1, I_2)$ rapidly decreases from 0.537593 to 0.475012. Further increase of the right border λ_2 from 650 to 1200 nm causes a slowly decrease of (1) to 0.454174. In the case of p -polarization one can observe the decrease of the functional $\Omega(I_1, I_2)$, from 0.563425 to 0.539841, while the right border λ_2 changes from 400 to 1200 nm. For both type of polarization the objective functional is decreasing, and for s -polarization its decrease is less than one in the case of p -polarization.



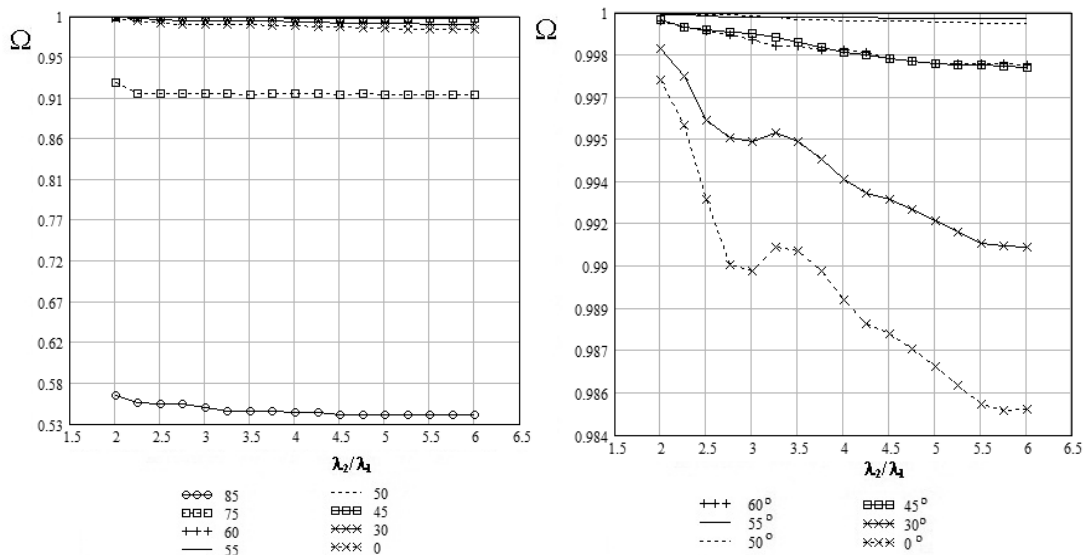
a) angle of incidence from 0° to 85°

b) angle of incidence from 0° to 60°

Fig. 3. Dynamics of the functional $\Omega(I_1, I_2)$ changes for a four-layer homogeneous structure in s -polarization case

From Fig. 3 it is possible to detect the dynamics of the changes of the functional (1) in the case of a four-homogeneous structure with s -polarization depending on the angle: for all angles within the right border λ_2 range from 400 nm to 1200 nm the objective functional tends to decrease. If $q_0 = 0^\circ$ or $q_0 = 30^\circ$, then for values of λ_2 between 600 nm and 650 nm a small growth of $\Omega(I_1, I_2)$ can be observed. In the cases $q_0 = 45^\circ$, $q_0 = 50^\circ$, $q_0 = 55^\circ$, $q_0 = 60^\circ$ this growth occurs in the range from 650 nm to 700 nm. For the angle $q_0 = 75^\circ$ a growth observed in the range from 900 nm to 950 nm, and for the angle $q_0 = 85^\circ$ the objective functional is everywhere decreasing. The figure shows that in the case of s -polarization the value of functional $\Omega(I_1, I_2)$ decreases.

In the case of p -polarization (Fig. 4) with increasing angle to $q_0 = 55^\circ$ the value of the functional (1) firstly grows, and then decreases. This effect is explained by the law of Brewster [7-9]. For p -polarization, as well as for s -polarization the objective functional decreases when λ_2 changes from 400 nm to 1200 nm for all angles, except the angles $q_0 = 0^\circ$ and $q_0 = 30^\circ$ for which the functional significantly increases in the range from 600 nm to 650 nm). In the case of p -polarization fluctuations of the objective function are less noticeable than ones in the case of s -polarization, and the functional $\Omega(I_1, I_2)$ value changes very slowly.



a) angle of incidence from 0° to 85°

b) angle of incidence from 0° to 60°

Fig. 4. Dynamics of the functional $\Omega(I_1, I_2)$ changes for a four-layer homogeneous structure in the case of p -polarization

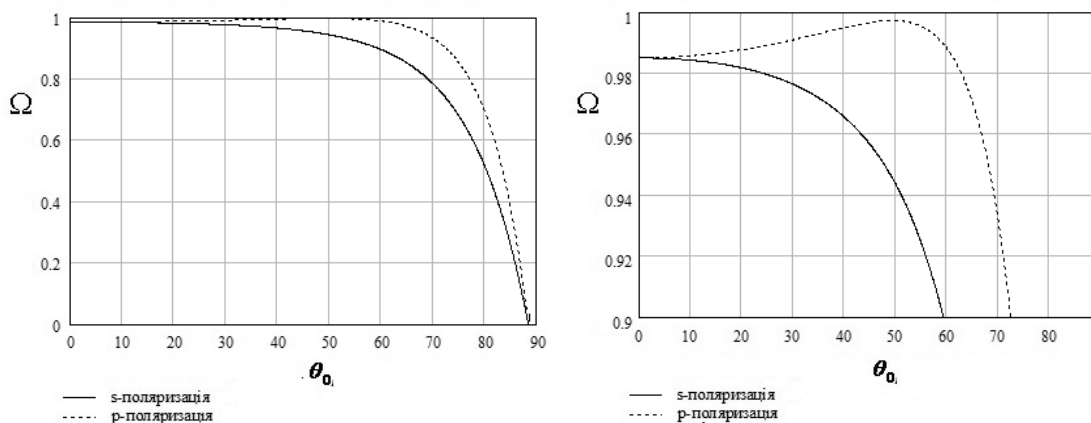


Fig. 5. Dynamics of changes of the functional $\Omega(I_2, I_1)$ for a four-homogeneous structure depending on q_0

The value of the functional (1) on the angle $q_0 = 55^\circ$ has smaller fluctuations than on other angles, and with a further increase of the angle the value of the functional decreases. The plots of functional (1) for the angles $q_0 = 50^\circ$ and $q_0 = 55^\circ$ intersect (Fig. 4b).

It is possible to see the following dependence of the objective function of the angle of incidence: in the case of s -polarization the objective function is decreasing, and in the case of p -polarization the

objective function first increases and then decreases (Fig. 5), which is also explained by the Brewster law [7–9].

Calculations have shown relatively high performance of software developed on the base of the r -algorithm for solving the problem of optimization of multilayer optical structures. The solving of our optimization problem on Intel (R) Core (TM) i5-3230@3.30 GHz takes a few minutes for each spectral band. Dynamics of changes of the functional (1) for a four-layer homogeneous structure as a function, depending on the angle of incidence is consistent with the Brewster law.

1. Furman Sh. *Basics of optics of multilayer systems* / Sh.Furman., A.V. Tikhonravov Editions *Frontiers, Gif-sur Yvette*, 1992. – 242 p. 2. Abeles F. *Matrix method* / *Ann.de Physique*. – 1950. – V.5. – P. 596-640. 3. Stetsyuk P.I., Mitsa A. V. *Parameters optimization for multilayer optical coatings* // *Kibernetika i sistemnyy analiz*. – Kiev, 2005. – pp. 107-115 (in Russian). 4.Mitsa A.V., Stetsyuk P.I. *Search problem of optimal parameters for homogenous optical coating* // *Teoriya optymalnyh rishen*. – № 2. – Kiev, 2003. – pp. 127-134 (in Ukrainian). 5. Shor N.Z., Zhurbenko N.G. *Minimization methods on base of spaces stretch in direction of the difference of two successive gradients* // *Kibernetika*. – 1971. – #3. – pp. 51–59 (in Russian). 6. Shor N.Z.. *Nondifferentiable Minimization and Application*. – Kiev: Naukova Dumka, 1979 (In Russian). 7. Krylova T.N. *Interference Coating*. – Leningrad: Mashinostroyeniye, 1973 (In Russian). 8.Putilin E. S. *Optical Coating. Textbook*. – St. Petersburg: St. Petersburg ITMO, 2010 (In Russian). 9. Ritter E. *Film dielectric for applications in optics* // *Physics of thin films*. – Moskva: Mir, 1978. – vol. 8. – pp. 7-27 (In Russian).

ПРОГНОЗУВАННЯ ЙМОВІРНОСТІ РОБОТИ ТА ВІДМОВИ ЗА ЗАДАНОЇ УМОВИ ГОТОВНОСТІ СИМЕТРИЧНИХ ІЄРАРХІЧНИХ СИСТЕМ, ГАЛУЖЕНИХ ДО 4-ГО РІВНЯ, ЗА ДОПОМОГОЮ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ

© Павлюк О., 2014

Запропоновано програмний модуль для розрахунку ймовірності роботи та відмови за заданої умови готовності ізотропних симетричних ієрархічних розгалужених систем (ІРС) з галуженням до n -го рівня, елементи якої підпорядковуються експоненційному закону. Здійснено прогнозування цих характеристик надійності за допомогою неітераційної штучної нейронної мережі (ШНМ). Розраховано середньоквадратичну приведену до діапазону значень похибку навчання і прогнозу, а також оцінено час навчання та прогнозування ШНМ.

Ключові слова: ієрархічна розгалужена система, штучна нейронна мережа, прогнозування, неітераційне навчання ШНМ, ШНМ з “вузьким горлом”.

The software module is developed. By the specified readiness parameters it calculates probabilities of the proper operation and failure-ability for the isotropic symmetric and hierarchical branched systems (HBS). The module is tested/restricted against the systems of the n -level branching whose elements obey the exponential rules. The non-iterative artificial neural network (ANN) has been deployed to the prediction of those characteristics. The reduced to the mean value range errors of the ANN learning and forecasting are calculated as well as the time estimations for the ANN learning and forecasting.

Key words: hierarchical branched system, artificial neural network forecasting non-iterational training ANN, ANN with “bottles neck”.

Вступ

В умовах ринкової економіки України надзвичайно актуальним є підвищення надійності складних систем. Без розрахунку характеристик надійності неможливо проектувати й експлуатувати системи. А тому виникає необхідність розробляти методи прогнозування показників надійності систем з врахуванням їх специфіки.

Більшість систем можна описати у вигляді ієрархічних деревоподібних розгалужених структур. За допомогою ІРС також описують структури комп'ютерних мереж. На нульовому рівні таких систем знаходиться сервер, на проміжних рівнях – концентратори інформації, а на вихідному – робочі станції. Такі мережі можуть мати симетричну, несиметричну, неізотропну, з переходами через рівні структуру з розгалуженням до n -го рівня. Також вони можуть бути з простим і складним підпорядкуванням.

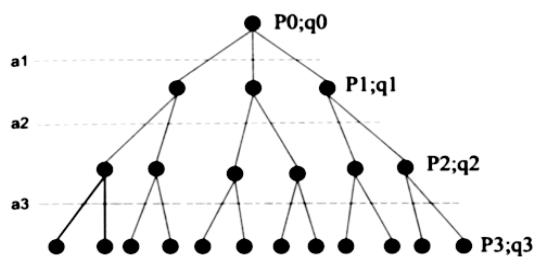


Рис. 1. Симетрична ІРС галужена до 3-го рівня.

Опис алгоритму

Найпростішим видом є симетрична ІРС. Для наочності на рис.1. показано симетричну ІРС галужену до 3-го рівня. Через $P_0, q_0, P_1, q_1, P_2, q_2, P_3, q_3$ позначено ймовірності роботи і відмови елементів системи для нульового, першого, другого і третього рівнів відповідно. Позначимо через a_1, a_2, a_3 коефіцієнти галуження для першого, другого і третього рівнів.

Для дослідження невідновлюваних симетричних ІРС використовують твірні функції і рекурентний вираз [1]. Твірна функція $S_m(Z)$ відображає розклад за степенями довільного параметра Z , який виділяє ймовірності P_m^n як коефіцієнти при Z . Твірну функцію для системи з рис. 1 можна записати так:

$$S_3(Z) = P_0(P_1(P_2(P_3Z + q_3)^{a_3} + q_2)^{a_2} + q_1)^{a_1} + q_0 \quad (1)$$

Розкладемо твірну функцію за формулою бінома Ньютона, позначивши через x_1, x_2, x_3 кількість працюючих елементів на першому, другому і третьому рівнях відповідно.

$$S_3(Z) = P_0 \sum_{x_1=0}^{a_1} C_{a_1}^{x_1} P_1^{x_1} q_1^{a_1-x_1} \sum_{x_2=0}^{a_2x_1} C_{a_2x_1}^{x_2} P_2^{x_2} q_2^{a_2x_1-x_2} \sum_{x_3=0}^{a_3x_2} C_{a_3x_2}^{x_3} P_3^{x_3} q_3^{a_3x_2-x_3} Z^{x_3} + q_0 \quad (2)$$

За формулою (2) запишемо рекурентний вираз (3), позначивши через функцію $\text{ceil}\left(\frac{x_i}{a_i}\right) i = \overline{1,3}$ заокруглення до найбільшого цілого.

$$S_3(Z) = P_0 \sum_{x_1=\text{ceil}\left(\frac{x_2}{a_2}\right)}^{a_1} C_{a_1}^{x_1} P_1^{x_1} q_1^{a_1-x_1} \sum_{x_2=\text{ceil}\left(\frac{x_3}{a_3}\right)}^{a_2x_1} C_{a_2x_1}^{x_2} P_2^{x_2} q_2^{a_2x_1-x_2} C_{a_3x_2}^{x_3} P_3^{x_3} q_3^{a_3x_2-x_3} \quad (3)$$

Позначимо через $P_3(x_3)$ ймовірність того, що працюють x_3 вихідних елементів системи. Якщо $0 < x_3 \leq a_1a_2a_3$, то отримаємо формулу (4) [2, 3]. Якщо $x_3 = 0$, то до формули (4) додають ще q_0 .

$$P_3(x_3) = P_0 \sum_{x_1=\text{ceil}\left(\frac{x_2}{a_2}\right)}^{a_1} C_{a_1}^{x_1} P_1^{x_1} q_1^{a_1-x_1} \sum_{x_2=\text{ceil}\left(\frac{x_3}{a_3}\right)}^{a_2x_1} C_{a_2x_1}^{x_2} P_2^{x_2} q_2^{a_2x_1-x_2} \sum_{x_3=0}^{a_3x_2} C_{a_3x_2}^{x_3} P_3^{x_3} q_3^{a_3x_2-x_3} \quad (4)$$

Нехай всі елементи системи підпорядковуються експоненційному закону. Тоді $P_3(x_3, t)$ ймовірність того, що працюють x_3 вихідних елементів системи, набуде такого вигляду:

$$P_3(x_3, t) = e^{-I_0t} \sum_{x_1=\text{ceil}\left(\frac{x_2}{a_2}\right)}^{a_1} C_{a_1}^{x_1} e^{-I_1x_1t} (1 - e^{-I_1t})^{a_1-x_1} \sum_{x_2=\text{ceil}\left(\frac{x_3}{a_3}\right)}^{a_2x_1} C_{a_2x_1}^{x_2} e^{-I_2x_2t} (1 - e^{-I_2t})^{a_2x_1-x_2} * \\ * C_{a_3x_2}^{x_3} e^{-I_3x_3t} (1 - e^{-I_3t})^{a_3x_2-x_3} \quad (5)$$

Розглянемо симетричну, галужену до n -го рівня ІРС, $n \in Z$. Запишемо $P_n(x_n, t)$ ймовірність того, що працюють x_n вихідних елементів системи.

$$P_n(x_n, t) = e^{-I_0t} \sum_{x_1=\text{ceil}\left(\frac{x_2}{a_2}\right)}^{a_1} C_{a_1}^{x_1} e^{-I_1x_1t} (1 - e^{-I_1t})^{a_1-x_1} * \dots * \sum_{x_{n-1}=\text{ceil}\left(\frac{x_n}{a_n}\right)}^{a_{n-1}x_{n-2}} C_{a_{n-1}x_{n-2}}^{x_{n-1}} e^{-I_{n-1}x_{n-1}t} (1 - e^{-I_{n-1}t})^{a_{n-1}x_{n-2}-x_{n-1}} * \\ * C_{a_nx_{n-1}}^{x_n} e^{-I_nx_nt} (1 - e^{-I_nt})^{a_nx_{n-1}-x_n} = e^{-I_0t} \sum_{x_1=\text{ceil}\left(\frac{x_2}{a_2}\right)}^{a_1} C_{a_1}^{x_1} e^{-I_1x_1t} (1 - e^{-I_1t})^{a_1-x_1} * \\ * \prod_{i=2}^{n-1} \sum_{x_i=\text{ceil}\left(\frac{x_i}{a_i}\right)}^{a_{i-1}x_{i-2}} C_{a_{i-1}x_{i-2}}^{x_{i-1}} C_{a_ix_{i-1}}^{x_i} e^{-I_ix_it} (1 - e^{-I_it})^{a_ix_{i-1}-x_i} * C_{a_nx_{n-1}}^{x_n} e^{-I_nx_nt} (1 - e^{-I_nt})^{a_nx_{n-1}-x_n} \quad (6)$$

Ймовірність відмови x_n вихідних елементів системи обчислюємо за формулою

$$Q_n(x_n, t) = 1 - P_n(x_n, t) \quad (7)$$

Розроблено програмний модуль на *Python* для розрахунку за формулами (6) та (7) ймовірності роботи і відмови за заданої умови готовності симетричної ізотропної ІРС, галуженої до n -го рівня, в якій всі елементи підпорядковуються експоненційному закону. Особливістю цього програмного модуля є динамічність змін: кількості параметрів ІРС; значення цих параметрів; кроку, з яким параметри необхідно перерахувати.

Вхідними даними для програмного модуля є:

- n – кількість рівнів симетричної ІРС, $n \in Z$;
- a_i – коефіцієнти галуження i -го рівня ІРС, $i = \overline{1, n}$;
- I_j – інтенсивність відмов елемента j -го рівня, $j = \overline{0, n}$;
- t – час, протягом якого працює ІРС;
- x_n – умова готовності.

Дані зібрані для подальшого їх використання у навчанні і тестуванні ШНМ, яку пропонується використовувати як універсальніший інструмент для передбачення надійності системи на ранніх стадіях проектування без виведення складних формул їх програмування, відлагодження та здійснення складних обчислень.

Результати обчислень автоматично записуються у форматі *.csv у файл даних для подальшого їх використання у навчанні і тестуванні ШНМ, яка є доволі універсальним інструментарієм. За стовбцями підряд зліва направо записуємо вхідні параметри, третій з кінця стовпець – значення часу з дискретністю в 1 годину, передостанній стовпець – $P_n(x_n, t)$, останній – $Q_n(x_n, t)$. Розрахунок і запис у файл вхідних значень і $P_n(x_n, t)$ та $Q_n(x_n, t)$ було здійснено за таких

значень вхідних параметрів: $n = 4$; $a_i = \overline{1,6}$, $i = \overline{1, n}$; $I_j = 10^{-3}$ [1/год], $j = \overline{0, n}$; $x_n = \prod_{l=1}^n x_l$. Такий файл

має 125900 рядків. Розмірність вихідного файла – 6.58 Мб. Запис результатів на ПК з 2-ма фізичними ядрами (i3; 2.1ГГц) та 3 Гб оперативної пам'яті тривав 18 секунд.

Для здійснення прогнозування $P_n(x_n, t)$ та $Q_n(x_n, t)$ було вибрано неітераційну ШНМ на основі нейропарадигми “функціонал на множині табличних функцій”, тобто нейромережу радіальних базових функцій (RBF), розроблену професором Р.О. Ткаченком [4, 5]. Основними перевагами такої нейромережі є швидкий процес навчання за рахунок неітераційного алгоритму, в основу якого покладено метод ортогоналізації в просторі векторів – реалізацій досліджуваного процесу. Це забезпечує швидкість виконання алгоритму в режимі реального часу. Точності відтворення результатів досягають налаштуванням мережі, тобто вибором її параметрів [6, 7]. При навчанні використовується зовнішній критерій – відповідність прогнозованих результатів значенням контрольної вибірки.

Результати обчислень $P_n(x_n, t)$ та $Q_n(x_n, t)$ поділяємо на два файли: навчання і тестування ШНМ. Для навчання ШНМ використовуємо 2000 значень, а перевіряємо прогноз на 136 значеннях. Для прогнозування експериментальним шляхом було вибрано ШНМ типу bottles neck (з “вузьким горлом”) з 8 нейронів на вхідному прошарку, 7 – у прихованому прошарку і 8 нейронів на вихідному прошарку. В неї кількість нейронів вихідного прошарку є більшою за кількість нейронів у прихованому прошарку і дорівнює кількості нейронів вхідного прошарку. Вхідний та прихований шари нейромережі виконують компресію даних, а прихований та вихідний – декомпресію. Така ШНМ забезпечує глибоку компресію даних з метою пониження розмірності “входів–виходів” саме для побудови карти в новій двовимірній системі координат. Реалізація цієї процедури є можливою завдяки використанню ШНМ з нелінійними синаптичними зв'язками, тобто нейромережі радіальних базових функцій.

Результати навчання ШНМ по розрахунку $P_n(x_n, t)$ для кожного виходу показано на рис.2, а прогнозування – на рис. 3.

Середньоквадратична приведена до діапазону похибка для обчислення $P_n(x_n, t)$ за допомогою ШНМ становила – 1.03 %, максимальна – 6.61 %. А похибка для обчислення $Q_n(x_n, t)$ становила 3.36 %, а максимальна – 8.87 %. Час прогнозування менший за 1 с, за задовільної точності. Такий час прогнозування приблизно співмірний з часом розрахунку по формулах (6) і (7) для малої кількості даних. Але якщо кількість значень $P_n(x_n, t)$ чи $Q_n(x_n, t)$ перевищує 500, то нейромережа здійснює прогноз швидше, ніж розрахунок за формулами (6) і (7).

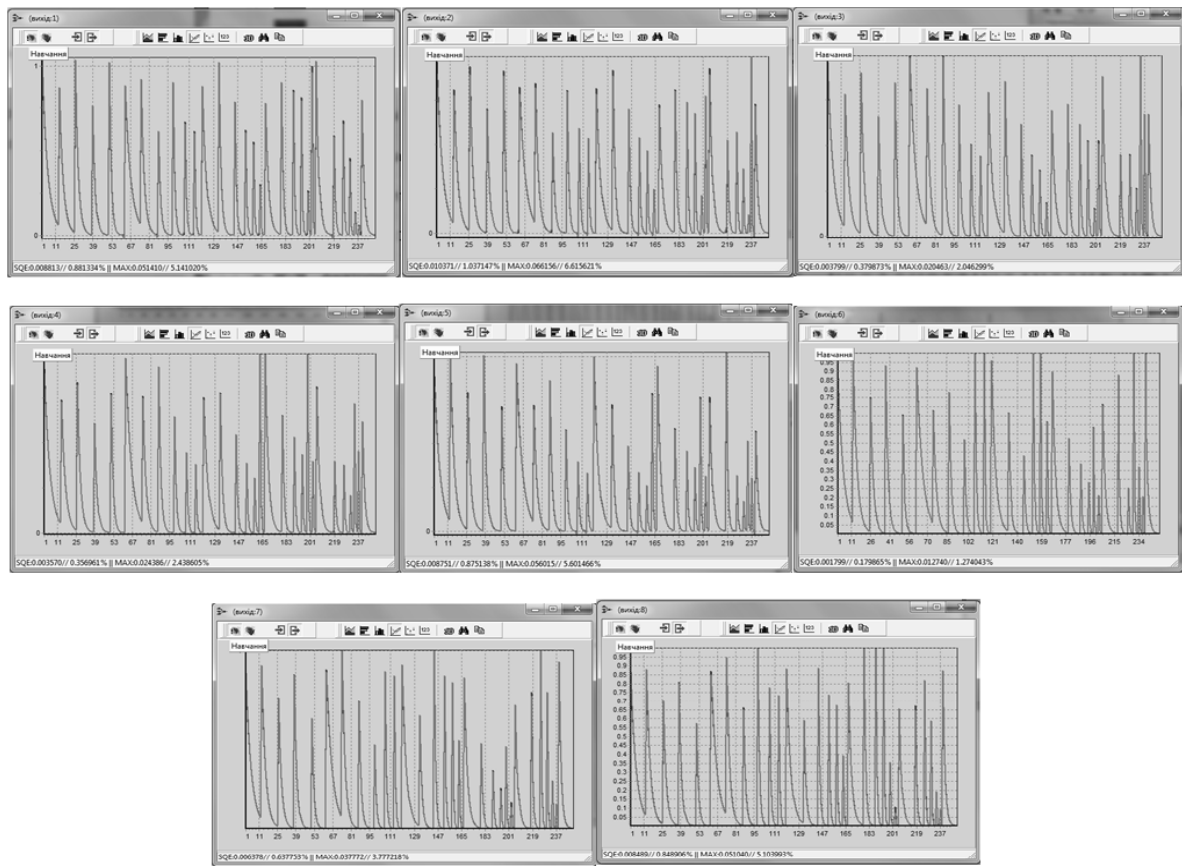


Рис. 2. Режим навчання ШМ для $P_n(x_n, t)$

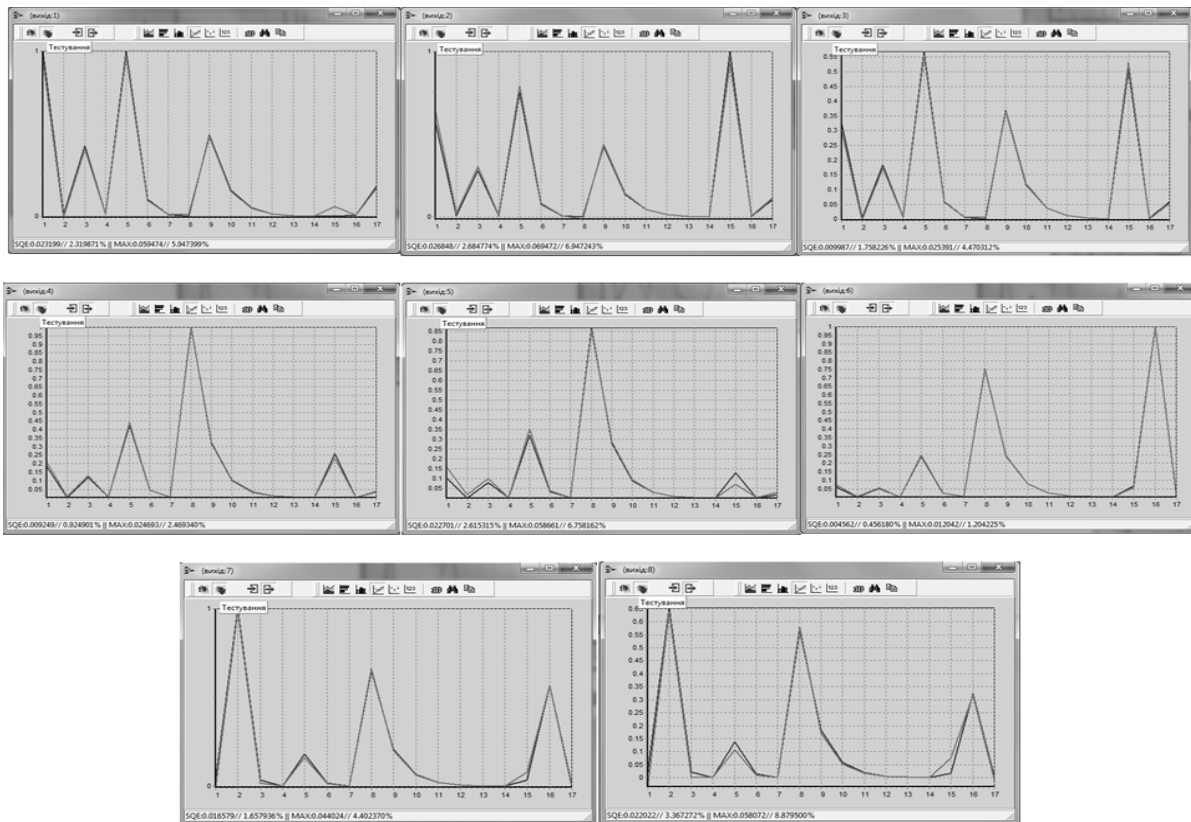


Рис. 3. Режим прогнозування ШМ для $P_n(x_n, t)$

Результати порівняння трьох змінених вхідних параметрів обчислень $P_n(x_n, t)$ та $Q_n(x_n, t)$ та їхніх прогнозованих нейромережею значень наведено у таблиці.

Результати досліджень

P_1	P_2	$ P_1 - P_2 $	Q_1	Q_2	$ Q_1 - Q_2 $	P_1	P_2	$ P_1 - P_2 $	Q_1	Q_2	$ Q_1 - Q_2 $
1	0,942	0,058	0,057	0,057	0,057	0,322	0,352	0,03	0,647	-0,000	0,000
0,568	0,634	0,066	0,365	0,336	0,336	0,243	0,249	0,006	0,750	0,092	0,092
0,323	0,297	0,026	0,702	0,646	0,646	0,183	0,160	0,023	0,839	0,171	0,171
0,184	0,206	0,022	0,793	0,710	0,710	0,138	0,108	0,03	0,891	0,214	0,214
0,104	0,160	0,056	0,839	0,729	0,729	0,104	0,098	0,006	0,901	0,214	0,214
0,059	0,071	0,012	0,928	0,793	0,793	0,078	0,082	0,004	0,917	0,222	0,222
0,034	-0,006	0,04	1,006	0,846	0,846	0,059	0,056	0,003	0,943	0,239	0,239
0,019	-0,034	0,053	1,034	0,850	0,850	0,044	0,045	0,001	0,954	0,241	0,241
0,011	-0,005	0,016	1,005	0,798	0,798	0,034	0,038	0,004	0,961	0,240	0,240
0,006	0,025	0,019	0,974	0,744	0,744	0,025	0,025	0	0,974	0,245	0,245
0,004	-0,003	0,007	1,003	0,751	0,751	0,019	0,015	0,004	0,984	0,247	0,247
0,002	0,008	0,006	0,991	0,718	0,718	0,014	0,009	0,005	0,990	0,246	0,246
0,001	0,017	0,016	0,982	0,688	0,688	0,011	0,010	0,001	0,989	0,238	0,238
0,001	0,004	0,003	0,995	0,681	0,681	0,008	0,008	0	0,991	0,232	0,232
1	0,987	0,013	0,012	-0,321	0,321	0,006	0,005	0,001	0,994	0,229	0,229
0,654	0,637	0,017	0,362	0,009	0,009	0,005	0,005	0	0,994	0,222	0,222
0,428	0,405	0,023	0,594	0,223	0,223	0,003	0,003	0	0,996	0,217	0,217
0,28	0,304	0,024	0,695	0,306	0,306	0,003	0,003	0	0,996	0,211	0,211
0,183	0,172	0,011	0,827	0,420	0,420	0,002	0,001	0,001	0,998	0,207	0,207
0,12	0,128	0,008	0,871	0,447	0,447	0,001	0,000	0,001	0,999	0,202	0,202
0,079	0,100	0,021	0,899	0,459	0,459	0,001	0,013	0,012	0,986	0,183	0,183
0,051	0,055	0,004	0,944	0,488	0,488	0,001	-0,013	0,014	1,013	0,204	0,204
0,034	0,018	0,016	0,981	0,509	0,509	0,001	0,006	0,005	0,993	0,179	0,179
0,022	0,001	0,021	0,998	0,511	0,511	1	0,995	0,005	0,004	-0,814	0,814
0,014	0,013	0,001	0,986	0,485	0,485	0,868	0,856	0,012	0,143	-0,680	0,680
0,009	0,009	0	0,990	0,474	0,474	0,753	0,751	0,002	0,248	-0,580	0,580
0,006	0,005	0,001	0,994	0,464	0,464	0,653	0,662	0,009	0,337	-0,496	0,496
0,004	0,004	0	0,995	0,452	0,452	0,567	0,579	0,012	0,420	-0,418	0,418
0,003	0,003	0	0,996	0,440	0,440	0,492	0,481	0,011	0,518	-0,325	0,325
0,002	0,002	0	0,997	0,428	0,428	0,427	0,439	0,012	0,560	-0,287	0,287
0,001	0,000	0,001	0,999	0,418	0,418	0,37	0,365	0,005	0,634	-0,218	0,218
0,001	0,000	0,001	0,999	0,406	0,406	0,321	0,325	0,004	0,674	-0,182	0,182
1	0,968	0,032	0,031	-0,573	0,573	0,279	0,289	0,01	0,710	-0,150	0,150
0,754	0,791	0,037	0,208	-0,407	0,407	0,242	0,244	0,002	0,755	-0,109	0,109
0,568	0,555	0,013	0,444	-0,182	0,182	0,21	0,202	0,008	0,797	-0,071	0,071
0,428	0,440	0,012	0,559	-0,078	0,078	0,182	0,171	0,011	0,828	-0,043	0,043

Позначення у таблиці є такими: P_1 – розрахована за формулою (6) ймовірність роботи за заданої умови готовності неізотропної симетричної галуженої до 4-го рівня ІРС; P_2 – спрогнозовані за допомогою неітераційної ШНМ ймовірність роботи за заданої умови готовності неізотропної симетричної, галуженої до 4-го рівня, ІРС; $|P_1 - P_2|$ – абсолютні значення різниці P_1 і P_2 ; Q_1 – розрахована за формулою (7) ймовірність відмови за заданої умови готовності неізотропної симетричної, галуженої до 4-го рівня, ІРС; Q_2 – спрогнозовані за допомогою неітераційної ШНМ ймовірність відмови за заданої умови готовності неізотропної симетричної, галуженої до 4-го рівня, ІРС; $|Q_1 - Q_2|$ – абсолютні значення різниці Q_1 і Q_2 .

Результати дійсних (синя крива) і прогнозованих (червона крива) за допомогою ШНМ значень $P_n(x_n, t)$ та $Q_n(x_n, t)$ показано на рис. 4.

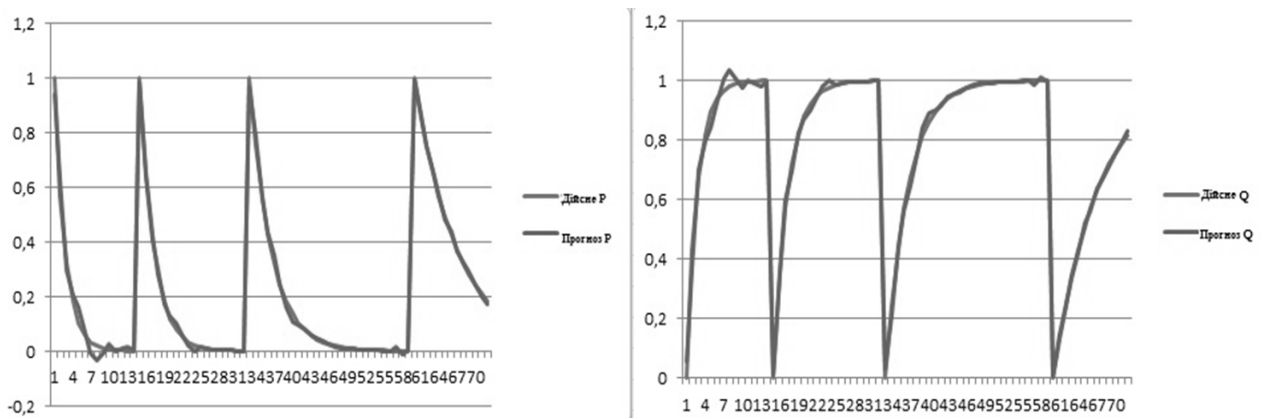


Рис. 4. Дійсні і прогнозовані за допомогою ШНМ для $P_n(x_n, t)$ та $Q_n(x_n, t)$

Аналіз отриманих результатів показує, що використання нейромереж для прогнозування характеристик надійності симетричних ІРС дає задовільні результати. Найбільшим є вигравш у швидкодії при здійсненні прогнозу характеристик надійності ІРС за допомогою ШНМ порівняно з розрахунком за формулами (6) і (7) за великої кількості рівнів галуження системи. Також швидкодія є вищою, якщо в системі є високонадійні елементи, тобто у випадках, коли необхідно опрацювати велику кількість статистичних даних.

Встановлено, що для здійснення прогнозу показників надійності ІРС якнайкраще підходить ШНМ типу bottles neck внаслідок можливості компресії великої кількості даних.

ШНМ можна використовувати для швидкого розрахунку параметрів ІРС із задовільною точністю, що є особливо актуальним при: розробці ІРС на етапі проектування; виявленні “слабкого” місця в ІРС, моделюванні роботи ІРС у режимі реального часу тощо.

1. Ушаков И.А. Вероятностные модели надежности информационно-вычислительных систем. – М.: Радио и связь, 1991. – 132 с. 2. Марунчак Д.Є., Сидор А.Р. Розрахунок надійності ієрархічних розгалужених систем зі старіючими вихідними елементами: Навч. посібник. – Львів: Видавництво Національного університету “Львівська політехніка”, 2006. – 72 с. 3. Марунчак Д.Є., Сидор А.Р. Надійність розгалужених систем: Навч. посібник. – Львів: Видавництво Національного університету “Львівська політехніка”, 2007. – 124 с. 4. Ткаченко Р.О. Неітераційне навчання штучних нейронних мереж прямого поширення // Технічні вісті. – 1999. – № 1(8), 2(9). – С. 41–42. 5. Ткаченко Р.О. Штучна нейронна мережа таблично-алгоритмічного типу // Вісник Держ. ун-ту “Львівська політехніка” “Комп’ютерна інженерія та інформаційні технології”. – 1999. – № 351. – С. 91–93. 6. Ткаченко Р.О. Павлюк О.М. Козут Р.М. Лінеаризація задач передбачення шляхом фазифікації // Вісник Нац. ун-ту “Львівська політехніка”. “Комп’ютерна інженерія та інформаційні технології”. – 2004. – № 521. – С. 214–219. 7. Ткаченко Р.О., Павлюк О.М., Ткаченко П.Р. Особливості тренування нейромереж прямого поширення за методом “К найближчих сусідів” // Відбір і обробка інформації. – 2004. – № 20(96). – С. 121–126.

МАРШРУТИЗАЦІЯ ПАКЕТІВ З ВРАХУВАННЯМ ПРОПУСКНИХ ЗДАТНОСТЕЙ КАНАЛІВ ТА КІЛЬКОСТІ ПРОМІЖНИХ ВУЗЛІВ

Ї Обельовська К., Роман В., Ліскевич О., 2014

Запропоновано алгоритм маршрутизації для комп’ютерних мереж, що враховує пропускні здатності каналів та кількість проміжних вузлів.

Ключові слова: комп’ютерні мережі, маршрутні таблиці, пропускна здатність каналу, кількість проміжних переходів.

Modified routing algorithm for computer networks based on channel throughput and number hops is proposed.

Key words: computer network, routing table, bandwidth, hop.

Вступ

Ефективність функціонування компютерних мереж залежать від багатьох чинників, серед яких важливе місце займає маршрутизація. В сучасних комп’ютерних мережах одним з основних протоколів, що використовуються для маршрутизації в межах автономної системи, є протокол OSPF (Open Shortest Path First), описаний в стандарті RFC 2328.

Постановка задачі

Протокол OSPF для обчислення найкоротшого шляху транспортування пакетів використовує алгоритм Дейкстри. Алгоритм розглядає кожен маршрутизатор як вузол дерева і визначає найкоротший шлях від нього до кожного вузла призначення на основі сукупної вартості доступу до нього. Як вартість шляху (метрики) протокол OSPF за замовчуванням використовує метрику, що ґрунтується на пропускній здатності каналів зв’язку. Що вища пропускна здатність каналу, то нижча вартість шляху, а отже, пакет буде відправлений по шляху з максимальною пропускною здатністю [1, 2].

На основі дерева найкоротших шляхів, знайденого за алгоритмом Дейкстри, будуються маршрутні таблиці маршрутизаторів.

Вхідними даними для алгоритму Дейкстри є зважений граф, що відповідає топології мережі, ваги ребер якого враховують пропускні здатності каналів мережі. Назвемо таку вагу – метрика 1 (M1). Недоліком вищеописаного підходу є те, що при знаходженні шляху транспортування пакета враховується тільки один параметр мережі, за замовчуванням – це пропускна здатність каналів. Проте цілком можливо, що для забезпечення інших критеріїв цей шлях не є тільки не оптимальним але навіть задовільним. Існує ряд робіт, присвячених вдосконаленню маршрутизації пакетів, прикладами яких можуть бути роботи [3 – 5]. Нами пропонується при формуванні маршрутних таблиць додатково до метрики M1, що ґрунтується на пропускній здатності каналів, враховувати кількість проміжних вузлів (маршрутизаторів) на шляху до вузла призначення. Назвемо цей параметр – метрика 2 (M2). Кожен маршрутизатор, через який транспортується пакет, вносить затримку, що збільшує загальний час транспортування пакета. Тому кількість проміжних маршрутизаторів на шляху проходження пакета для багатьох прикладних задач є одним з важливих критеріїв оптимізації. Інколи, наприклад, при передаванні мови, буває доцільніше відправляти пакети по шляху, пропускна здатність якого не є високою, проте містить малу кількість проміжних вузлів.

3. Побудова маршрутних таблиць за комбінованою метрикою, що ґрунтується на пропускній здатності шляху та кількості транзитних вузлів

Вхідними даними для побудови маршрутних таблиць з врахуванням пропускних здатностей каналів є топологія мережі та пропускні здатності її каналів. Отримане за алгоритмом Дейкстри дерево шляхів є оптимізовано за критерієм максимальної пропускної здатності сукупного шляху.

Для побудови маршрутних таблиць маршрутизаторів з урахуванням пропускних здатностей каналів, а також кількості проміжних вузлів додаткових вхідних даних не потрібно. Щоб мінімізувати кількість проміжних вузлів на шляхах, достатньо як вхідні дані для алгоритму Дейкстри прийняти, що вага всіх ребер графу, що відповідає топології мережі, дорівнює одиниці. Критерієм оптимізації в цьому випадку буде кількість переходів (хопів) на шляху від вузла джерела до вузла призначення.

Записи маршрутної таблиці пропонується формувати на основі аналізу шляхів, знайдених за алгоритмом Дейкстри як для метрики M1, так і для метрики M2.

Проілюструємо вищеописаний підхід на прикладі мережі, показаної на рис. 1.

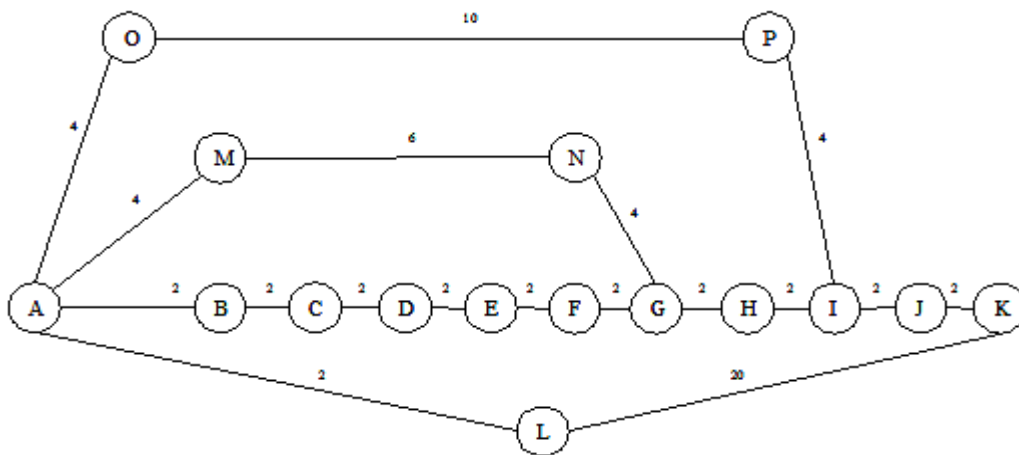


Рис. 1. Топологія мережі, ваги каналів якої враховують пропускні здатності каналів

За алгоритмом Дейкстри для наведеної мережі сформуємо для вузла А таблицю шляхів (рис. 2) згідно з метрикою 1, що враховує пропускні здатності каналів:

Вершини	Шлях	L1	N1
A	A	0	0
B	AB	2	1
C	ABC	4	2
D	ABCD	6	3
E	ABCDE	8	4
F	ABCDEF	10	5
G	ABCDEFG	12	6
H	ABCDEFGH	14	7
I	ABCDEFGHI	16	8
J	ABCDEFGHIJ	18	9
K	ABCDEFGHIJK	20	10
L	AL	2	1
M	AM	4	1
N	AMN	10	2
O	AO	4	1
P	AOP	14	2

Рис. 2. Таблиця шляхів від вузла А до всіх інших вузлів, знайдена за метрикою 1

У таблиці шляхів рис. 2: *L1* – критерій оптимізації (вага сукупного шляху, отриманого за метрикою 1, що ґрунтується на пропускній здатності каналів); *N1* – кількість переходів (хопів) на шляху до віддаленого вузла.

Сформуємо для мережі, наведеної на рис. 1, таблицю шляхів від вузла А згідно з метрикою 2, в якій повинна бути мінімізована кількість хопів на шляху між взаємодіючими вузлами. Оскільки

оптимізація буде здійснюватись за критерієм мінімальної кількості проміжних вузлів (метрика 2), то ваги каналів, пов'язані з пропускними здатностями каналів, у цьому випадку не беруть участі в роботі алгоритму Дейкстри. Вага кожного каналу приймається рівною одиниці. В результаті отримаємо таблицю шляхів (рис. 3):

Вершини	Шлях	L2	N2
A	A	0	0
B	AB	2	1
C	ABC	4	2
D	ABCD	6	3
E	ABCDE	8	4
F	AMNGF	16	4
G	AMNG	14	3
H	AMNGH	16	4
I	AOP I	18	3
J	ALKJ	24	3
K	ALK	22	2
L	AL	2	1
M	AM	4	1
N	AMN	10	2
O	AO	4	1
P	AOP	14	2

Рис. 3. Таблиця шляхів, знайдена за метрикою 2 від вузла А до всіх інших вузлів

У таблиці шляхів (рис. 3) є такі позначення: N2 – критерій оптимізації (метрика 2, що визначає кількість проміжних вузлів на шляху між взаємодіючими вузлами); L2 – вага сукупного шляху, що є сумою ваг окремих шляхів, які ґрунтуються на пропускних здатностях каналів мережі.

Як видно з аналізу рис. 2 та 3, таблиці шляхів, отримані за різними критеріями, різні. Відповідно, побудовані на їх основі маршрутні таблиці маршрутизаторів будуть також різними.

Як приклад проаналізуємо шляхи від вузла А до вузла К, отримані за двома метриками.

Згідно з алгоритмом Дейкстри за метрикою 1 отримано такий шлях: А, В, С, D, E, F, G, H, I, J, К, що має мінімально можливу для даної мережі вагу сукупного шляху, що дорівнює 20 та проходить через 9 проміжних вузлів.

Для ілюстрації правил формування маршрутних таблиць присвоїмо вузлам мережі (рис. 1) такі IP-адреси:

Таблиця 1

Відповідності IP-адрес деяким інтерфейсам вузлів мережі

A-O	10.0.0.1	O-A	10.0.0.2	K-J	50.0.0.2	G-F	80.0.0.3
A-M	20.0.0.1	M-A	20.0.0.2	C-B	60.0.0.2	I-P	90.0.0.2
A-B	30.0.0.1	B-A	30.0.0.2	F-E	70.0.0.2	I-H	100.0.0.2
A-L	40.0.0.1	L-A	40.0.0.2	G-N	80.0.0.2		

У маршрутній таблиці вузла А, що використовуватиме дерево шляхів, знайдене за метрикою 1, для наведеного прикладу (шляху від вузла А до вузла К) буде введено запис (табл. 2):

Таблиця 2

Фрагмент маршрутної таблиці вузла А, побудованої за метрикою 1

Destination node	Gateway	Interface	Metrica
50.0.0.0	30.0.0.2	30.0.0.1	20

У цій таблиці: Destination node – IP адреса вузла призначення; Gateway – IP адреса наступного вузла; Interface – IP адреса порту вузла А, на який буде відправлено пакет; Metrica – вага шляху від поточного вузла до вузла призначення, знайденого за критерієм максимізації пропускної здатності сукупного шляху.

Згідно з алгоритмом Дейкстри за метрикою 2 отримано цілком інший шлях від вузла А до вузла К. Це шлях А, L, К, що має всього 1 проміжний вузол на шляху та вагу сукупного шляху, що дорівнює 22.

У маршрутній таблиці вузла А, що використовуватиме дерево шляхів, знайдене за метрикою 2, для наведеного прикладу буде введено інший запис:

Таблиця 3

Фрагмент маршрутної таблиці вузла А, побудованої за метрикою 2

Destination node	Gateway	Interface	Metrica
50.0.0.0	40.0.0.2	40.0.0.1	22

Наведений приклад ілюструє значну відмінність шляхів між вузлами А та К, отриманих за різними метриками, параметри яких наведено у табл. 4.

Таблиця 4

Параметри шляхів між вузлами А та К за двома метриками

	Метрика	Кінцеві вузли		Вага шляху, що ґрунтується на пропускній здатності	Кількість проміжних вузлів
1	Метрика 1	А	К	20	9
2	Метрика 2	А	К	22	1

З наведених результатів видно, що шлях, оптимізований за критерієм мінімізації кількості проміжних вузлів, має пропускну здатність, лише на 10 % меншу від максимально можливої. При цьому кількість проміжних вузлів шляху зменшилась аж у 9 разів, замість дев'яти проміжних маршрутизаторів один. Очевидно, що це значно зменшить час затримки, зумовлений обробкою пакетів в маршрутизаторах, а також зменшить шанс потрапляння пакетів у перевантажені маршрутизатори. Платою за це є зменшення пропускної здатності на 10 відсотків. Тому, якщо для прикладного застосування така пропускну здатність є достатньою, доцільно в результуючу маршрутну таблицю для шляху між вузлами А та К включити відповідний запис з табл. 3.

У загальному випадку для формування записів кінцевої маршрутної таблиці, що будується за комбінованою метрикою, необхідно сформулювати правила їх вибору з таблиць, отриманих за метриками 1 та 2.

Рішення про вибір шляху за комбінованою метрикою приймається тільки за наявності альтернативних шляхів, тобто у випадку, коли шляхи до одного і того самого кінцевого вузла, знайдені за метриками 1 та 2, відрізняються між собою. Аналізуємо ці шляхи та обираємо той, що є ефективніший з погляду одночасного врахування двох параметрів. Цей шлях записуємо до кінцевої таблиці шляхів.

Для аналізу ефективності шляху введемо такі параметри:

- $K1$ – відношення ваг шляхів знайдених за метриками 2 та 1;
- $K2$ – відношення кількості проміжних вузлів на шляхах, знайдених за метриками 1 та 2.

Для $L1 > 0$ та $N2 > 1$ відношення $K1$ та $K2$ можна записати:

$$K1 = L2/L1, \tag{1}$$

$$K2 = (N1-1)/(N2-1). \tag{2}$$

Для вибору результуючого маршруту потрібно визначити значення $K1_{max}$ та $K2_{min}$. У кожній конкретній ситуації значення цих коефіцієнтів можна вибрати різне. При побудові результуючої маршрутної таблиці метрику 1 можна використовувати за замовчуванням, а коли виконуються одночасно умови (3) та (4):

$$K1 \leq K1_{max}, \tag{3}$$

$$K2 \geq K2_{min}, \tag{4}$$

відповідні записи результуючої таблиці замінюються на записи, знайдені за метрикою 2.

Для прикладу приймемо $K1_{max} = 1,2$ та $K2_{min} = 1,5$. Оцінивши за такими критеріями шляхи з двох таблиць шляхів (рис. 2 та рис. 3), отриманих за різними метриками, сформуємо кінцеву таблицю шляхів (рис. 4) вузла А:

Вершини	Шлях	L3	N3
A	A	0	0
B	AB	2	1
C	ABC	4	2
D	ABCD	6	3
E	ABCDE	8	4
F	ABCDEF	10	5
G	AMNG	14	3
H	AMNGH	16	4
I	AOP	18	3
J	ABCDEFGHIJ	18	9
K	ALK	22	2
L	AL	2	1
M	AM	4	1
N	AMN	10	2
O	AO	4	1
P	AOP	14	2

Рис. 4. Таблиця шляхів вузла А для комбінованої метрики

У таблиці шляхів (рис. 4) є такі позначення: N3 – кількість хопів на шляху між взаємодіючими вузлами; L3 – вага сукупного шляху, що є сумою ваг окремих шляхів, які основані на пропускних здатностях каналів мережі.

На основі отриманої таблиці шляхів можна сформувати маршрутну таблицю для вузла А. Фрагмент маршрутної таблиці, побудованої за комбінованою метрикою, показано на рис. 5:

Destination node	Gateway	Interface	Metrica L3	Metrica N3
60.0.0.0	30.0.0.2	30.0.0.1	4	2
80.0.0.0	20.0.0.2	20.0.0.1	14	3
100.0.0.0	10.0.0.2	10.0.0.1	18	3
50.0.0.0	40.0.0.2	40.0.0.1	22	2

Рис. 5. Фрагмент маршрутної таблиці вузла А для комбінованої метрики

Висновок

Запропоновано модифікований алгоритм побудови маршрутних таблиць, що формує шляхи передавання пакетів з врахуванням двох характеристик мережі: пропускної здатності каналів та кількості проміжних вузлів на шляхах. Вибір шляхів з меншою кількістю проміжних вузлів за достатньої пропускної здатності зменшує затримки пакетів у мережі.

Запропонований підхід можна розширити, додаючи нові метрики, що характеризують інші параметри мережі. При внесенні нових метрик необхідно вносити додаткові дані про мережу і додаткові коефіцієнти оцінювання ефективності шляхів, отриманих за різними метриками.

1. *OSPF Design Guide*. – [Електронний ресурс]. Режим доступу: <http://www.cisco.com/c/en/us/support/docs/ip/open-shortest-path-first-ospf/7039-1.html#16>. 2. Олифер В.Г., Олифер Н.А. *Компьютерные сети. Принципы, технологии, протоколы: Учебник для вузов*. 4-е изд. – СПб.: Питер, 2012. – 944 с. 3. Кузнецов Н.А., Фетисов В.Н. *Алгоритм Дейкстры с улучшенной робастностью для управления маршрутизацией в IP-сетях. Автоматика и телемеханика*. – 2008. – № 2. – С. 80–85. 4. Чуба І. В. *Метод маршрутизації у гетерогенних комп'ютерних мережах на основі аналізу ієрархій: автореф. дис. на здобуття наук. ступеня канд. техн. наук: 05.13.05 / Чуба Ірина Вікторівна; Національний авіаційний університет*. – К.: [б. в.], 2008. – 20 с. 5. Обельовська К., Русаков А. *Модифікований алгоритм маршрутизації для зменшення перевантажень ресурсів комп'ютерних мереж. Вісник НУ "Львівська політехніка" N 663 "Комп'ютерні науки та інформаційні технології"*. – Львів, 2010. – С. 223 – 228.

SELF-ORGANIZING MAP AND ITS LEARNING IN THE FUZZY CLUSTERING-CLASSIFICATION TASKS

© *Bodyanskiy Ye., Vynokurova O., Mulesa P., Slipchenko O., 2014*

Запропоновано комбінований метод самонавчання-навчання самоорганізованої мапи (SOM-LVQ), що дає змогу підвищити якість обробки інформації за умов класів, що перетинаються внаслідок раціонального вибору параметра кроку навчання і введення спеціальної процедури нечіткого виведення в процесі класифікації-кластеризації, який проходить як з зовнішнім навчальним сигналом, так і без нього. Як міру подібності функцій сусідства і належності використовуються косинусоїдальні конструкції, що дають змогу забезпечити процесам самонавчання-навчання більшу гнучкість і надати їм низку нових корисних властивостей.

Ключові слова: комбінований метод самонавчання-навчання самоорганізованої мапи, кластеризація, класифікація, класи, що перетинаються.

In the paper, combined self-learning and learning method of self-organizing map (SOM-LVQ) is proposed. Such method allows to increase quality of information processing under condition of overlapping classes due to rational choice of learning rate parameter and introducing special procedure of fuzzy reasoning in the clustering-classification process, which occurs both with external learning signal (“supervised”), and without one (“unsupervised”). As similarity measure of neighborhood function or membership one, cosine structures are used, which allow to provide a high flexibility due to self-learning-learning process and to provide some new useful properties.

Key words: combined self-learning and learning method of self-organizing map, clustering, classification, overlapping classes.

Introduction

Self-organizing maps (SOM) and neural networks of learning vector quantization (LVQ) have seen extensive use for solving different problems in Data Mining domain (clustering, classification, fault detection and compression of information to name a few). This type of neural networks was proposed by T. Kohonen [1, 2] and represents, in fact, a single-layer feedforward architecture, which provides an operator for mapping of input space into the output space. Operation-wise SOM and LVQ are quite similar: each neuron is fed input signal (sample) producing output, which is used during competition stage to determine winning neuron – usually the one with maximum output signal value. Vector of synaptic weights for winning neuron is the one closest to the input sample in terms of the metric chosen (which is Euclidian metric in most cases). Next is neurons adjustment phase. Synaptic weights of the winning neuron gets moved closer to input sample. Alternatively, a subset of neurons (rather than a single one) can be adjusted – those determined to be “reasonably close” to the input sample are updated. Resulting network output, however, is determined exclusively by winning neuron (this principle is usually referred to as “Winner-Takes-All” (WTA)). It is this principle (WTA) which negatively affects accuracy in case when there are overlapping clusters in underlying data.

Taking into account the above mentioned properties of SOM and LVQ networks, it makes sense to introduce fuzzy classification capabilities on top of them, while preserving online operation. In [8, 9] fuzzy

self-organizing map was proposed, in which conventional neurons are replaced by fuzzy rules. This neural network shows enough high efficiency, but its learning properties was significantly lost especially in on-line mode. In [5, 10, 11] fuzzy clustering Kohonen network and fuzzy linear vector quantization network are described. In fact such networks are neural representation of fuzzy c-means (FCM) [3], which is far enough from SOM and LVQ mathematical tool and designed for operation in batch mode.

Problem statement

Let us consider single-layer neural network with lateral connections containing receptors and neurons in the Kohonen layer with each neuron being characterized by a vector of its synaptic weights. During learning stage input vector is fed to the inputs of all neurons (usually adaptive linear associators) (here – either the number of observation in a table “object-properties”, or current discrete time for on-line processing mode) and neurons produce the scalar signals on their outputs

$$y_j(k) = w_j^T(k)x(k), \quad j = 1, 2, \dots, m. \quad (1)$$

Note that neuron’s output depends on current values of synaptic weights vector, assuming iterative learning algorithm.

Each input vector can activate either single neuron (w_j) or a set of neighboring neurons – this also depends on learning algorithm chosen.

Self-organization procedure is based on the competitive learning approach (self-learning), and begins with the initialization of synaptic weights. Selecting initial values for weights from a uniform random distribution over input space has become a common practice. In addition, weights are usually normalized to reside on unit hypersphere:

$$\|w_j(k)\|^2 = w_j^T(k)w_j(k) = 1. \quad (2)$$

Self-organizing is usually divided into three stages [2]: competition, cooperation and synaptic adaptation. This paper introduces additional stage to this process – fuzzy inference, which allows an online learning algorithm to classify data samples belonging to several classes simultaneously.

Learning algorithm for SOM

The competition process is started to analysis of current pattern $x(k)$, which is fed to all neurons of Kohonen’s layer from receptive (zero) layer. For each neuron the distance between input sample and a vector of synaptic weights is computed:

$$D(w_j(k), x(k)) = \|x(k) - w_j(k)\|. \quad (3)$$

In case when all inputs and synaptic weights are normalized according to (2), i.e.

$$\|w_j(k)\| = \|x(k)\| = 1, \quad (4)$$

and Euclidian metric is used, the proximity measure between the vectors $w_j(k)$ and $x(k)$ can be in the following way:

$$w_j^T(k)x(k) = y_j(k) = \cos(w_j(k), x(k)) = \cos q_j(k). \quad (5)$$

So the expression (3) takes the form

$$D(w_j(k), x(k)) = \sqrt{2(1 - y_j(k))}, \quad (6)$$

and type of dependence $D(w_j(k), x(k))$ from output signal value $y_j(k)$ is shown on fig. 1.

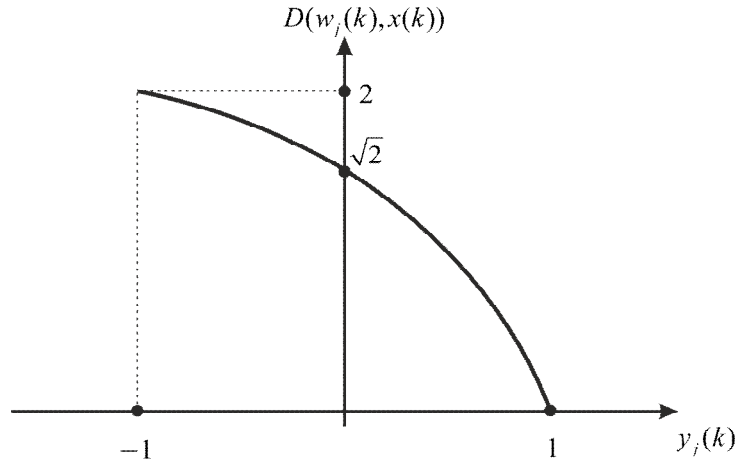


Fig. 1. Dependence of distance $D(w_j(k), x(k))$ from output signal level for j -th neuron

Using relation (5), we can replace metric (3) with a simpler construction, usually referred to as a measure of similarity [12]

$$y(w_j(k), x(k)) = \max\{0, \cos(w_j(k), x(k))\} = \max\{0, \cos q_j(k)\}, \quad (7)$$

which satisfies softer conditions compared to classic measure requirements:

$$\begin{aligned} y(w_j(k), x(k)) &\geq 0, \\ y(w_j(k), w_j(k)) &= 1, \\ y(w_j(k), x(k)) &= y\{x(k), w_j(k)\} \end{aligned} \quad (8)$$

and which has a form shown on fig 2.

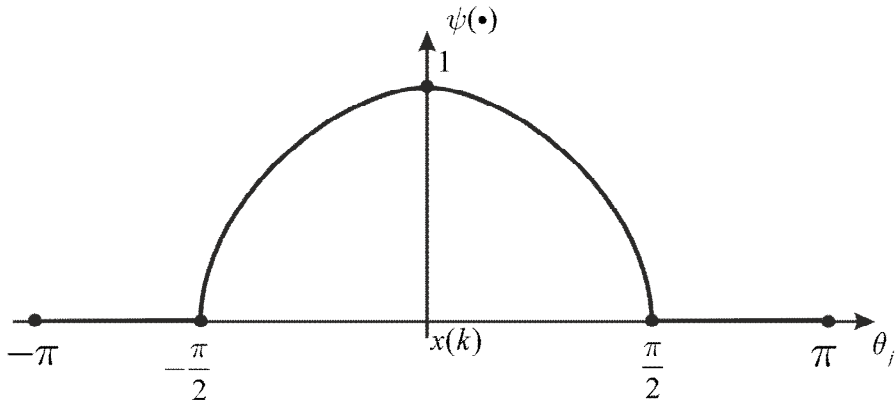


Fig.2. The measure of similarity (7)

Next we look for a winning neuron that has a biggest value of similarity to the input vector in the sense (6) or (7), i. e.

$$D(w_j^*(k), x(k)) = \min_i D(w_i(k), x(k)) \quad (9)$$

or

$$y(w_j^*(k), x(k)) = \max_i y(w_i(k), x(k)). \quad (10)$$

After that we tune values of synaptic weights using WTA self-learning rule is realized in form

$$w_j(k+1) = \begin{cases} w_j^*(k) + h(k)(x(k) - w_j^*(k)), & \text{if } j\text{-th neuron won in the competition,} \\ w_j(k) & \text{otherwise.} \end{cases} \quad (11)$$

In a nutshell, expression (11) means that synaptic weights vector of winner is moved closer to the input pattern $x(k)$ by a value which is proportional to learning rate $0 < h(k) < 1$.

Choice of learning rate value $h(k)$ defines the convergence rate of self-learning process, and is usually based on empirical reasons, at that the parameter must monotonically decrease in on time.

It is easy to see, that first relation of the rule (11) minimizes the following criterion

$$E_j^k = \sum_{t=1}^{k_j} \|x(t) - w_j\|^2 \quad (12)$$

(here k_j is a number of data in the sampling with dimensions k , when j -th neuron was winner), notably in practice as synaptic weights estimation conventional arithmetic mean is used:

$$w_j(k) = \frac{1}{k_j} \sum_{t=1}^{k_j} x(t). \quad (13)$$

Therefore, in fact self-learning rule (3) is stochastic approximation procedure [13], and learning rate parameter $h(k)$ must be selected according to the condition of A. Dvoretzky [14]. It is know that using the following value

$$h(k) = \frac{1}{k_j} \quad (14)$$

leads to slowdown of the learning process.

Requirement of monotonic decreasing of the parameter $h(k)$ leads to expression in form

$$h(k) = r^{-1}(k), \quad r(k) = ar(k-1) + \|x(k)\|^2, \quad 0 \leq a \leq 1. \quad (15)$$

When taking into account normalization to unit hypersphere (4) we have:

$$r(k) = ar(k-1) + 1, \quad (16)$$

Which with $a=1$ gives a well-known expression, that was introduced in [15].

Changing the forgetting parameter a provides enough variance for the learning rate to both satisfy Dvoretzky condition and adjust for characteristics of different data sets:

$$\frac{1}{k_j} \leq h(k) \leq 1. \quad (17)$$

Note that adjusting synaptic weights with (13) in genera breaks normalization (4). In order to maintain it, we should slightly modify our learning algorithm:

$$w_j(k+1) = \begin{cases} \frac{w_j^*(k) + h(k)(x(k) - w_j^*(k))}{\|w_j^*(k) + h(k)(x(k) - w_j^*(k))\|}, & \text{if } j\text{-th neuron won in the competition,} \\ h(k) = r^{-1}(k), \quad r(k) = ar(k-1) + 1, \quad 0 \leq a \leq 1, \\ w_j(k), & \text{otherwise.} \end{cases} \quad (18)$$

In order to better adjust to input data, a so called “cooperation stage” is frequently introduced to SOM learning process. During this stage a winning neuron defines a local domain of topological neighborhood, where weights are adjusted for a set of neurons rather than only for a winning one. Those neurons closer to the winner receive a bigger adjustment.

This topological domain is defined by neighborhood function $j(j, p)$, which depends on distance $D(w_j^*(k), w_p(k))$ between winner $w_j^*(k)$ and any of Kohonen’s layer neurons $w_p(k)$, $p=1, 2, \mathbf{K}, m$ and some parameter s , which sets its width.

Usually $j(j, p)$ is the bell-shaped function, symmetrical with respect to the maximum in point $w_j^*(k)$ ($j(j, j)=1$) and decreasing when distance $D(w_j^*(k), w_p(k))$ increases. Gaussian function is commonly used here:

$$j(j, p) = \exp\left(-\frac{\|w_j^*(k) - w_p(k)\|^2}{2s^2}\right). \quad (19)$$

Adding neighborhood function results in the following learning algorithm:

$$w_p(k+1) = w_p(k) + h(k)j(j, p)(x(k) - w_p(k)), \quad p=1, 2, \mathbf{K}, m, \quad (20)$$

which minimizes criterion

$$E_p^k = \sum_{t=1}^k j(j, p) \|x(t) - w_p\|^2 \quad (21)$$

This criterion is commonly referred to as “Winner Takes More” (WTM).

The necessity to maintain normalization to unit hypersphere (4) leads to the expression in form

$$\begin{cases} w_p(k+1) = \frac{w_p(k) + h(k)j(j, p)(x(k) - w_p(k))}{\|w_p(k) + h(k)j(j, p)(x(k) - w_p(k))\|}, \quad p=1, 2, \mathbf{K}, m, \\ h(k) = r^{-1}(k), \quad r(k) = ar(k-1) + 1, \quad 0 \leq a \leq 1. \end{cases} \quad (22)$$

It is possible to skip the entire competition by tuning synaptic weights of network depending on their similarity with the current vector-pattern $x(k)$. In this case instead of conventional neighborhood function $j(j, p)$ depending on winner $w_j^*(k)$, we can use similarity measure (7), that depends on $x(k)$.

In this case instead of (20) we can use the rule in form:

$$\begin{aligned} w_p(k+1) &= w_p(k) + h(k)y(w_p(k), x(k))(x(k) - w_p(k)) = \\ &= w_p(k) + h(k) \max\{0, \cos(w_p(k), x(k))\}(x(k) - w_p(k)) = \\ &= w_p(k) + h(k) \max\{0, w_p^T(k)x(k)\}(x(k) - w_p(k)) = \\ &= w_p(k) + h(k) \max\{0, \cos q_p(k)\}(x(k) - w_p(k)) = \\ &= w_p(k) + h(k) \max\{0, y_p(k)\}(x(k) - w_p(k)), \end{aligned} \quad (23)$$

which reminds “INSTAR” learning rule of S. Grossberg [16].

For maintaining (4) the rule (23) has to be rewritten in the form

$$\begin{cases} w_p(k+1) = \frac{w_p(k) + h(k)y(w_p(k), x(k))(x(k) - w_p(k))}{\|w_p(k) + h(k)y(w_p(k), x(k))(x(k) - w_p(k))\|}, \\ h(k) = r^{-1}(k), \quad r(k) = ar(k-1) + 1, \quad 0 \leq a \leq 1. \end{cases} \quad (24)$$

In many real world problems clusters can overlap as shown on fig 3, where * denotes patterns, belonging to j -th clusters, and o - p -th ones. In this case vector $x(k)$ belongs j -th cluster with proportional membership level $\cos q_j(k)$, and with proportional level $\cos q_p(k)$ - to p -th one. Synaptic weights concentrated in crosshatched region, don't have relationship to the pattern $x(k)$ according to (7).

Using similarity measure, that is shown on fig.2, we can introduce the membership estimate of pattern $x(k)$ to j -th class in form:

$$0 \leq m_{w_j(k)}(x(k)) = \frac{y(w_j(k), x(k))}{\sum_{l=1}^m y(w_l(k), x(k))} \leq 1. \quad (25)$$

Learning algorithm for LVQ

Learning vector quantization neural networks in contrast to self-learning SOM adjust their parameters based on external learning (reference) signal, which fixes the membership of each pattern $x(k)$ to a particular class.

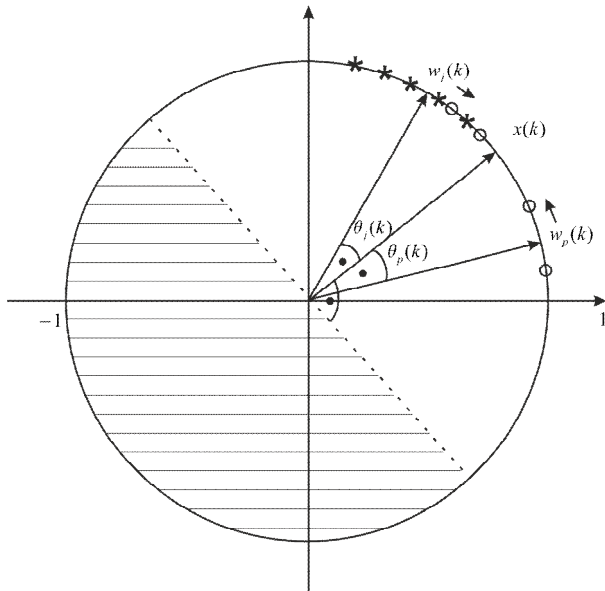


Fig. 3. The overlapping clusters

The main idea of LVQ neural network is to build a compact representation of large data set based on a restricted set of prototype samples (centroids) $w_j(k)$, $j=1,2,\mathbf{K},m$, that provide sufficiently accurate approximation of the original space X .

For each input vector $x(k)$ which is normalized according to (4), we look for a winning neuron $w_j^*(k)$ that corresponds to a centroid of a certain class. In other words, winner is defined as a neuron with minimal distance to the input vector (9) or, which is the same, with maximal similarity measure (10).

Since the learning is supervised, membership of the vector $x(k)$ to specific domain X_j of the space X is known, which allows considering two typical situation occurring in the trained linear vector quantization:

- the input vector $x(k)$ and neuron-winner $w_j^*(k)$ belong to the same cell of Voronoy [2];
- the input vector $x(k)$ and neuron-winner $w_j^*(k)$ belong to the different cells of Voronoy.

Than corresponding learning LVQ-rule can be written in form:

$$w_j(k+1) = \begin{cases} w_j^*(k) + \mathbf{h}(k)(x(k) - w_j^*(k)), & \text{if } x(k) \text{ and } w_j^*(k) \text{ belong to the same cell,} \\ w_j^*(k) - \mathbf{h}(k)(x(k) - w_j^*(k)), & \text{if } x(k) \text{ and } w_j^*(k) \text{ belong to the different cells,} \\ w_j(k) & \text{for neurons, which aren't won in instant } k. \end{cases} \quad (26)$$

The rule (26) has a clear physical interpretation: if the winning neuron and presented sample belong to the same class, than prototype $w_j^*(k)$ is moved to $x(k)$; otherwise prototype $w_j^*(k)$ is pushed away from $x(k)$, effectively increasing the distance $D(w_j^*(k), x(k))$, i.e. solving the maximization task based on criterion (12).

As for the choice of learning rate parameter $\mathbf{h}(k)$, that common recommendations are the same that for SOM, i.e. the learning rate parameter must monotonically decrease in process controlled learning.

In [17] it was proved that LVQ tuning algorithm converges in case of learning rate $\mathbf{h}(k)$ satisfies Dvoretzky conditions. This, in turn, allows choosing $\mathbf{h}(k)$ according to Goodwin-Ramadge-Caines algorithm [13], or in the previously defined form (16) with $a = 1$, which is essentially the same. When $a < 1$, the algorithm gets tracking properties and handles the case when class centroids are drifting.

For providing to fulfillment of the condition (4) it possible to introduce modification of rule (26) in the form (18):

$$w_j(k+1) = \begin{cases} \frac{w_j^*(k) + \mathbf{h}(k)(x(k) - w_j^*(k))}{\|w_j^*(k) + \mathbf{h}(k)(x(k) - w_j^*(k))\|}, & \text{if } x(k) \text{ and } w_j^*(k) \text{ belong to the same cell,} \\ \frac{w_j^*(k) - \mathbf{h}(k)(x(k) - w_j^*(k))}{\|w_j^*(k) - \mathbf{h}(k)(x(k) - w_j^*(k))\|}, & \text{if } x(k) \text{ and } w_j^*(k) \text{ belong to the different cells,} \\ w_j(k) & \text{for neurons, which are not won in instant } k. \end{cases} \quad (27)$$

First and third expressions in formula (27) are completely identical to WTA – self-learning algorithm, which the second one represents a “push-back” scenario.

Let's consider a situation, shown on fig. 4, when neuron $w_j^*(k)$ won the competition. At the same time current sample $x(k)$ belongs to a class with different centroid $w_p(k)$. Now we need to “push” vector $w_j^*(k)$ away so, that $x(k)$ was equidistant from $w_j^*(k)$ and from $w_p(k)$.

Applying elementary transformations, we get the following formulas:

$$w_j(k+1) = w_j^*(k) - \mathbf{h}(k)(x(k) - w_j^*(k)), \quad (28)$$

$$x^T(k)w_j(k+1) = x^T(k)w_j^*(k) - \mathbf{h}(k)\|x(k)\|^2 - \mathbf{h}(k)x^T(k)w_j^*(k), \quad (29)$$

$$\cos(w_j(k+1), x(k)) = \cos(w_j^*(k), x(k)) - \mathbf{h}(k)(1 + \cos(w_j^*(k), x(k))) \quad (30)$$

In order to satisfy

$$\cos(w_j(k+1), x(k)) = \cos(w_p(k), x(k)) \quad (31)$$

We need to set $h(k)$ in form

$$\begin{aligned} h(k) &= \frac{\cos(w_j^*(k), x(k)) - \cos(w_p(k), x(k))}{\cos(w_j^*(k), x(k)) + 1} = \frac{\cos(w_j^*(k), x(k)) - \cos(w_p(k), x(k))}{\cos(w_j^*(k), x(k)) + \cos(x(k), x(k))} = \\ &= \frac{\cos q_j(k) - \cos q_p(k)}{\cos q_j(k) + 1} = \frac{x^T(k)w_j^*(k) - x^T(k)w_p(k)}{x^T(k)w_j^*(k) - x^T(k)x(k)}. \end{aligned} \quad (32)$$

After one step of the weights tuning the pattern $x(k)$ will belong equally to both $w_j(k+1)$ and $w_p(k) = w_p(k+1)$, i.e.

$$m_{w_j(k+1)}(x(k)) = m_{w_p(k+1)}(x(k)) = 0,5. \quad (33)$$

In case when several classes are overlapping, we can use estimate (25) for computing membership levels.

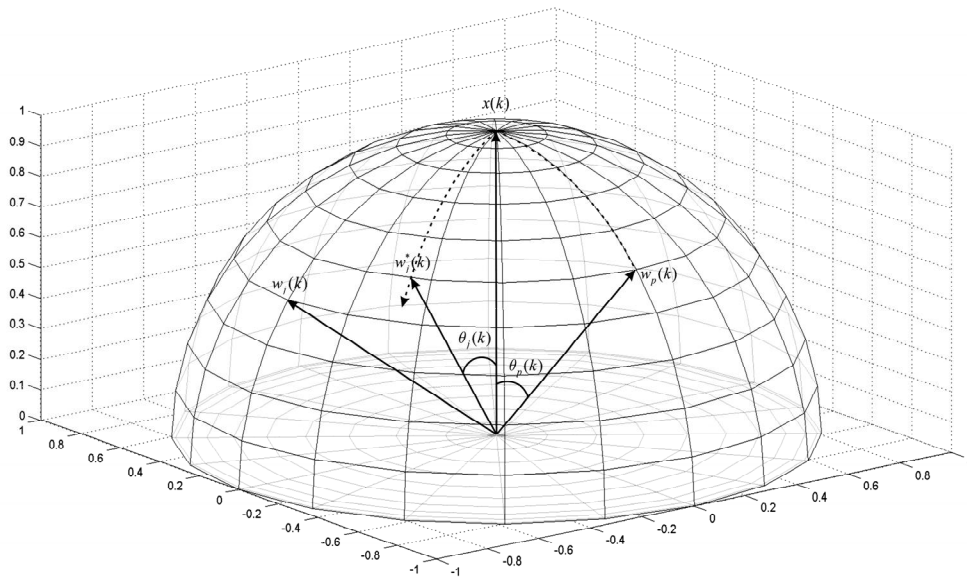


Fig. 4. Learning of LVQ

Joining SOM and LVQ learning algorithms

A promising application of Kohonen neural network is adaptive pattern recognition, which implemented by the systems. The implementation consists of sequentially connected SOM and LVQ [2] layers. First network operates in self-learning mode, while the second one adds supervised component to the process. From an input vector $x(k)$ which is fed to the input system, SOM network extracts a feature sample $y(k)$ in a space with sufficiently reduced dimensionality. This simplifies a problem in hand without significant loss of information. On the second stage LVQ is trained to classify of incoming pattern $y(k)$ using supervised learning. A distinctive feature of proposed system is the connection of two identical architectures, that are tuned by different learning algorithms.

For many data mining problems, especially in healthcare domain, substantial shares of input samples lack clear class association. Indeed, a diagnosis might be either missing, be ambiguous or yet unknown.

In this case it is possible to tune the synaptic weights with unified architecture with lateral connections using different learning methods. Each learning method is initialized according to the level of apriori information about $x(k)$ available.

A resulting learning algorithm for combined (SOM+LVQ) neural network can be presented in the following form:

$$w_j(k+1) = \begin{cases} \frac{w_j^*(k) + \mathbf{h}(k)(x(k) - w_j^*(k))}{\|w_j^*(k) + \mathbf{h}(k)(x(k) - w_j^*(k))\|}, \\ d_L(k) \frac{w_j^*(k) - \mathbf{h}(k)(x(k) - w_j^*(k))}{\|w_j^*(k) - \mathbf{h}(k)(x(k) - w_j^*(k))\|}, \\ d_L = \begin{cases} 1, & \text{if the network works in supervised mode} \\ & \text{and sample } x(k) \text{ does not belong to class } j, \\ 0, & \text{otherwise,} \end{cases} \\ w_j(k) & \text{for neurons, which did not get activated by sample } x(k). \end{cases} \quad (34)$$

where $w_j^*(k)$ is winning neuron.

It is clear, that for $d_L = 0$, i.e. in self-learning mode, the first expression of formula (34) can be replaced by expression (24).

Conclusion

The combined self-learning procedure for Kohonen neural network is proposed. Such method allows data processing under the overlapping classes condition, when memberships of training data to specific classes can be unknown at all, and have both crisp and fuzzy nature. This method is based on using similarity measure, recurrent optimization and fuzzy inference and differs with high speed, possibility of operating in on-line mode and simplicity of computational realization.

1. Kohonen, T. *Self-Organizing Maps*. – Berlin: Springer-Verlag, 1995 – 362 p.
2. Haykin, S. *Neural Networks: A Comprehensive Foundation*. - Upper Saddle River, N.Y.: Prentice Hall, 1999. – 842 p.
3. Bezdek, J.C. *Pattern Recognition with Fuzzy Objective Function Algorithms*. - New York: Plenum Press, 1981. – 272 p.
4. Hoepfner, F., Klawon, T., Kruse, R. *Fuzzy Clusteranalyse: Verfahren fuer die Belderkennung, Klassifikation und Datenanalyse*. – Braunschweig: Vieweg, Reihe Computational Intelligence, 1996. – 280 S.
5. Bezdek, J.C., Keller, J., Krisnapuram, R., Pal, N. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*. Springer, 1999 – 777 p.
6. Hoepfner F., Klawonn F., Kruse R., Runkler T. *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition*. - Chichester: John Wiley & Sons, 1999. – 300 p.
7. Sato-Ilic M., Jain L. *Innovations in Fuzzy Clustering: Theory and Applications*. – Berlin: Springer, 2006. – 152 p.
8. Vuorimaa, P. *Use of the fuzzy self-organizing map in pattern recognition // Proc. 3-rd IEEE Int. Conf. Fuzzy Systems "FUZZ-IEEE'94"*. - Orlando, USA, 1994 – P.798-801.
9. Vuorimaa, P. *Fuzzy self-organizing map // Fuzzy Sets and Systems*. - 1994. - 66. - P. 223-231.
10. Tsao E.C.-K., Bezdek J.C., Pal N.R. *Fuzzy Kohonen clustering networks // Pattern recognition*. – 1994. – 27. – P. 757-764.
11. Pascual-Marqui R.D., Pascual-Montano A.D., Kochi K., Caraso J.M. *Smoothly distributed fuzzy C-means: a new self-organizing map // Pattern Recognition*. – 2001. – 34. – P. 2395-2402.
12. Sepkovski, J.J. *Quantified coefficients of association and measurement of similarity // J. Int. Ass. Math.* - 1974. – 6 (2). – P. 135-152.
13. Wasan M.T. *Stochastic Approximation*. – Cambridge: The University Press, 2004. – 216 p.
14. Dvoretzky, A., *On stochastic approximation // Proc. 3-rd Berkley Symp. Math. Statistics and Probability*. – 1956. – 1. – P. 39-55.
15. Goodwin, G.C., Ramadge, P.J., Caines, P.E. *A globally convergent adaptive predictor // Automatica*. – 1981 – 17 (1). – P.135-140.
16. Grossberg S. *Classical and instrumental learning by neural networks // In "Progress in Theoretical Biology"*. – N.Y.: Academic Press, 1974. – 3 – P. 51-141.
17. Baras J.C., La Vigna A. *Convergence of Kohonen's learning vector quantization // Proc. Int. Joint Conf. on Neural Networks*. – San Diego, CA, 1990. – 3. – P.17-20.

ПРОЕКТУВАННЯ СИСТЕМИ АВТОМАТИЗАЦІЇ ПРОЦЕСУ SEO-ОПТИМІЗАЦІЇ

© Басюк Т., 2014

Проаналізовано передумови виникнення процесу seo-оптимізації, описано основні види послуг, які надаються підприємствами із просування інтернет-ресурсів та проаналізовано основні етапи роботи оптимізаторів. Обґрунтовано методологічні засади побудови автоматизованої системи Seo-оптимізації та сформульовано висновки щодо можливості її проектування із використанням сучасних технологій програмування.

Ключові слова: структура, seo-оптимізація, веб-сайт, інтернет.

This paper analyzes the background of the process of SEO-optimization, carried description of the main types of services provided to businesses with promotion of internet resources and analyzes the main stages of SEO. The methodological basis for building an automated system optimization and Seo-formulated conclusions on the possibility of its design using modern programming technologies.

Key words: structure, seo-optimization, web site, internet.

Вступ

Seo-оптимізація – дії, скеровані на поліпшення видимості сайту в пошукових системах за тематичними запитами, користувачів і, як наслідок, збільшення цільового трафіку, що надходить з пошукових систем [1]. Сьогодні просування сайту в пошукових системах є одним з найефективніших способів залучення цільової аудиторії. З огляду на це вона з кожним днем привертає увагу дослідників, метою яких є розроблення нових методів та засобів для її автоматизації [2].

Порівнюючи ефективність Seo-оптимізації та інших видів інтернет-реклами, не можна віддати перевагу тому або іншому напряму інтернет-маркетингу. Кожний з них має свої сильні і слабкі сторони. У ряді випадків слабкі сторони пошукового просування роблять його менш привабливим для замовників. До таких недоліків можна віднести: високу конкуренцію, що склалася за тематикою, відсутність гарантій результатів, періодичну зміну в алгоритмах роботи пошукових систем, що здатне вплинути на стабільність досягнутих результатів. Проте саме зазначений механізм просування сайту дозволяє за мінімальною ціною забезпечити якнайширше інформування аудиторії і тим самим привернути відвідувачів до власного ресурсу. Особливо цей підхід є актуальним для ресурсів з великою кількістю варіантів пошукових запитів та корисних сторінок [3].

З огляду на це актуальною є задача аналізу всіх етапів проведення пошукової оптимізації сайту з можливістю їх “оптимізації” відповідно до вимог користувача. При цьому критерієм оптимізації слугуватиме мінімізація часу проведення Seo-оптимізації.

Зв’язок висвітленої проблеми із важливими науковими та практичними завданнями

Проектування системи автоматизації процесу Seo-оптимізації є важливим завданням у межах розвитку як IT-галузі, так і промисловості загалом. Немає жодного підприємства чи суб’єкта господарювання, який так чи інакше не є представлений в глобальній мережі Інтернет, що, своєю чергою, сприяє розвитку відповідного математико-алгоритмічного апарату для реалізації даного процесу [5].

Внаслідок новизни поставленої задачі немає загального її вирішення для широкого кола вхідних даних [1]. З певними допущеннями можна застосовувати методи системного аналізу та технології штучного інтелекту [3]. Тому побудова нових методів на основі даного конгломерату сприятиме вирішенню завдання автоматизації процесу Seo-оптимізації.

Аналіз останніх досліджень та публікацій

Автоматизація процесу Seo-оптимізації є відносно новим науково-практичним завданням [1, 6] оскільки її виникнення відбулося на фоні інтенсивного розвитку мережі Інтернет та збільшення зацікавленості користувачів до її можливостей із просування товарів та послуг. Ця ситуація сприяла виникненню цілої галузі – Seo-оптимізації – та відповідних компаній, які надають відомості про аналізований інтернет-ресурс [4]. Особливістю цих організацій є те, що зазвичай дії, які вони здійснюють, полягають у “просуванні” конкретного ресурсу в межах однієї пошукової системи, що не вирішує проблеми оптимізації загалом [5].

З огляду на здійснений аналіз [2,3] актуальною є задача проектування системи автоматизації процесу Seo-оптимізації, що надасть необхідних засобів із покращення позицій інтернет-ресурсів.

Основні завдання дослідження та їх значення

Основними завданнями дослідження є: аналіз основних етапів здійснення процесу Seo-оптимізації, розгляд методологічної структури та формулювання рекомендацій щодо побудови структурної схеми системи Seo-оптимізації.

Відповідно до цих завдань дослідження для їх вирішення необхідно: сформулювати загальну схему процесу Seo-оптимізації; відобразити та проаналізувати основні її етапи; визначити методи оптимізації та окреслити головних акторів [2], які братимуть участь у процесі Seo-оптимізації. Вирішення зазначених завдань дасть змогу спроектувати прикладну систему, яка надасть засоби: накопичення даних щодо параметрів оптимізації та передбачить можливості з їх систематизації; аналізу, впровадження та підтримки нових сервісів пошукових систем; надання рекомендацій щодо оптимальних варіантів Seo-оптимізації.

Основні результати досліджень

Внаслідок проведеного дослідження вдалося сформулювати методологію із просування інтернет-ресурсів, що складається з таких етапів: складання семантичного ядра, підготовка текстів сайту (копірайтинг), оптимізація коду, розміщення посилань, моніторинг і аналіз результатів.

Складання семантичного ядра полягає у відборі пошукових фраз, за якими просуватиметься інтернет-ресурс та розподілятиметься по сторінках сайту. В загальному випадку це повний набір слів і словосполучень, що описують певний предмет та його характеристики. Правильне складання семантичного ядра має дуже важливе значення для пошукової оптимізації, тому що саме на його основі будується пошуковий механізм, без якого не можливе просування сайту. Семантичне ядро визначає стратегію створення контенту на сайті і може або звужуватися, або розширюватися, доповнюватися синонімами, новими поняттями. Залежно від типу ресурсу і цілей просування розмір семантичного ядра може складатися як з десятка, так і з тисяч ключових фраз. Існує багато різних програм і сервісів, що автоматизують цю процедуру, проте найкраще складати семантичне ядро вручну, використовуючи можливості пошукових систем, зокрема таких, як сервіс Wordstat в Яндексі або сервіс підбору ключових слів від компанії Google. Так можна повністю проконтролювати процес із усвідомленням виконуваних дій.

Під час цієї процедури виділяють три групи запитів: високочастотні, середньочастотні і низькочастотні. Належність запиту до певної групи визначається за відповідною кількістю запитів, зроблених користувачами протягом місяця. Крім частоти введення, необхідно враховувати градації запитів: висококонкурентні, середньоконкурентні і низькоконкурентні. Визначення конкурентності запиту є важкорезалізованим і довшим у часі, оскільки вимагає детального “ручного” аналізу відповідних ресурсів. Зокрема процес підбору ключових слів із використанням сервісу Wordstat (Wordstat.Yandex.ua) полягає у введенні міста (регіону), для якого здійснюється аналіз, та відповідного ключового слова. У результаті отримується статистика у різних варіаціях (рис. 1). Цифри поруч з кожним запитом у результатах підбору надають відомості щодо попередніх прогнозів кількості показів на місяць, які можна отримати, обравши запит як ключове слово. Потім з отриманого списку підібраних запитів потрібно обрати цільові запити (запити, що стосуються тематичного наповнення ресурсу) та після формування останніх завершити роботу над створенням семантичного ядра.

Яндекс
підбір слів

просування сайту ✕ Підібрати

● За словами ○ За регіонами ● Історія запитів

Всього показів за фразу «просування сайту»: 1 288 Список Карта

Все	Регіони	Міста	Показів на місяць	Регіональна популярність
			709	1 602%
			33	380%
			31	1 512%
			15	2 161%
			13	684%
			6	8%
			5	49%
			4	697%
			3	45%
			3	1%
			2	6%
			1	16%
			1	6%
			1	472%

Рис. 1. Підбір слова за допомогою сервісу Wordstat

Підготовка текстів сайту (копірайтинг) полягає в підготовці текстів, привабливих для відвідувачів сайту і одночасно оптимізованих під пошукові системи. Існує кілька видів копірайтингу: простий та складний, при цьому їх часто називають рерайтингом. Простий рерайтинг полягає в переробленні однієї статті в іншу, в якій описується все те саме, але іншими словами. На виході формується “унікальний контент”, що не порушує авторських прав джерела. При цьому складний рерайтинг здійснюється аналогічно, але на вході існує множина джерел. У результаті цього процесу створюється оптимізований текст щодо запитів пошукових систем.

Під оптимізацією коду сайту розуміють технічний процес щодо зменшення розміру коду, який передається користувачам і пошуковим системам під час завантаження. Оптимізація коду має значні переваги, зокрема: зменшення розміру сторінок сайту і прискорення їх завантаження; зниження навантаження на сервер хостингу, що позитивно позначається на його стабільності (особливо актуально для проектів з високим рівнем відвідуваності) [4].

Розміщення посилань належить до “зовнішніх чинників” оптимізації – без їх врахування неможливе успішне просування сайту в конкурентних тематиках. Причому ці роботи необхідно здійснювати постійно, створюючи зовнішнє “посилальне оточення” [2] та відображаючи необхідний текст навколо посилань, тобто текст, що відповідає тематиці товарів.

Моніторинг і аналіз результатів полягає в постійному контролі позицій інтернет-ресурсу в результатах пошуку за всіма відібраними запитамі і аналізі змін, що відбуваються. Ця процедура дає змогу оперативно реагувати на зміни в алгоритмах ранжування пошукових машин, оцінювати і своєчасно вносити необхідні корективи.

Схема процесу SEO-оптимізації

Що стосується розглянутого процесу Seo-оптимізації із залученням відповідних фахівців, то вона має вигляд, як на рис. 2. Тобто процес аналізу та просування веб-сайту складається з таких етапів:

1. *Аналіз інтернет-ресурсу та створення нового проекту.* При надходженні проекту в роботу за ним закріплюється відповідальний фахівець – проект-менеджер. На всіх етапах робіт із просування з клієнтом взаємодіє закріплений за проектом акаунт-менеджер. На початковому етапі на підставі побажань клієнта визначають мету та пріоритети проекту. Проект-менеджер аналізує сайт й ресурси конкурентів. Після обліку отриманих акаунт-менеджером під час діалогу із клієнтом даних проект-менеджер розподіляє цілі сторінками ресурсу.

2. Створення плану робіт. На підставі отриманих даних (у результаті аналізу сайту та ресурсів конкурентів з врахуванням рівня складності кожного із ключових запитів) проект-менеджер за допомогою співробітників відділу просування формує план робіт за проектом.

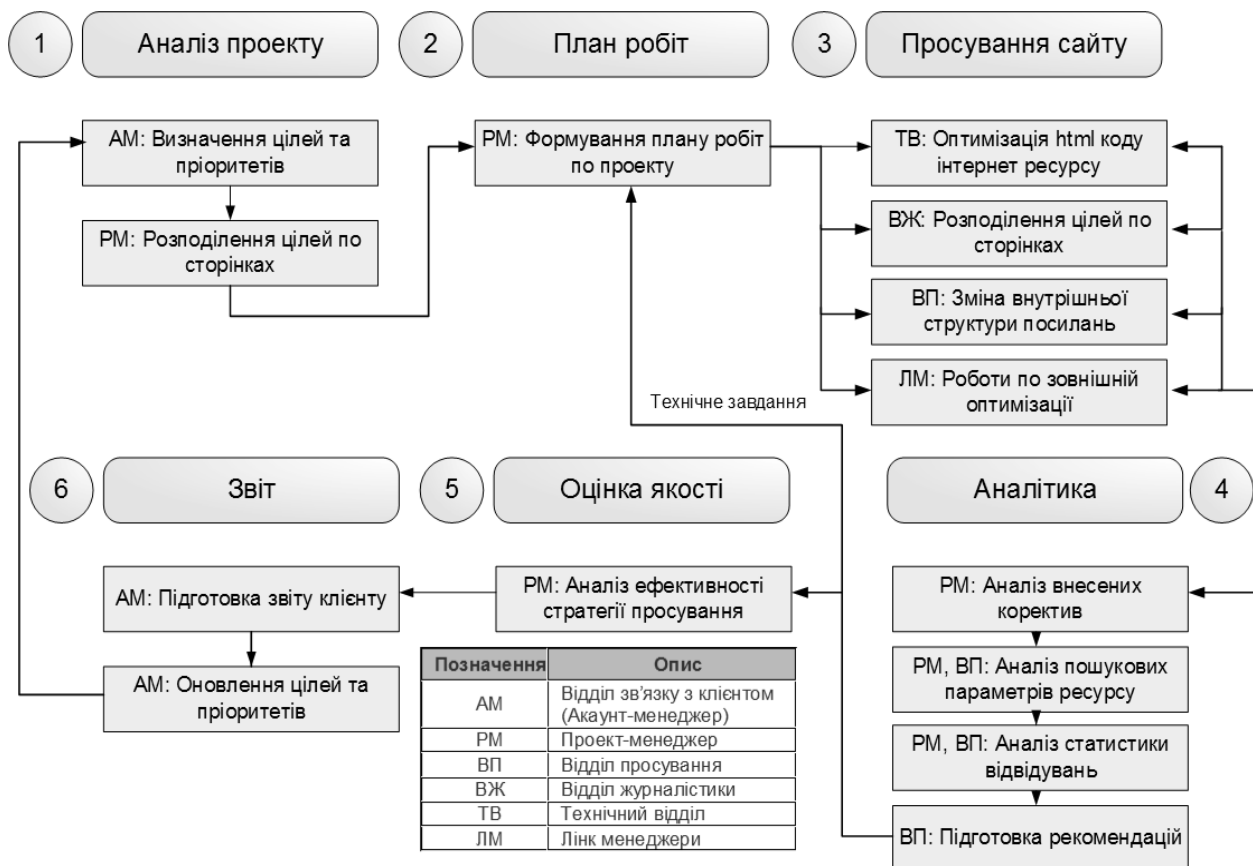


Рис. 2. Схема процесу SEO-оптимізації

3. Просування сайту складається з внутрішньої та зовнішньої оптимізації.

3.1. Внутрішня оптимізація інтернет-ресурсу

3.1.1. Програмісти технічного відділу за підготовленим проект-менеджером технічним завданням вносять необхідні для просування ресурсу виправлення в html-код. У випадку відсутності повного доступу до сайту акаунт-менеджер передає проект-менеджеру підготовлене технічне завдання на внесення необхідних корективів в html-код сайту технічним фахівцям з боку клієнта. Після внесення корективів проект-менеджер перевіряє коректність виконання технічного завдання і в разі потреби готує нове завдання із необхідними уточненнями й доповненнями.

3.1.2. Копірайтери відділу журналістики на підставі підготовленого проект-менеджером технічного завдання готують необхідні для просування сайту корективи тексту та статті. Акаунт-менеджер погоджує нові текстові матеріали із клієнтом і за необхідності копірайтери вносять повторні зміни в текстовий матеріал на підставі отриманої від клієнта інформації. Потім технічний відділ розміщує погоджені варіанти текстів.

3.1.3. Фахівці відділу просування готують технічне завдання на коригування зовнішньої "посилальної структури" для оптимізації передавання інформації між сторінками ресурсу. Технічний відділ вносить відповідні корективи в ресурс.

3.2. Зовнішня оптимізація сайту. На цьому етапі проводиться робота над зовнішніми факторами ранжування ресурсу. Згідно з поставленим завданням лінк-менеджери розміщують у ресурсі посилання із сайтів, близьких по тематиці (для збільшення ТІЦ), а також високо ранжованих сайтів (для збільшення загальної важливості інтернет-ресурсу) [2].

4. Аналітика. На цьому етапі аналізуються основні внутрішні й зовнішні параметри інтернет-ресурсу, який просувається, розглядається динаміка позицій сайту в різних пошукових системах за

кожним із ключових запитів та моделюється статистика відвідуваності сайту. На підставі отриманих даних відділ просування формує список рекомендацій з подальшої роботи.

5. Оцінка якості. На підставі даних, отриманих на попередньому етапі, проект-менеджер оцінює основні параметри просування сайту з врахуванням мети і пріоритетів, сформульованих на першому етапі. Потім проект-менеджер збирає дані, необхідні для формування звіту клієнтові.

6. Звіт клієнтові. На підставі наданих проект-менеджером даних акаунт-менеджер формує звіт для клієнта. Після аналізу звіту клієнтом у випадку корегування вхідних даних за цілями і пріоритетами акаунт-менеджер передає відповідну інформацію проект-менеджеру для зміни стратегії просування інтернет-ресурсу. У випадку одержання нових цілей або коригування пріоритетів роботи проводяться, починаючи з другого етапу.

Якщо список цілей і пріоритетів залишився незмінним, роботи тривають, починаючи з третього етапу. У зв'язку з тим, що в алгоритм ранжування сайтів роботами пошукових систем періодично вносяться корективи, необхідною умовою просування є корегування внутрішніх параметрів сторінок, які впливають на ранжування.

Етап зовнішньої оптимізації сайту також є циклічним, тобто роботи з нарощування й корегування зовнішнього “посилального оточення” сайту, що просувається, необхідно проводити безупинно для підтримки й поліпшення досягнутих позицій.

З огляду на зазначені етапи аналізу та просування інтернет-ресурсу було розроблено структурну схему модульної системи автоматизації процесу Seo-оптимізації (рис.3).

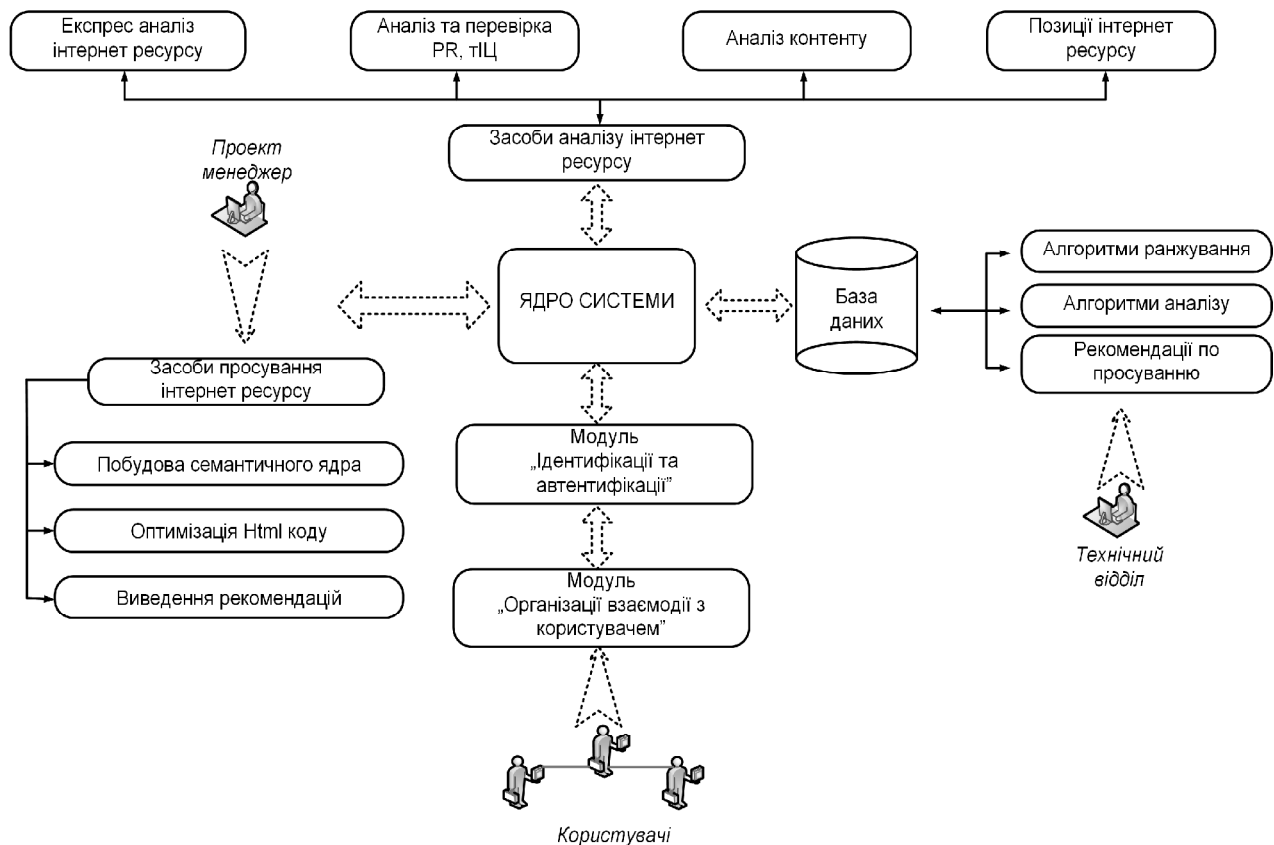


Рис. 3. Структурна схема системи автоматизації процесу Seo-оптимізації

Основними модулями системи є: засоби аналізу та просування інтернет-ресурсу, ядро, модуль організації взаємодії з користувачем та модуль ідентифікації та автентифікації.

Система надає два рівні доступу: на рівні менеджера (адміністратора) та на рівні користувача. Менеджер (адміністратор), володіючи повними правами, може: активувати та видаляти користувачів, змінювати інформацію бази даних тощо, а права користувача обмежуються аналізом

інтернет-ресурсу та виведенням рекомендацій щодо його просування. Важливою операцією, яку виконує модуль адміністрування, є опрацювання граничних ситуацій, що реалізується визначенням реакції кожного об'єкта й всієї системи на певні зовнішні дії: ініціалізацію, термінацію й обвал. Зокрема процес ініціалізації полягає в приведенні системи до фіксованого початкового стану: оголошення всіх початкових змінних та параметрів задачі та формування ієрархії об'єктів. Процес термінації полягає у вивільненні всіх апаратних та програмних ресурсів, зайнятих під задачі системи. Процес обвалу являє собою незаплановану термінацію системи [6].

Ядро системи – основний модуль системи, в якому опрацьовуються всі операції та роботи з її базою даних. Ядро системи координує роботу решти модулів з метою формування цілісної системи. Особливою функцією зазначеного модуля є організація процесу введення вхідної інформації, її опрацювання та збереження.

Модуль організації взаємодії з користувачем призначений для забезпечення особливостей роботи залежно від прав користувача. Реалізується через графічний інтерфейс користувача, що об'єднує всі елементи та компоненти програми, які впливають на взаємодію з користувачем.

Засоби аналізу інтернет-ресурсу містять команди, що реалізують: експрес-аналіз, аналіз та перевірку PR та ТІЦ, аналіз контенту та визначають позицію ресурсу в пошукових системах. Засоби просування реалізують: побудову семантичного ядра, оптимізацію HTML-коду та виведення рекомендацій щодо просування ресурсу.

Висновки

У статті описано особливості проектування системи автоматизації процесу Seo-оптимізації. Описані підходи та сформовані етапи із просування інтернет ресурсів дали змогу спроектувати структурну схему системи, що надає методологічне підґрунття для її проектування з використанням сучасних засобів розроблення. Зазначений підхід дає змогу: застосовувати нові адаптивні алгоритми просування ресурсів; спростити процес формування “семантичного ядра”; підвищити інформативність та конкурентоздатність послуг, які існують на ринку Seo-оптимізації. Подальші дослідження будуть скеровані на вивчення методів організації та проектування кожного з модулів системи автоматизації процесу Seo-оптимізації.

1. Ашманов І.С. *Просування сайту в пошукових системах* / І.С. Ашманов, А.А. Іванов. – К.: Собор, 2010. – 340 с. 2. Басюк Т.М. *Принципи побудови системи аналізу та просування інтернет ресурсів* / Т.М. Басюк // *Вісник Нац. ун-ту “Львівська політехніка”*. – 2012.– № 784. – С.43-48. 3. Райцин М. *Ефективне просування сайтів в Яндексі* / М. Райцин, О. Кураков. – К.: Діалектика, 2009. – 456 с. 4. Тєроу Ш. *Видимість в Інтернеті. Поисковая оптимизация сайтов* / Ш. Тєроу. – М.: Символ, 2009. – 288 с. 5. Vasyuk T.M. *Optimization of internet resource under plurality of user requests* / Т.М. Vasyuk // *Тези доповідей міжнародної науково-практичної конференції “Інформаційні управляючі системи та технології”*. – Одеса: 2013. – С. 305–307. 6. Grappone J. *Search Engine Optimization (SEO): An Hour a Day.* / J. Grappone – United States: Wiley Publishing, Inc., 2013. – 435 p.

ПІДХІД ДО ОЦІНЮВАННЯ РИЗИКІВ У ЗАДАЧАХ ПЛАНУВАННЯ

© Гожий О., Кобилінський І., Лугінець Д., 2014

Проаналізовано головні типи ризиків та особливості методів аналізу ризиків у задачах планування. Визначено підхід до аналізу та моделювання ризиків у задачах планування на основі VaR методу. Досліджено ефективність параметричних та непараметричних методів оцінювання VaR.

Ключові слова: планування, ризик, методи оцінювання, ефективність.

The paper analyzes the main types of risks and specific methods of risk analysis in planning problems. The approach to analysis and risk modeling in problems planning based on VaR method. The efficiency of parametric and nonparametric methods for evaluating VaR.

Key words: planning, risk, assessment methods, efficiency.

Вступ

Ризики притаманні будь-якій діяльності, а з розвитком іноваційних технологій та підприємницької сфери кількість наявних та потенційних ризиків значно зростає. Сьогодні процес управління ризиками розглядається як ключовий напрям прикладного менеджменту, значна увага приділяється вивченню ризикових сфер і основних видів ризиків, пошуку ефективних методів їх оцінювання, контролю та моніторингу, а також створенню відповідних систем ризик-менеджменту. Аналіз ризиків у різних прикладних задачах – дуже широка та динамічна галузь досліджень. Ефективність вирішення будь-якої проблеми переважно залежить від правильності й обґрунтованості прийняття рішень на всіх етапах розв'язання задач і незалежно від складності розв'язуваних задач, що, своєю чергою, неможливо без урахування ризиків. Для управління будь-яким процесом чи для розв'язування задач планування потрібно вміти аналізувати ризик, оцінювати його ступінь, передбачати наслідки від прийнятого рішення і не виходити за допустимі межі ризику. Тобто для ефективного вирішення задач динамічного планування необхідно виявляти ризик, передбачати його, маючи на меті зниження його до якомога нижчого рівня.

Постановка задачі

Метою роботи є: 1) розглянути особливості методів аналізу ризиків у задачах динамічного планування; 2) розробити підхід до моделювання ризиків у задачах планування на основі VaR методу.

Огляд основних типів ризиків та сучасних методів і моделей аналізу ризиків

Основною причиною виникнення ризиків є невизначеність середовища діяльності, що, своєю чергою, зумовлено такими чинниками, як брак повної та достовірної інформації про зовнішнє середовище; обмеженість можливостей щодо сприйняття та опрацювання інформації про процес або систему; випадковість появи несприятливих подій у процесі діяльності або планування; свідомо протидія учасників процесу, зокрема конкурентів; виникнення конфліктів; порушення договірних зобов'язань; політичні рішення, які істотно впливають на економіку тощо.

Потреба оцінити ризик під час вирішення різноманітних завдань виникає постійно. Розрізняють два типи оцінювання ризиків: “суб’єктивне” – це оцінювання ризику типу “великий”, “середній”, “малий”; таке оцінювання ризику впливає на суб’єктивні рішення, які приймаються [1]. Є “технічне” оцінювання ризиків, яке використовують в інформаційних системах для детальнішого оцінювання ризиків, наприклад: економічних, страхових, екологічних тощо.

З погляду джерел виникнення ризику поділяють на *зовнішні* (ендогенні) і *внутрішні* (екзогенні) [1, 3].

До зовнішніх належать ризики, які виникають у зовнішньому щодо об'єкта середовищі і безпосередньо не залежать від його діяльності. Це політичні, правові, соціальні та загальноекономічні ризики, що виникають у разі загострення економічної кризи в країні, політичної нестабільності, війни, заборони на платежі за кордон, консолідації боргів, введення ембарго, відміни імпорتنих ліцензій, стихійного лиха (пожежі, повені, землетруси), приватизації, націоналізації, неадекватного правового регулювання тощо.

Основними чинниками, які впливають на рівень зовнішніх ризиків, є політичні, технічні та економічні.

Усі інші чинники – демографічні, соціальні, географічні – розглядають крізь призму політичних та економічних чинників.

Серед зовнішніх ризиків виділяють п'ять основних груп:

- *ризик форс-мажорних обставин* – виникнення непередбачених обставин, які негативно впливають на діяльність будь-якого підприємства чи установи;
- *ризик країни* – можливість настання несприятливих для діяльності умов у політичній, правовій чи економічній сферах країни, де проводить діяльність підприємство чи установа;
- *зовнішньополітичний ризик* – імовірність зміни міжнародних відносин, а також політичної ситуації в одній із країн, які впливають на діяльність підприємства чи установи (війни, міжнародні скандали, закриття кордонів тощо);
- *правовий ризик* – імовірність настання несприятливої ситуації, пов'язаної зі змінами законодавства різних країн;
- *макроекономічний ризик* – імовірність несприятливих змін кон'юнктури на окремих ринках або економічній ситуації загалом (економічна криза).

Реалізація зовнішніх ризиків може поставити під загрозу безперервність діяльності підприємства чи установи. Вплив зовнішніх ризиків на результативність роботи установи виключно високий, управляти цими ризиками найскладніше, а найчастіше – неможливе. Тому підприємство чи установа має враховувати можливий вплив зовнішніх ризиків та оцінювати ймовірність виникнення екстремальних обставин за допомогою стрес-сценаріїв. Для оцінювання зовнішніх ризиків застосовуються переважно експертні і логічні методи аналізу. Також розробляють відповідні нагальні заходи у формі плану дій на випадок настання кризових обставин. Цей план слід регулярно оновлювати та тестувати. Такий план є невід'ємною складовою механізмів контролю за рівнем ризиковості діяльності.

До внутрішніх належать ризики, які виникають безпосередньо внаслідок діяльності конкретного підприємства чи установи. Чим ширше коло клієнтів, партнерів, зв'язків, банківських операцій, послуг, тим більше внутрішніх ризиків супроводжує його роботу. Порівняно із зовнішніми внутрішні ризики краще піддаються ідентифікації та квантифікації.

Згідно з прийнятою класифікацією, серед внутрішніх ризиків виокремлюють декілька категорій ризиків, які поділяють на дві групи.

До першої групи віднесено ризики, які піддаються кількісному оцінюванню, тобто квантифіковані. *Квантифікованими* називаються ризики, щодо яких існує безпосередня залежність між рівнем ризику і доходами. Мета процесу управління цими ризиками полягає в їх оптимізації. До другої групи належать ризики, які не піддаються кількісному оцінюванню, тобто *неквантифіковані* [1]. Ризики, щодо яких немає безпосередньої залежності між ризиком і доходами банку, є неквантифікованими. Мета управління полягає в їх мінімізації. Це такі ризики, як юридичний, стратегічний, репутацій.

Процес ідентифікації та менеджменту ризиків у задачах, динамічного планування спрямований на виявлення наявних у процесі розробки плану ризиків, їх оцінювання, планування та реалізацію заходів стосовно зниження впливу та ліквідацію факторів ризиків. Загальну схему процесу ідентифікації та менеджменту ризиків зображено на рис. 1.

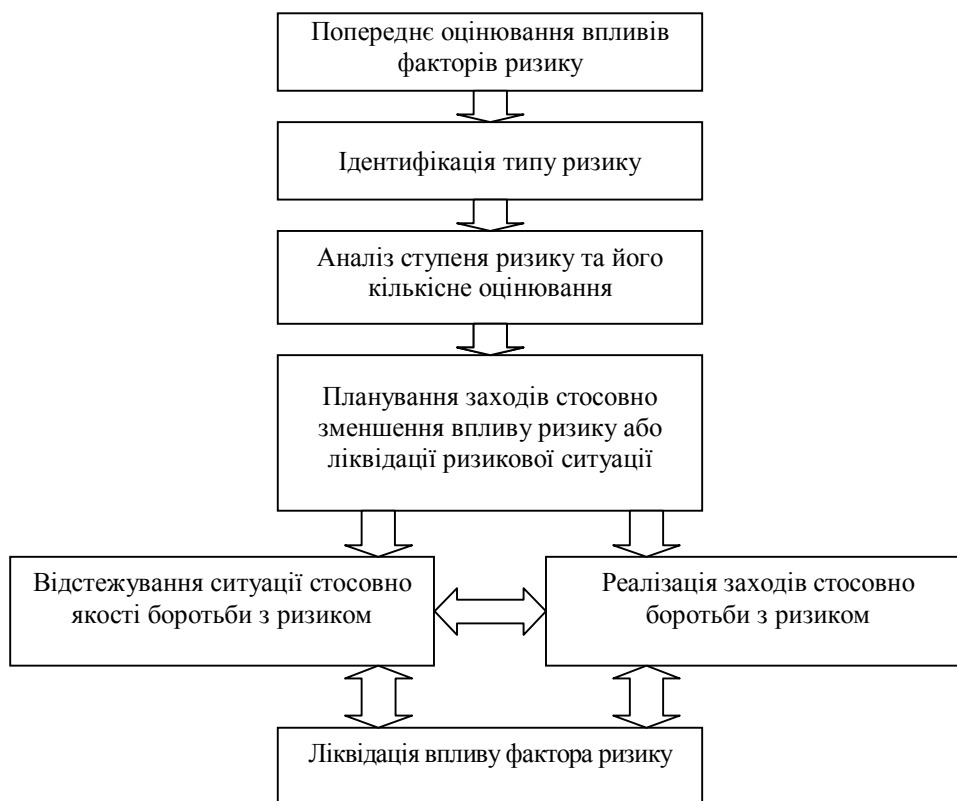


Рис. 1. Процес ідентифікації та менеджменту ризиків

Для оцінювання зовнішніх ризиків застосовуються переважно логічні та експертні методи аналізу. Також розробляють відповідні нагальні заходи на випадок настання кризових обставин. Такий план слід регулярно оновлювати та тестувати. До внутрішніх належать ризики, які виникають безпосередньо у зв'язку з розробленням конкретних етапів плану. Порівняно із зовнішніми внутрішні ризики краще піддаються ідентифікації та квантифікації.

Міри ризику

Для оцінювання величини ризиків у задачах динамічного планування можливо використовувати міри ризиків. Два важливі аспекти характеризують ризик: **волатильність** (мінливість) індикаторів, імовірність або частота подій, і, по-друге, **чутливість** критеріїв діяльності до їх наслідків. Відповідно можна запропонувати дві основні категорії виміру ризиків: показники чутливості і імовірнісні (статистичні) величини [6]. Розподіл стає ще більш умовним, якщо взяти до уваги суб'єктивність, що є принциповою властивістю оцінювання ризику.

Волатильність. На практиці розподіл показників ефективності зазвичай оцінюється за ретроспективою, допускаючи, що спостереження ідентичні і незалежно розподілені. Якщо N – кількість спостережень, то очікувану дохідність m можна оцінити простою середньою \bar{x} , а ризик, варіацію – оцінкою дисперсії. Квадратний корінь із оцінки дисперсії дохідності – стандартне відхилення – називають волатильністю (мінливістю):

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (1)$$

Вона вимірює величину розсіювання значень показників ефективності навколо очікуваного рівня [2].

Необхідно зазначити, що **невизначеність ризику** характеризують не певні показники розподілу, а сам розподіл. Існують розподіли, наприклад, розподіл Коші, для якого не існує дисперсії, і апроксимація розподілом Коші випадкових процесів, які характеризуються кінцевим

математичним сподіванням і кінцевою дисперсією, є неправомірною [6]. Тому вибрати для оцінювання волатильності ту чи іншу характеристику розподілу необхідно, враховуючи особливості розподілу даних і задачі, що розглядаються. Тобто *волатильність* — це характеристика, що визначає мінливість результатів залежно від використовуваного інструмента. Одним з прикладів є стандартне відхилення.

Інтервали часу, за які розраховується волатильність та інші параметри, можуть бути різноманітні: години, дні, тижні, місяці, квартали, роки. Виникає проблема агрегування – вираження волатильності і очікуваної ефективності для різних періодів.

Розглянуті міри ризику характеризуються такими недоліками:

1. Більша частина з них не можуть бути агреговані, тобто зведені в один показник такого самого типу однаково до всіх факторів ризику. Фактори ризику не можуть бути агреговані для різних задач.

2. Традиційні міри ризику недостатньо контролюють ризик. Ліміти позицій, що визначаються факторами ризику або показниками чутливості, часто неефективні.

Все це пояснює ту велику популярність, якою в сучасному ризик-менеджменті користується підхід до оцінювання ризиків на основі показника “вартості під ризиком” (value at risk – VAR), який ефективно справляється з вищеперахованими проблемами.

Показник VaR (Value-at-Risk). VAR – це виражена у певних грошових одиницях (базовій валюті) оцінка величини, яку не перевищать протягом даного періоду втрати із заданою імовірністю [2].

Нехай зафіксовано деякий портфель відкритих позицій. **VaR** портфеля для даного довірчого рівня $(1-a)$ і даного періоду підтримки позиції t визначається як таке значення, яке забезпечує покриття можливих втрат x утримувача портфеля за час t з імовірністю $(1-a)$:

$$P(\text{Var} \geq x) = 1 - a. \quad (2)$$

Як впливає із визначення, величина VaR для портфеля заданої структури – це найбільший очікуваний збиток, що спричинений коливанням цін на фінансових ринках, який розраховується:

- на визначений період часу у майбутньому (часовий горизонт);
- із заданою імовірністю його перевищення (рівень довіри);
- при даному припущенні про характер поведінки ринку (метод розрахунку).

Ключовими для VaR є такі параметри:

• **Очікувана кількість ризику**, яку можна розраховувати в абсолютному вимірі або у процентному відношенні до значення показника на певну дату.

• **Часовим горизонтом**, який характеризується очікуваною кількістю ризику (тобто термін, за який можна реалізувати на ринку цей інструмент (закрити позиції) без істотних втрат). На практиці залежно від специфіки використання VaR такими горизонтами найчастіше можуть бути день, тиждень, декада, місяць.

• **Глибина періоду розрахунку VaR** – це об’єм ретроспективних або штучно змодельованих даних, на основі яких визначається оцінка. Наприклад, фраза “глибина розрахунку тижневого VaR становить 2 роки” означає, що для розрахунку втрат на тиждень бралися дані за 2 роки [5].

• **Рівень довіри (імовірність)**, з якою максимальні збитки не перевищують розрахованої очікуваної кількості ризику, визначається залежно від переваг за ризиками, що виражені в регламентуючих документах наглядових органів або в корпоративній практиці. Наприклад, Базельський комітет з банківського нагляду рекомендує рівень довіри 99 %, на який орієнтуються наглядові органи, на практиці найпопулярніший рівень довіри – 95 % [6]. Тобто формулу (2) трактують так: очікувана кількість ризику VaR перевищить реальну кількість ризику x за часовий горизонт t з імовірністю a ($a=0,01; 0,05$ і т.п.).

Розглянемо підходи до оцінювання VaR. Методи оцінювання наведено на рис. 2.

Параметричний метод розрахунку VaR передбачає аналітичне обчислення необхідної оцінки ризику за статистичною моделлю фінансового результату портфеля.

В основу практично будь-якого параметричного методу покладено дві основні складові:

- модель залежності вартості фінансового результату портфеля від змін факторів ризику;
- модель волатильностей і кореляцій факторів ризику.

Локальне оцінювання означає лінійну, або складнішу апроксимацію функції вартості фінансового інструмента [6, 7].

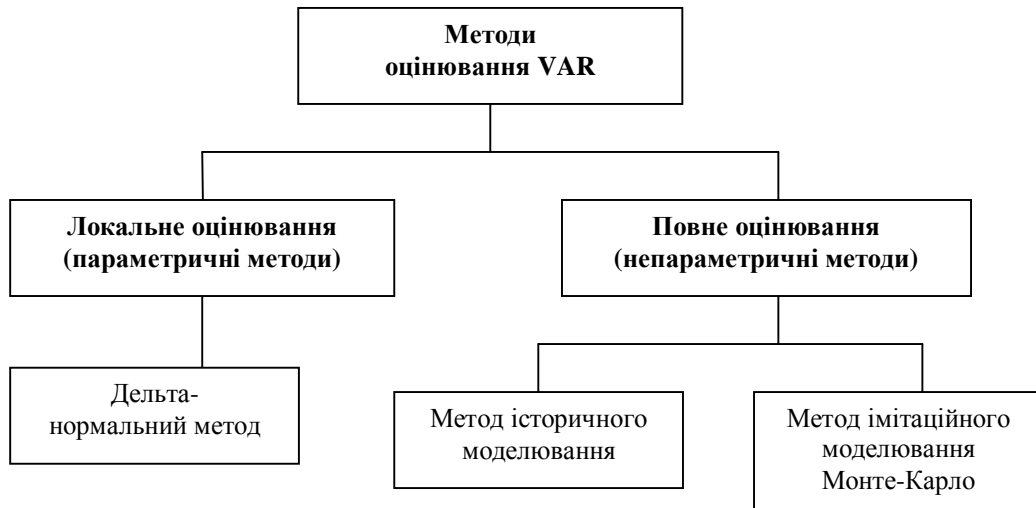


Рис. 3. Методи оцінювання VaR

На практиці зазвичай використовують два параметричні методи розрахунку VaR:

- Дельта-нормальний VaR;
- Дельта-гамма VaR.

На практиці найбільшої популярності здобув дельта-нормальний VaR

Дельта-нормальний метод розрахунку VaR дає змогу отримати оцінку VaR в замкненому вигляді. В його основу покладено припущення про нормальний закон розподілу логарифмічних дохідностей факторів ризику:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \sim N(m, s^2). \quad (3)$$

Припущення про нормальний закон розподілу факторів ризику значно полегшує знаходження величини VaR, тому що в цьому випадку розподіл дохідностей інструментів, які є лінійними комбінаціями факторів ризику, також буде нормальним. Ця фундаментальна властивість зберігатиметься для будь-якого портфеля, який складається із інструментів з лінійними ціновими характеристиками, як, наприклад, акцій або валют.

У випадку нормально розподіленої випадкової величини довірчий інтервал $(1-\alpha)$ завжди характеризується єдиним параметром – квантиллю $k_{1-\alpha}$, яка показує положення шуканого значення випадкової величини (симетрично в обох хвостах розподілу) відносно середнього $(M[r_t])$, що виражене в кількості стандартних відхилень доходності портфеля (s_t) . Так, для значень довірчого інтервалу, які найчастіше використовуються, 95 і 99 % відповідні квантилі дорівнюватимуть 1,65 і 2,33 стандартних відхилень доходності портфеля.

Переваги та недоліки дельта-нормального методу

Перевагами дельта-нормального методу є:

1. Порівняно проста реалізація
2. Порівняно невеликі витрати на збирання поточних даних
3. Припустима точність оцінки VaR у більшості випадків практичного застосування при відповідності вхідних даних.

Недоліками дельта-нормального методу є:

1. Низька точність оцінювання ризику нелінійних інструментів через недоліки виміру чутливості нелінійних інструментів до факторів ризику.

2. Для розподілів доходностей більшості фінансових активів характерні відхилення на краях розподілу щільності розподілу ймовірностей від нормального розподілу. Внаслідок цього оцінки VaR, що розраховані за нормальним розподілом, виявляються завищеними або заниженими.

3. Ігнорування ризику екстремальних ринкових подій, які можуть привести до аномальних збитків і не є достатньо частими для того, щоб бути представленими в останніх історичних даних (на основі яких оцінюється кореляція і волатильність доходностей).

Для врахування ризику екстремальних подій використовується стрес-тестування, яке є різновидом сценарного аналізу. Стрес-тестування – метод кількісного оцінювання ризику, який полягає у визначенні величини неузгодженої позиції, яка визначається при шоківій зміні зовнішнього фактора – валютного курсу, процентної ставки тощо [2]. Поєднання цих величин дає уявлення про те, яку суму збитків чи доходів можна отримати у разі, якщо події розвиватимуться за закладеними припущеннями.

Непараметричні методи (методи повного оцінювання)

Під непараметричними методами розумітимемо ті методи, в яких оцінювані функції (наприклад, функція розподілу) не визначаються кінцевим числом параметрів. Повне оцінювання означає, що проводиться повний перерахунок вартості фінансового інструмента без апроксимуючих припущень [5].

Найпоширенішими непараметричними методами є:

- Метод історичного моделювання;
- Метод імітаційного моделювання Монте-Карло.

Суть обох методів полягає в побудові емпіричної функції розподілу майбутніх змін цін і потенційних прибутків та втрат, що впливають .

Метод історичного моделювання. Історичний варіант VaR-методу розрахунку показника ризикової вартості (VaR) полягає у застосуванні для обчислень реальних історичних значень часового ряду випадкової величини, що аналізується. Для розрахунку VaR необхідно побудувати розподіл змодельованих змін вартостей активу або портфеля активів за вибраний попередній історичний період, які залежать від змін одного або декількох факторів ризику за цей самий період. Процедура реалізації методу є такою:

1) визначається початковий ряд показників – базових значень (наприклад, цін), що розглядається для всіх зафіксованих в історичному періоді станів ринку;

2) визначається часовий інтервал, на якому розраховуватиметься ризикова вартість (VaR);

2) визначається ймовірність (рівень довіри), з якою розраховуватиметься ризикова вартість;

3) з використанням базових значень, відповідної випадкової величини послідовно розраховуються зміни базових значень і відповідні їм зміни вартості одиничного активу або портфеля активів;

4) зміни вартостей, розраховані на попередньому етапі, впорядковуються за зростанням із формуванням часового ряду змін;

5) на часовому ряді змін відповідно до вибраної ймовірності, починаючи з найбільших від'ємних значень, виокремлюється стільки значень, щоб відношення їх кількості до загальної кількості значень у часовому ряді змін становило не більше $1-\alpha$ % для ймовірності α (наприклад, не більше 1 % для ймовірності 99 %);

6) значення з множини значень, що залишилися після виокремлення даних на попередньому етапі, з найменшим номером індексу і буде шуканим значенням ризикової вартості VaR [5].

Переваги методу історичного моделювання:

- відносна простота реалізації методу;
- відсутність припущення про нормальний закон розподілу доходностей факторів ризику або якої-небудь іншої стохастичної моделі динаміки цін на ринку, крім тієї, яка реально спостерігалася в минулому (що дозволяє враховувати ефект “товстих хвостів” такого розподілу;
- непогана точність оцінювання ризику нелінійних фінансових інструментів;

- відсутність ризику використання помилкової моделі для оцінювання вартості інструмента;
- інтуїтивна простота та наочність.

Недоліки методу історичного моделювання:

- некоректність результатів у випадку, якщо вибірка, отримана в базовому періоді, не є репрезентативною, зокрема і за кількістю спостережень;
- використання лише однієї траєкторії цін;
- ігнорування різниці між старими і останніми спостереженнями, тоді як видалення із вибірки найстаріших значень може значно покращити точність моделі;
- великий об'єм обчислень для великих диверсифікованих портфелів.

Метод імітаційного моделювання Монте-Карло. Метод Монте-Карло, або метод стохастичного моделювання оснований на моделюванні випадкових процесів із заданими характеристиками. На відміну від методу історичного моделювання, у методі Монте-Карло *змінюються псевдовипадковим чином* відповідно до заданих параметрів розподілу, наприклад, математичним сподіванням m і волатильністю s . Імітований розподіл може бути будь-яким, а кількість сценаріїв – доволі великою (до декількох десятків тисяч). В іншому метод аналогічний до методу історичного моделювання.

У таблиці 1 наведено головні особливості методів оцінювання .

Метод Критерії	Дельта - нормальний	Історичного моделювання	Метод імітаційного моделювання Монте-Карло
Оцінювання	Локальне	Повне	Повне
Врахування історичного розподілу	Як оцінка нормального розподілу	Аналогічно, тому, яке було у минулому	Повністю
Врахування “допустимої” волатильності	Можливе	Ні	Так
Припущення про нормальний розподіл дохідностей	Так	Ні	Ні
Оцінка екстремальних подій	Погана	Погана	Можлива
Модельний ризик	Може бути значним	Допустимий	Високий
Об'єм ретроспектив	Середній	Дуже великий	Малий
Обчислювальна складність	Невисока	Висока	Дуже висока
Наочність	Середня	Висока	Низька
Обчислювальні потужностей	Низькі	Середні	Високі

Висновки

Міра ризику VaR має низку переваг і недоліків. Серед її переваг слід зазначити простоту представлення інформації про ризик у вигляді лише одного значення вартості збитків. Серед її недоліків слід згадати відсутність інформації про випадки, вірогідність яких мала, відсутності інформації про вигляд розподілу збитків і можлива неоднозначність оцінювання збитків залежно від декомпозиції ресурсів за факторами ризику. Для оцінювання VaR використовують методи локального та повного оцінювання. Головним недоліком методів локального оцінювання є припущення про нормальний розподіл ретроспективних даних, за якими і оцінюють волатильність факторів ризику. Недоліками методів повного оцінювання є високі вимоги до облікових потужностей для роботи в режимі он-лайн та відсутність інформації про вигляд розподілу збитків,

коли при різних розподілах за заданим рівнем довіри може спостерігатися однакове значення VaR. Адекватність моделі для оцінювання ризику при плануванні потрібно визначати періодично, обчислюючи кількість помилок прогнозу можливих втрат за річний період.

1. Лобанов А.А., Чугунов А.В. *Энциклопедия финансового риск-менеджмента*. – М.: Альпина Паблишер, 2003. – 786 с. 2. Jorion Ph. *Financial risk-management: Second edition*. – Hoboken, New Jersey: John Wiley & Sons, 2003.–708 p. 3. Базельский комитет по банковскому надзору. *Международная конвергенция измерения капитала и стандартов капитала: новые подходы (рабочий вариант перевода “Банка России”)*. – Базель, 2004. – <http://www.cbr.ru> 4. *Управління банківськими ризиками: навч. посібник / Л.О. Примостка, П.М. Чуб, Т.Г. Карчева та ін; За ред. О.Л. Примостки*. – К.: КНЕУ, 2007. – 600 с. 5. Яблоков А.И. *Методика оцінки та управління валютним ризиком VaR // Економіко-математичне моделювання соціально-економічних систем*. – 2007. – № 13. – С. 121–128. 6. Барковський В.В., Барковська Н.В., Лопатін О.К. *Теорія ймовірностей та математична статистика*. – К.: ЦУЛ, 2002. – 448 с. 7. Шора О.Є. *Застосування VAR – методології в практичній діяльності комерційних банків // Облік і фінанси АПК*. – 2005. – № 12. – С. 142–145.

УДК 004.738.5

А. Берко¹, В. Висоцька², Л. Чирун³

Національний університет “Львівська політехніка”,

¹кафедра загальної екології та екоінформаційних систем,

²кафедра інформаційних систем та мереж,

³кафедра програмного забезпечення

ТЕОРЕТИКО-МНОЖИННА МОДЕЛЬ СИСТЕМ ЕЛЕКТРОННОЇ КОНТЕНТ-КОМЕРЦІЇ

© Берко А., Висоцька В., Чирун Л., 2014

Проаналізовано основні проблеми електронної контент-комерції та функціональних сервісів опрацювання контенту. Запропонована модель дає можливість створити методи та засоби опрацювання інформаційних ресурсів в системах електронної контент-комерції та реалізувати підсистеми формування, управління та супроводу контенту.

Ключові слова: інформаційний ресурс, комерційний контент, контент-аналіз, контент-моніторинг, контентний пошук, система електронної контент-комерції.

General problems of electronic content commerce and functional services of content processing are analyzed in the article. Proposed model gives an opportunity to create a method and instrument of information resources processing in electronic commerce systems and to implement the subsystem of content formation, management and support.

Key words: information resources, commercial content, content analysis, content monitoring, content search, electronic content commerce systems.

Вступ. Загальна постановка проблеми

Активний розвиток Інтернету сприяє зростанню потреб у даних виробничого/стратегічного характеру і реалізації нових форм інформаційного обслуговування [1, 2]. Документована інформація, підготовлена відповідно до потреб користувачів і призначена для їх задоволення, є інформаційним продуктом або комерційним контентом, що є основним об'єктом процесів електронної контент-комерції [1, 2, 5]. Характерною рисою таких систем є можливість автоматично опрацьовувати інформаційні

ресурси для збільшення обсягів продажу контенту постійному користувачу, активно залучати потенційних користувачів та розширяти межі цільової аудиторії [4, 5, 7, 8]. Принципи і технології електронної контент-комерції активно застосовують, створюючи системи on-line/off-line продажу та аналізу/обміну/збереження контенту, Інтернет-магазину, cloud storage/computing [2]. Проблематика проектування, створення, впровадження та супроводу систем електронної контент-комерції є актуальною з огляду на такі фактори, як активний розвиток досліджень у галузі е-бізнесу та відсутність теоретичного обґрунтування стандартизованих методів і потреба в уніфікації програмних засобів опрацювання інформаційних ресурсів у таких системах [2]. З'являються нові підходи та способи вирішення цієї проблеми, але існує невідповідність між відомими методами і засобами опрацювання інформаційних ресурсів та принципами побудови систем електронної контент-комерції. Відсутні загальні підходи створення систем електронної контент-комерції та уніфіковані методи опрацювання інформаційних ресурсів у цих системах.

Аналіз останніх досліджень та публікацій

Практичний чинник опрацювання інформаційних ресурсів у системах електронної контент-комерції пов'язаний з вирішенням проблем зростання обсягів контенту в Інтернеті, у швидких темпах поширення доступності до Інтернету, активному розвитку електронного бізнесу, розширенні набору інформаційних товарів та послуг, зростанні попиту на інформаційні товари та послуги, створенні технологій і засобів та розширенні областей застосування методів опрацювання інформаційних ресурсів. У цьому напрямі активно працюють провідні світові виробники засобів опрацювання інформаційних ресурсів, зокрема розроблено такі системи: для on-line продажу контенту інтернет-газети (New York Times, Washington Post), інтернет-журнали (Wall Street Journal, Chip), дистанційне навчання (Ashford University), інтернет-видання у вигляді словників/довідників (Oxford English Dictionary Online), інтернет-видавництва (Online Publishers Association), портали розважального (tochka.net), інформативного (Weather Channel, Karpaty.ua) та дитячого (teremoc.ru, vshkolu.com) комерційного контенту; для off-line продажу контенту copywriting services (Textbroker, Copywriting 911), Marketing Services Shop (The Copy Box, Local Internet Marketing), RSS Subscription Extension (Apple, Google, Intel, Microsoft); Інтернет-магазин для продажу eBooks, Software, video, music, movies, picture, digital art, manuals, articles, certificates, forms, files тощо (Google, Amazon, Apple, Android, Opera, Yahoo!); для збереження контенту cloud storage (та cloud computing (EMC, Google, Amazon, Apple, Microsoft, Mac, Linux, iPhone, Android, Palm).

Теоретичний чинник опрацювання інформаційних ресурсів у системах електронної контент-комерції пов'язаний із розробленням методів та засобів формування, управління та супроводу контенту. В наукових роботах Д. Ланде, С. Брайчевського, А. Григор'єва та В. Фурашева досліджено та розвинуто математичні моделі електронних інформаційних потоків [3, 4, 6, 8]. Г. Зіпф (G. Zipf) запропонував емпіричну закономірність розподілу частоти слів природної мови, а Дж. Селтон (G. Salton) та Р. Папка (R. Papka) – виявлення нових подій в потоках контенту. В роботах С. Макківер (S. McKeever), Б. Бойко (B. Boiko), Дж. Макговерн (G. McGovern), Дж. Хаскос (J. Haskos), Е. Роклі (A. Rockley), Р. Накано (R. Nakano), Б. Дойл (B. Doyle), Р. Вудс (R. Woods), Халверсон (Halverson) описані моделі життєвого циклу контенту. Методологію контент-аналізу започаткували А. Тенні (A. Tenni), Б. Метьюз (B. Matthews), Д. Спііда (D. Spiida), Ж. Кайзер (J. Kaiser), Гласер (Glaser) та активно розвинули Г. Лассуел (H. Lasswell), О. Холсті (O. Holsti). Ф. Джобіш (F. Joubish) запропонував методологію дослідження текстів для визначення авторства, автентичності або сенсу. К. Нойендорф (K. Neuendorf) та К. Кріпендорф (K. Krippendorff) розробили методи кількісного та якісного аналізу тексту. Корпорації EMC, IBM, Microsoft Alfresco, Open Text, Oracle і SAP розробили специфікації Content Management Interoperability Services на інтерфейс Web-сервісів, покликаних забезпечити взаємодію між системами управління контентом електронного бізнесу. Інструментарій Interoperable Content Application взаємодіє з контентом із різних репозиторіїв за допомогою сервісного інтерфейсу і спеціальної надбудови Content Management Interoperability Services Implementation.

Виділення проблем

Кількість потоків Web-контенту в системах електронного бізнесу більша, ніж шляхів переміщення товарів на промислових підприємствах. Значна частина контентних потоків складається з легко формалізованих і автоматизованих процедур. Основна проблема – відсутність загального підходу до процесу моделювання, проектування та розроблення систем електронної контент-комерції(СЕКК).

Модель системи управління контентом – це множини величин

$$S = \langle X, Q, C, V, H, Function, T, Z, Y \rangle, \quad (1)$$

де $X = \{x_1, x_2, \mathbf{K}, x_{n_x}\}$ – множина контенту з різних джерел даних; $Q = \{q_1, q_2, \mathbf{K}, q_{n_q}\}$ – множина інформаційних запитів користувачів системи; $C = \{c_1, c_2, \mathbf{K}, c_{n_c}\}$ – множина комерційного контенту; $V = \{v_1, v_2, \mathbf{K}, v_{n_v}\}$ – множина умов супроводу контенту та зовнішніх впливів середовища на систему, $H = \{h_1, h_2, \mathbf{K}, h_{n_h}\}$ – множина умов формування та управління контентом; $Z = \{z_1, z_2, \mathbf{K}, z_{n_z}\}$ – множина компонентів інформаційного ресурсу, $T = \{t_1, t_2, \mathbf{K}, t_{n_t}\}$ – час транзакцій управління контентом; $Y = \{y_1, y_2, \mathbf{K}, y_{n_y}\}$ – множина вихідних характеристик роботи системи, *Function* – оператор формування вихідних характеристик системи [2].

Процес роботи системи управління контентом S описують оператором

$$y_j(t_p + \Delta t) = Function(x_i, q_d, c_r, v_l, h_k, t_p, z_w), \quad (2)$$

де x_i – це запити відвідувачів/користувачів до системи управління контентом. Згідно із Google Analytics значення y_j подають як множину $y_j = \{a_1, a_2, \dots, a_g\}$, де a_1 – це кількість відвідувань за період часу Δt ; a_2 – середній час перебування на інформаційному ресурсі (хв:с); a_3 – показник відмов (%); a_4 – досягнута мета; a_5 – динаміка (%); a_6 – кількість всього перегляду сторінок; a_7 – кількість перегляду сторінок за одне відвідування; a_8 – нові відвідування (%); a_9 – абсолютно унікальні відвідувачі; a_{10} – джерело трафіка у % (пошукові системи, прямий трафік або інші інформаційні ресурси) [5].

Впливи величин c_r , v_l , h_k , на z_w та y_j як результат роботи CMS є невідомими та не дослідженими. Модель не розкриває зв'язків між вхідною інформацією, контентом, вихідною інформацією та процеси опрацювання Web контенту в системі. Вивчення динаміки потоку комерційного Web контенту та побудова моделей опрацювання інформаційних ресурсів у СЕКК є важлива та актуальними. У статті необхідно розв'язати задачу опрацювання інформаційних ресурсів у системах електронної контент-комерції у вигляді теоретично обґрунтованої концепції, уніфікованих методів та програмних засобів шляхом автоматизації процесів формування, управління та супроводу контенту, що ґрунтуються на принципах побудови та функціональних можливостях таких систем.

Формулювання мети

Метою роботи є розширення функціональних можливостей та підвищення ефективності роботи систем електронної контент-комерції шляхом зменшення затрат на розроблення і збільшення відвідуваності системи за рахунок застосування уніфікованих методів та програмних засобів опрацювання інформаційних ресурсів. Метою роботи визначено необхідність вирішення таких завдань:

- створення формальної моделі систем електронної комерції для визначення недоліків існуючих методів та засобів опрацювання інформаційних ресурсів;
- розроблення уніфікованих методів опрацювання інформаційних ресурсів для проектування архітектури системи електронної контент-комерції;

- розроблення узагальненої типової архітектури системи електронної контент-комерції для реалізації етапів життєвого циклу комерційного контенту;
- реалізація програмних засобів опрацювання інформаційних ресурсів у системах електронної контент-комерції для скорочення часу і затрат на формування, управління та супровід комерційного контенту, підвищення якості опрацювання інформаційних ресурсів шляхом використання випробуваних вирішень.

Аналіз отриманих наукових результатів

Основними підсистемами опрацювання інформаційних ресурсів в СЕКК є формування, управління та супровід контенту, які мають наступні відношення: *контент* ® *формування контенту* ® *база даних* ® *управління контентом* ® *інформаційний ресурс*, або *запит користувача* ® *управління контентом* ® *інформаційний ресурс* ® *супровід контенту* ® *база даних*, тобто

$$Function = Formation \cup Management \cup Support, \quad (3)$$

де *Function* – оператор формування вихідних даних у системі; *Formation* – оператор формування Web контенту; *Management* – оператор управління Web контентом; *Support* – оператор супроводу Web контенту.

Модель системи електронної контент комерції подано як

$$S = \langle X, Q, Formation, H, C, V, Management, Support, Z, T, Y \rangle, \quad (4)$$

де $X = \{x_1, x_2, \mathbf{K}, x_{n_x}\}$ – множина контенту з різних джерел даних; $Q = \{q_1, q_2, \mathbf{K}, q_{n_Q}\}$ – множина інформаційних запитів користувачів системи; *Formation* – оператор формування контенту; $H = \{h_1, h_2, \mathbf{K}, h_{n_H}\}$ – множина умов формування та управління контентом; $C = \{c_1, c_2, \mathbf{K}, c_{n_C}\}$ – множина комерційного контенту; $V = \{v_1, v_2, \mathbf{K}, v_{n_V}\}$ – множина умов супроводу контенту та зовнішніх впливів середовища на систему; *Management* – оператор управління контентом; *Support* – оператор супроводу контенту; $Z = \{z_1, z_2, \mathbf{K}, z_{n_Z}\}$ – множина компонентів інформаційного ресурсу; $T = \{t_1, t_2, \mathbf{K}, t_{n_T}\}$ – час транзакцій опрацювання інформаційних ресурсів; $Y = \{y_1, y_2, \mathbf{K}, y_{n_Y}\}$ – множина статистичних даних роботи системи

Процес формування Web контенту описано оператором *Formation* вигляду

$$c_r = Formation(u_f, x_i, t_p), \quad (5)$$

де u_f – множина умов формування контенту, тобто $u_f = \{u_1(x_i), \mathbf{K}, u_{n_U}(x_i)\}$.

Комерційний Web контент подано так:

$$c_r = \left\{ \bigcup_f u_f \mid (x_i \in X) \wedge (\exists u_f \in U), U = U_{x_i} \vee U_{\overline{x_i}}, i = \overline{1, m}, f = \overline{1, n} \right\}. \quad (6)$$

Етап управління комерційним Web контентом описано оператором *Management* вигляду

$$z_w = Management(q_d, c_r, h_k, t_p), \quad (7)$$

де $Q = \{q_1, q_2, \mathbf{K}, q_{n_Q}\}$ – множина запитів користувачів; H – множина умов управління контентом, тобто $H = \{h_1(c_{i+1}, q_d), \mathbf{K}, h_{n_H}(c_{i+n_H}, q_d)\}$. Управління контентом подано як

$$z_w = \left\{ \bigcup_{k=1}^{n_H} h_k(c_{i+1}, q_d) \mid (c_{i+k} \in C) \wedge (q_d \in Q) \wedge (h_k \in H_q), H = H_{q_d} \vee H_{\overline{q_d}}, i = \overline{1, n_C}, d = \overline{1, n_Q}, k = \overline{1, n_H} \right\}. \quad (8)$$

Етап супроводу комерційного Web контенту *Support* описано функцією вигляду

$$y(t_p + \Delta t) = \text{Support}(v_l, h_k, c_r, z_w, t_p, \Delta t), \quad (9)$$

де v_l – множина умов супроводу комерційного Web-контенту та зовнішніх впливів середовища на систему, тобто $v_l = (v_l(q_i, h_k, c_r, z_w, t_p), \mathbf{K}, v_{n_v}(q_i, h_k, c_r, z_w, t_p))$. Вихідні дані реалізовано як

$$y_j = \left\{ \bigcup_l v_l \left| \begin{array}{l} (\exists q_d \in Q) \wedge (\exists z_w \in Z) \wedge (\forall v_l \in V) \wedge (\forall (c_r \wedge q_d) \in h_k), \\ V = V_{q_d} \vee \overline{V_{q_d}}, d = \overline{1, n_Q}, l = \overline{1, n_V}, w = \overline{1, n_Z}, r = \overline{1, n_C}, k = \overline{1, n_H} \end{array} \right. \right\}. \quad (10)$$

Для повнофункціональної системи електронної контент-комерції характерна складна система взаємозв'язаних операцій, методів, прийомів опрацювання комерційного Web-контенту.

Метод формування комерційного контенту. Процес формування комерційного Web-контенту для інформаційного ресурсу забезпечує зв'язок між множиною вхідних даних з різних джерел інформації та множиною сформованого комерційного контенту, збереженого у відповідній базі даних у системах електронної контент-комерції, тобто $\text{Source}(x_i) \rightarrow x_i \rightarrow X \rightarrow \text{Formation}(u_f, x_i, t_p) \rightarrow c_r \rightarrow C \rightarrow \text{DataBase}(C)$, де $\text{Source}(x_i)$ – джерело x_i контенту, x_i – i -й контент з джерела; X – множина даних з відповідного джерела; $\text{Formation}(u_f, x_i, t_p)$ – оператор формування комерційного контенту з x_i у фіксований час t_p при умовах u_f ; c_r – сформований за умовою u_f r -й контент; C – множина контенту; $\text{DataBase}(C)$ – база даних контенту.

Модель формування комерційного Web контенту подано як

$$\text{Formation} = \left\langle \begin{array}{l} X, \text{Gathering}, \text{Formatting}, \text{KeyWords}, \text{Backup}, \\ \text{Caterization}, \text{BuDigest}, \text{Dissemination}, T, C \end{array} \right\rangle, \quad (11)$$

де X – множина вхідних даних $x_i \in X$ з різних інформаційних ресурсів або від модераторів; Gathering – оператор збирання/створення контенту з різних джерел; Formatting – оператор форматування комерційного контенту; KeyWords – оператор виявлення ключових слів і понять контенту; Caterization – оператор автоматичної рубрикації комерційного контенту; Backup – оператор виявлення дублювання змісту контенту; BuDigest – оператор формування дайджестів комерційного контенту; Dissemination – оператор вибіркового поширення контенту; T – час $t_p \in T$ транзакції формування контенту; C – множина контенту $c_r \in C$.

Оптимальне рішення, здатне допомогти орієнтуватися в динамічній вхідній інформації з різних джерел, надають процеси збирання інформації з джерел і поширення її фрагментів за потребами користувачів $C = \text{Gathering}(X, U_G, T)$, де X – множина контенту з різних джерел даних, U_G – множина умов збирання даних з різноманітних джерел; Gathering – оператор збирання/створення контенту; T – час збирання/створення контенту.

Задачу виявлення дублювання змісту контенту описано оператором вигляду $C = \text{Backup}(\text{Gathering}(X, U_G, T), U_B)$, де X – множина контенту з різних джерел даних, U_B – множина умов виявлення дублювання змісту контенту; Backup – оператор виявлення дублювання контенту, C – множина контенту. Дубльований за змістом контент у системах електронної контент-комерції виявляють за допомогою лінгвостатистичних методів знаходження загальних термів, ланцюжки яких утворюють словесні сигнатури контенту.

Технологія синдикації контенту містить процес навчання програм збирання даних структурним особливостям окремих джерел (з інформаційних ресурсів, від модераторів, користувачів, відвідувачів, журналістів, редакторів), безпосереднє сканування контенту та приведення до загального формату в XML $C = \text{Formatting}(\text{Backup}(\text{Gathering}(X, U_G, T), U_B), U_{FR})$,

де *Formatting* – оператор форматування контенту; U_{FR} – множина умов форматування контенту. Опрацювання множини контенту C для виявлення значущих ключових слів побудовано за принципом знаходження ключових слів за змістом (термами), ґрунтується на законі Зіпфа і полягає у виборі слів із середньою частотою появи (найвживаніші слова ігноруються за допомогою стоп-словника, а рідкісні слова з текстів повідомлень не враховуються). Виявлення ключових слів та понять визначається оператором $KeyWords(C)$ та описано оператором вигляду $C = KeyWords(Formatting(Backup(Gathering(X, U_G, T), U_B), U_{FR}), U_K)$, де U_K – множина умов виявлення ключових слів в тексті, тобто $U_K = \{Unicity, Noun, NumbSymb, UseFrequency\}$, де *Unicity* – унікальність терму ≥ 80 , *Noun* – терм є іменником, словосполученням іменників або прикметника з іменником, *NumbSymb* – кількість знаків без пробілів, *UseFrequency* – частота вживання терму в тексті комерційного контенту. Множина умов рубрикації $U_{Ct} = \{TKeyWords, FKeyWords, Comparison, QuantitativelyTKey\}$, де *TKeyWords* – тематичні ключові слова; *FKeyWords* – частота вживання ключових слів; *Comparison* – порівняння вживання ключових слів різних тематик, *QuantitativelyTKey* – частота вживання тематичних ключових слів у контенті.

Задачі класифікації та розподілу контенту реалізують через інформаційно-пошукову систему вибіркового поширення контенту. Контент аналізують на відповідність запитам оператором $C_{Ct} = Categorization(KeyWords(C, U_K), U_{Ct})$, де *Categorization* – оператор рубрикації контенту згідно із виявленими ключовими словами, U_{Ct} – множина умов автоматичної рубрикації, C_{Ct} – множина рубрикованого релевантного контенту. Множину дайджестів C_D формують залежністю $C_D = BuDigest(C_{Ct}, U_D)$, де *BuDigest* – оператор формування дайджестів, U_D – множина умов формування дайджестів контенту, тобто $C_D = BuDigest(Categorization(KeyWords(C, U_K), U_{Ct}), U_D)$. Релевантний контент розсилають користувачам та завантажують до баз даних. Вибіркове поширення контенту описано $C_{Ds} = Dissemination(C_D, U_{Ds})$, де C_{Ds} – множина вибірково поширеного контенту, U_{Ds} – множина умов вибіркового поширення контенту, *Dissemination* – оператор вибіркового поширення контенту.

Метод управління комерційним Web-контентом. Процес управління комерційним контентом подано такою загальною схемою зв'язків: $User(q_d) \rightarrow q_d \rightarrow Q \rightarrow H(c_{i+1}, q_d) \rightarrow Management(q_d, c_r, h_k, t_p) \rightarrow z_w \rightarrow User(z_w)$. Методи управління комерційним контентом класифіковано відповідно до виду генерації сторінок інформаційного ресурсу: генерація сторінок за запитом користувача, генерація сторінок при редагуванні модератором та змішаний тип.

1. Модель управління комерційним Web-контентом типу генерації сторінок за запитом користувача системи електронної контент-комерції подано як

$$Management_Q = \langle H, C, Q, Presentation, Edit, Weight, T, Z \rangle, \quad (12)$$

де C – множина контенту; Z – множина сформованих сторінок інформаційного ресурсу; Q – множина запитів користувачів; *Presentation* – оператор формування та подання сторінки інформаційного ресурсу; *Edit* – оператор редагування та модифікації контенту; *Weight* – вага блоку контенту; T – час транзакцій управління контентом; H – множина умов управління контентом.

Етап редагування та модифікації Web контенту подано оператором вигляду $c_j(t_{r+1} + \Delta t) = Edit(c_j(t_r), t_{r+1}, \Delta t)$. Етап формування множини сторінок описано оператором $Z(t_r + \Delta t) = Presentation(q_i, C, Weight(C), t_r, \Delta t)$, де *Weight(C)* – загальна вага контентного блоку інформаційного ресурсу, тобто

$$z_i = \left\{ \bigcup c_j(q_i, t_r) \mid (\forall c_j \in C_q) \wedge (\exists q_i \in C_q) \wedge Weight(C_q), C = C_q \vee C_{\bar{q}}, j = \overline{1, m}, i = \overline{1, n} \right\}. \quad (13)$$

2. Модель генерації сторінок при редагуванні інформаційного ресурсу в системі електронної контент-комерції має вигляд

$$Management_E = \langle C, H, Edit, Weight, T, Z \rangle, \quad (14)$$

де C – множина контенту; H – множина умов управління контентом, Z – множина статичних сторінок інформаційного ресурсу; $Weight$ – вага блоку контенту; $Edit$ – оператор редагування/модифікації контенту. Етап формування сторінок описано оператором вигляду $Z(t) = Edit(C, Weight, H, t)$.

3. Модель змішаного управління комерційним Web контентом подано як

$$Management_M = \langle C, Q, H, Presentation, Edit, Weight, Caching, T, Z \rangle, \quad (15)$$

де C – множина контенту; Z – множина сформованих сторінок інформаційного ресурсу; Q – множина запитів; H – множина умов управління контентом; $Presentation$ – оператор формування та подання сторінки інформаційного ресурсу; $Edit$ – оператор редагування та модифікації контенту, $Caching$ – оператор формування кешу або інформаційних блоків контенту, де $Cache = Caching(C, Weight, t, \Delta t)$ або $Cache = Caching(Z, Weight, t, \Delta t)$, тобто, відповідно:

$$Cache = \left\{ \bigcup c_i \mid (\forall c_i \in C) \wedge Weight(c_i), t = t + \Delta t, i = \overline{1, n} \right\}, \quad (16)$$

$$Cache = \left\{ \bigcup z_i \mid (\exists z_i \in Z) \wedge (\forall c_j \in C) \wedge Weight(c_j) \wedge (\forall c_j \in z_i), t = t + \Delta t, i = \overline{1, n} \right\}. \quad (17)$$

Метод супроводу комерційного Web контенту. Процес супроводу комерційного контенту подано такою загальною схемою зв'язків $User(q_d, z_w) \rightarrow q_d \rightarrow z_w \rightarrow V(q_d, z_w, t_p) \rightarrow Support(v_l, h_k, c_r, z_w, t_p, \Delta t) \rightarrow y_j \rightarrow Moderator(y_j)$. Моделю супроводу комерційного контенту має вигляд

$$Support = \langle Q, C, H, V, T, BuInfPort, IdThemTop, ConCorrTablConc, CalRankConc, Z, Y \rangle, \quad (18)$$

де Q – множина запитів користувачів; C – множина Web контенту; $BuInfPort$ – оператор формування інформаційних портретів; $IdThemTop$ – оператор виявлення тематичних сюжетів; $ConCorrTablConc$ – оператор побудови таблиць взаємозв'язку контенту; $CalRankConc$ – оператор розрахунку рейтингів контенту; H – множина внутрішніх параметрів системи; V – множина параметрів впливу зовнішнього середовища; Z – множина сторінок інформаційного ресурсу; T – множина часу транзакцій опрацювання інформаційних ресурсів; Y – множина статистичних даних роботи системи.

У результаті аналізу функціонування системи електронної контент-комерції S та супроводу контенту C формується множина $Y = \{Y_P, Y_T, Y_C, Y_R\}$ відповідно до умов $V = \{V_P, V_T, V_C, V_R\}$, де $Y_P = Y_{P_c} \vee Y_{P_q}$ – підмножина інформаційних портретів контенту Y_{P_c} та користувачів Y_{P_q} , Y_T – підмножина тематичних сюжетів контенту; Y_C – підмножина таблиць взаємозв'язку контенту; Y_R – підмножина рейтингів контенту; $V_P = V_{P_c} \vee V_{P_q}$ – множина умов формування інформаційних портретів; V_T – множина умов виявлення тематичних сюжетів; V_C – множина умов побудови таблиць взаємозв'язку контенту; V_R – множина параметрів розрахунку рейтингів контенту. Множина інформаційних портретів комерційного контенту Y_{P_c} подано як $Y_{P_c} = BuInfPort(V_{P_c}, C, H, Q, T)$, а множина портретів

користувачів Y_{Pq} подано як $Y_{Pq} = BuInfPort(V_{Pq}, Q, H, Z, T)$, де $V_P = V_{Pc} \vee V_{Pq}$ – множина умов формування портретів; $BuInfPort$ – оператор формування портретів $Y_P = Y_{Pc} \vee Y_{Pq}$.

Множину тематичних сюжетів комерційного Web контенту Y_T подано як $Y_T = IdThemTop(C, H, Q, V_T, T)$, де V_T – множина умов виявлення сюжетів контенту; $IdThemTop$ – оператор визначення тематичних сюжетів контенту Y_T . Множину таблиць взаємозв'язку контенту Y_C подано як $Y_C = ConCorrTablConc(C, V_C, T)$, де V_C – множина умов побудови таблиць взаємозв'язку, $ConCorrTablConc$ – оператор побудови таблиць взаємозв'язку. Множину рейтингів комерційного контенту Y_{Rc} як $Y_{Rc} = CalRankConc(C, Q, H, Y_C, V_{Rc}, Spam, Tonality, T)$, а множину рейтингів модераторів Y_{Rm} подано як $Y_{Rm} = CalRankConc(C, Q, H, Y_C, V_{Rm}, T)$, де $V_R = V_{Rc} \vee V_{Rm}$ – множина параметрів для розрахунку рейтингів контенту, $Tonality(Q^+, Q^0, Q^-, T, H)$ – критерій тональності комерційного контенту, $Spam(Q, T)$ – оператор визначення фільтрування коментарів, $CalRankConc$ – оператор визначення рейтингів контенту та модераторів $Y_R = Y_{Rc} \vee Y_{Rm}$.

Множину вихідних статистичних даних Y подано як

$$Y = \{Y_P, Y_T, Y_C, Y_R\} = Support(V, C, Q, H, Z, T, \Delta T); \quad (19)$$

$$Y = \{Y_P, Y_T, Y_C, Y_R\} = Support(V_P, V_T, V_C, V_R, C, Q, H, Z, T, \Delta T), \quad (20)$$

де $Y_P = Y_{Pc} \vee Y_{Pq}$ – підмножина інформаційних портретів контенту та користувачів; Y_T – підмножина тематичних сюжетів контенту; Y_C – підмножина таблиць взаємозв'язку контенту; $Y_R = Y_{Rc} \vee Y_{Rm}$ – підмножина рейтингів контенту та модераторів; $Support$ – оператор супроводу контенту.

Подано програмні реалізації розроблених систем з підсистемами опрацювання інформаційних ресурсів при організації е-комерції через інтернет-газети та інтернет-журнали (табл. 1).

Таблиця 1

Підсистеми опрацювання комерційного контенту в розроблених системах електронного бізнесу

№ з/н	Information Resource Name	System Appointment	Address	Subsystem availability		
				Formation	Management	Support
1	Fotoghalereja-vysocjkykh	Online shop	fotoghalereja-vysocjkykh.com	+/-	+	+/-
2	Vgolos	Online Newspaper	vgolos.com.ua	+	+	+
3	Tatjana	Online Newspaper	tatjana.in.ua	-	+/-	+/-
4	Presstime	Online Newspaper	presstime.com.ua	+/-	+	+
5	AutoChip	Online shop	autochip.vn.ua	-	+	+/-
6	Kursyvalyut	Online Journal	kursyvalyut.com	+	+	-
7	Good morning, accountant!	Online Newspaper	dobryjranok.com	+	+	+/-
8	Victana	Online Journal	victana.lviv.ua	+/-	+/-	+/-
9	Information for Businesses	Online Journal	goodmorningua.com	-	-	-

Таблиця 2

Результати роботи систем за період часу з 10.2010 to 03.2013р.

Characteristics	The developed e-business systems								
	1	2	3	4	5	6	7	8	9
Visiting	4144	5997052	646	3654456	7838	9124	8724	517	25
Visits time	4:50	2:14	6:22	2:04	1:58	1:30	2:27	8:38	8:12
Failures/refusals	55,62	71,90	47,5	83,08	54,45	82,76	68,15	35,8	48,0
Goal Conversion	4,35	0,00	0,00	0,00	10,97	0,00	0,00	0,00	0,00
Unique visitors	2674	2501402	256	1501202	5704	7632	4996	250	7

Characteristics	The developed e-business systems								
	1	2	3	4	5	6	7	8	9
Repeat visits	1470	3495641	393	2153254	2134	1500	3728	267	18
Pageviews	19082	11588861	2924	11588861	20828	16952	18892	3760	81
Pages/visit	4,60	1,93	4,51	1,67	2,66	1,86	2,17	7,27	3,24
% New Visits	64,48	41,68	39,3	39,88	72,75	83,56	57,23	47,9	28,0
New Visitor	64,53	41,71	39,5	34,76	72,77	83,58	57,27	48,4	28,0
Returning Visit	35,47	58,29	60,5	65,24	27,23	16,42	42,73	51,6	72,0
Search traffic	56,47	36,10	32,5	26,12	69,67	79,68	24,34	23,0	20,0
Direct Traffic	19,50	11,20	38,5	31,22	17,15	14,17	48,73	32,9	80,0
Transitions (%)	24,03	52,48	28,97	42,46	13,18	6,15	26,94	44,10	0,00
Language – ru	45,15	55,57	33,90	43,54	66,87	54,39	69,01	49,7	52,0
Language – uk	36,20	23,70	57,3	33,34	10,23	25,51	10,49	39,3	48,0
Ukraine	81,25	91,22	83,9	87,56	67,17	66,01	89,47	87,8	60,0
Russia	9,10	1,84	6,78	4,67	24,04	5,57	0,62	0,39	0,00
Poland	0,19	0,29	0,01	0,12	0,11	0,61	0,11	0,19	0,00
United States	1,21	0,66	0,01	0,32	0,01	0,57	0,07	0,01	0,00
Germany	0,34	0,32	0,15	0,11	0,20	0,26	0,08	0,01	0,00
City – Kyiv	17,25	23,62	29,1	25,45	20,09	22,86	36,16	28,4	48,0
City – L'viv	38,92	43,41	48,4	34,54	10,17	14,43	9,70	45,1	12,0
Dnipropetrovsk	2,82	2,72	0,00	0,45	3,22	2,51	6,63	1,35	0,01
City – Kharkiv	1,50	2,82	0,62	0,65	3,56	1,65	6,08	1,16	0,01
City – Ternopil	2,24	0,01	0,01	0,01	0,01	2,17	0,01	0,01	0,01

На рис. 1, 2 подано результати роботи розроблених систем у вигляді графіків, з яких випливає, що за наявності всіх етапів життєвого циклу контенту суттєво збільшується обсяг відвідувань та унікальних користувачів.

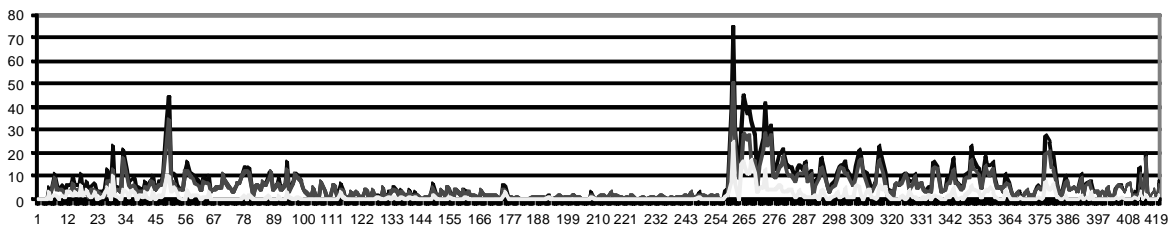


Рис. 1. Розподіл часу перебування на сайті та трафіку

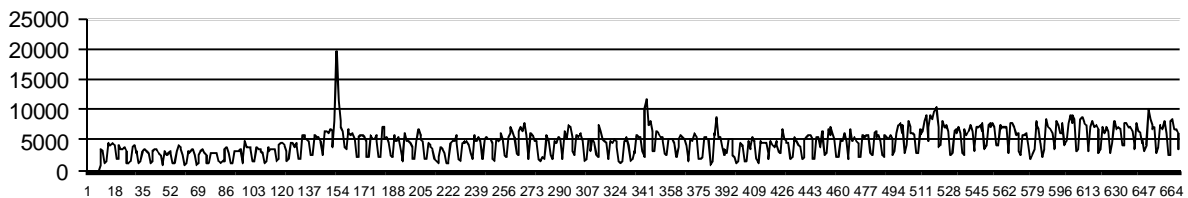


Рис. 2. Розподіл відвідувань та відмов відвідування сайтів

Сервіс ведення статистики відвідувань інформаційного ресурсу дає змогу оцінити трафік інформаційного ресурсу та ефективність маркетингових заходів, наприклад, для Victana (рис. 3).

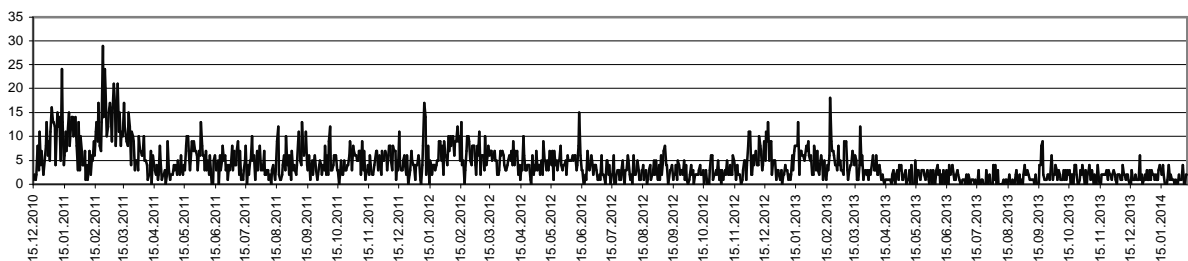


Рис. 3. Загальна кількість відвідувачів за 2010–2014 рр.

Наявність підсистем формування, управління та супроводу комерційного контенту в системах електронної контент-комерції сприяє збільшенню кількості унікальних відвідувачів, переглянутих сторінок, часу відвідування та відвідування інформаційних ресурсів і збільшення цільової та регіональної аудиторії.

Висновки і перспективи подальших наукових розвідок

У роботі розв'язано актуальну наукову задачу дослідження і розроблення методів та засобів опрацювання інформаційних ресурсів систем електронної контент-комерції із використанням розробленої класифікації, математичного та програмного забезпечення та узагальненої архітектури систем електронної контент-комерції. Розроблено модель системи електронної комерції, що дало змогу описати життєвий цикл комерційного контенту, побудувати структуру системи електронної контент-комерції та методи опрацювання інформаційних ресурсів в системах електронної контент-комерції. Розроблено метод формування контенту, метод управління контентом та метод супроводу контенту для досягнення ефекту від роботи на рівні розробника систем (скорочення часу і затрат на розроблення, підвищення якості шляхом використання випробуваних вирішень). Вдосконалено структуру системи електронної контент-комерції, відмінну від існуючих наявністю підсистем опрацювання інформаційних ресурсів, що дало можливість реалізації етапів життєвого циклу контенту. Розроблено рекомендації з проектування структури систем електронної контент-комерції, відмінні від існуючих більшою деталізацією етапів та наявністю засобів формування, управління та супроводу комерційного контенту, що дають змогу ефективно реалізовувати опрацювання інформаційних ресурсів на рівні розробника систем (зменшення ресурсів та часу на розроблення, покращення якості роботи систем електронної контент-комерції). Розроблено та впроваджено прикладне програмне забезпечення формування, управління, супроводу комерційного контенту для досягнення ефекту від роботи на рівні власника (підвищення рентабельності, зростання інтересу користувачів) та користувача (зрозумілості, спрощення інтерфейсу, уніфікація, розширення вибору) СЕKK.

1. Береза А. Електронна комерція / А. Береза, І. Козак, Ф. Левченко. – К.: КНЕУ, 2002. – 326 с.
2. Берко А. Системи електронної контент-комерції / А. Берко, В. Висоцька, В. Пасічник. – Л: НУЛП, 2009. – 612 с.
3. Большакова Е. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика / Е. Большакова, Д. Ландэ, А. Носков, Э. Клышинский, О. Пескова, Е. Ягунова. – М: МИЭМ, 2011. – 272 с.
4. Брайчевский С. Современные информационные потоки / С. Брайчевский, Д. Ландэ // Научно-техническая информация. – 2005. – № 11. – С. 21-33.
5. Клифтон Б. Google Analytics: профессиональный анализ посещаемости веб-сайтов / Б. Клифтон. – М: Вильямс, 2009. – 400 с.
6. Корнеев В. Базы данных. Интеллектуальная обработка информации / В. Корнеев, А. Гареев, С. Васютин, В. Райх. – М: Нолидж, 2000. – 352 с.
7. Ландэ Д. Основы моделирования и оценки электронных информационных потоков / Д. Ландэ, В. Фурашев, С. Брайчевский, О. Григорьев. – К: Інжиніринг, 2006. – 348 с.
8. Ландэ Д. Основы интеграции информационных потоков: монография / Д. Ландэ. – К: Інжиніринг, 2006. – 240 с.

ФОРМИ НЕЧІТКОСТЕЙ У БАЗАХ ДАНИХ ТА ЗНАНЬ

© Седушев О., Буров Є., 2014

Описано та проаналізовано різноманітність форм нечіткостей у базах даних та знань. Було досліджено та порівняно усі дані, що є неточними, розмитими, невизначеними, неповними і т.д. Показано, що нечіткі знання представляються у вигляді нечітких продукційних правил, які містять нечіткі дані у антецедентах та консеквентах.

Ключові слова: форми нечіткостей, нечіткі дані, нечіткі знання, нечіткі продукційні правила.

This paper describes and analyzes the whole variety of forms of fuzziness in databases and knowledge bases. All kinds of data that is imprecise, vague, uncertain, incomplete etc. have been researched and compared. It is shown that fuzzy knowledge is represented through fuzzy rules, which contain fuzzy data in the antecedent and consequent parts.

Key words: forms of fuzziness, fuzzy data, fuzzy knowledge, fuzzy production rules.

Вступ та постановка проблеми

Класичні моделі даних часто страждають від власної неспроможності зобразити чи маніпулювати даними та знаннями, які є неточними, невизначеними, розмитими тощо. Проте саме такі дані та знання усе частіше знаходяться у сьогодишніх інформаційних системах, базах та сховищах даних, базах знань, оскільки багато предметних областей, що є джерелами даних, містять прояви різних форм нечіткостей.

Постає необхідність у виявленні форм таких даних та знань, здатності їх розрізнати, вносити у моделі даних та знань у певному вигляді, перетворювати та навіть усувати за необхідності.

Усю сукупність різнорідних форм даних, що є неточними, нечіткими, невизначеними, розмитими, неповними тощо, називатимемо неklasичними даними. Приводом для цього є те, що природа і походження усіх форм таких даних різняться, а відтак виникає потреба у неоднакових прийомах для роботи з тією чи іншою формою неklasичних даних. Klasичними можна вважати такі дані, що є точними, визначеними, цілісними, чіткими, не мають втрачених чи пропущених значень і т.д. Klasичні дані переважали у апаратах багатьох наук доти, поки Л. Заде не почав формувати математичні засади нечіткої логіки (60-ті рр. ХХ ст.) [1].

Станом на сьогодні, коло задач, вирішуваних за допомогою використання апарату нечітких множин та нечіткої логіки, значно розрослося та містить такі області, як аналіз даних, розпізнавання, дослідження операцій, моделювання складних систем, підтримка прийняття рішень тощо.

Переважно важлива інформація для системи надходить з двох джерел:

1) від людей-експертів, які описують свої знання про предметну область тощо, використовуючи природну мову, що означає породження суб'єктивності та нечіткості;

2) від приладів, датчиків, вимірювачів, математичних моделей, що означає породження неточності та нецілісності.

Тому збереження експертних оцінок та думок, неточних даних вимагає знання неklasичних даних, а також уміння з ними працювати. Це також стосується видобування та інтерпретації даних з баз нечітких даних та знань, де неklasичні дані зустрічаються найчастіше.

Загалом переважна більшість сучасних систем СУБД та СУБЗ, контролерів, приладів та програмних застосунків вимагають наявності механізму, що дозволяє підтримувати та керувати неklasичними даними та нечіткими знаннями, а також їх аналізувати.

Аналіз останніх досліджень і публікацій

У [2] йдеться про те, що зберігання неklasичних даних у базах даних дозволяє СУБД краще відповідати на запити користувачів, оскільки запити користувачів є зазвичай розмитими та неточними. Також автори роблять першу спробу класифікувати форми проявів неklasичних даних. Подібно до цього, у [3] описується онтологія для представлення нечітких знань та неklasичних даних за допомогою класів, слотів, екземплярів. Ця онтологія є інтуїтивним засобом, що дає змогу користувачам, які не є експертами, задавати особливу нечітку інформацію без допомоги експерта, що знає структуру каталогів.

Сьогодні актуальним завданням є створення експертних систем діагностування, які можуть опрацювати нечітку діагностичну інформацію. У базах знань таких систем міститься кількісна та якісна інформація (представлена переважно лінгвістичними змінними), яка характеризує стан об'єкта діагностики. У [4] автор описує механізм нечіткого логічного висновку для експертної системи діагностування комп'ютерних засобів, який дає змогу враховувати нечіткі якісні знання у процесі діагностики, що покращує якість роботи такої системи за неповноти опису ситуації.

У [5] йдеться про форми проявів неklasичних даних, що можуть зберігатися у сучасних нечітких базах даних. Наведено різноманітні приклади відносно різних типів нечітких баз даних. Акцент роблять на нечітких об'єктно-орієнтованих базах даних, які можуть справлятися як з комплексними об'єктами, так і з нечіткими запитам та різними формами нечіткості та невизначеності у кортежах та сутностях.

Автори [6] розглядають проблему прийняття рішень в умовах невизначеності на основі застосування продукційних правил нечіткої логіки. Вони описують структуру та функції системи нечіткого прийняття рішень.

Формулювання цілі статті

Метою цієї статті є побудова таксономії форм неklasичних даних та нечітких знань, їхнє дослідження та порівняння. Побудова такої таксономії, визначення особливостей походження та форм нечіткостей необхідна для аналізу та розроблення методів видобування, інтерпретації, збереження та опрацювання нечітких даних та знань.

Виклад основного матеріалу

Неповні (incomplete), нецілісні (inconsistent), невизначені або сумнівні (uncertain), дво- та багатозначні (ambiguous), нечіткі або розмиті (fuzzy, vague), неточні (imprecise), null дані є можливими формами неklasичних даних. Нижче буде подано характеристики, приклади та підформи кожної з вищезазначених форм. Варто зазначити, що кожна форма неklasичних даних містить прояв чогось (наприклад, неповноти, нецілісності, неточності, нечіткості тощо).

1. Неповні дані

Неповнота зазвичай означає відсутність або недостачу значення, а також неточну інформацію, де множина можливих значень охоплює увесь домен можливих значень. Неповне опрацювання даних трапляється, коли процес отримання даних не відбувся вчасно. Неповні дані виникають внаслідок втрачених оновлень, поганого зчитування, відсутності значень, проходження файлу даних за недостатню кількість циклів [7].

2. Нецілісні дані

Поняття нецілісності стосується швидше зберігання самих даних у різних моделях, аніж даних як таких. Нецілісність – це семантичний конфлікт, який означає, що один і той самий аспект даного чи одне і те саме значення даного відображене по-різному, частіше за все помилково. Наприклад, зарплатня співробітника *X* збережена одночасно у вигляді 1000 у.о. та 2000 у.о. у тій самій базі даних або в різних базах даних. Нецілісність інформації зазвичай спричинена процесом інтеграції (об'єднання) інформації з різних вхідних джерел. Поверхнево нецілісні дані свідчать про те, що вони є взаємосуперечливими або ненадійними.

– нецілісність даних, спричинена ненадійними джерелами

Така нецілісність може виникати, якщо при злитті інформації джерела даних були ненадійними або недостовірними (табл. 1).

Вибірка з відношень бази даних “Потримані авто”

Власник	Марка автомобіля	Номерний знак	Рік випуску	Пробіг, км
Гнатів Іван	BMW X5	AA 6549 ЛО	2001	100000
Іван Гнатів	BMW X5	AA 6549 ЛО	2003	200000

Згідно з табл. 1 нецілісність у значеннях атрибутів *Пробіг* та *Рік випуску*. У цьому випадку важко зрозуміти, який із кортежів (перший чи другий) є коректним. Якщо клієнт бази даних захоче придбати авто, то на дані з якого кортежу він має орієнтуватися? Ця нецілісність могла виникнути тоді, коли власник авто вніс свою інформацію з помилками у різні бази даних.

3. Невизначені або сумнівні дані

Невизначеність породжується тоді, коли носій інформації утворив суб'єктивну думку чи дав суб'єктивну оцінку істинності деякого факту, у якому він не впевнений на усі 100 %. Спотвореність такого роду інформації робить достеменно неможливим визначення її правдивості чи фальшивості. Усе, що можна зробити в цій ситуації, – це оцінити ймовірність такої інформації бути правдивою чи хибною на деякому нескінченному проміжку значень (зазвичай це інтервали $[0, 1]$ та $[0, 100]$, де перше та останнє значення ідентифікують 100 % вірність та 100 % хибність інформації відповідно) [8].

Отже, невизначеність даних пов'язана із ступенем достовірності їхніх значень. Для прикладу, ймовірність (ступінь достовірності) того, що автомобіль *BMW X5* (табл. 1) був випущений у 2001 році, становить 50 %.

– невизначеність даних, спричинена статистичними показниками чи аналізами

Багато даних (виміри, дані анкетувань тощо) отримані статистично і по суті є завжди сумнівними. Крім того, вимірювальні прилади можуть вимірювати щось із певною похибкою (це спричинено неідеальністю пристроїв), а отже, продукують виміряні дані хоч і з мінімальною, але сумнівністю (невизначеністю).

– невизначеність даних, пов'язана із засекреченням та захистом інформації

Багато даних є невизначеними чи сумнівними через те, що їхнє розголошення є небажаним.

4. Дво- та багатозначні дані

Такі дані являють собою певну неясність, оскільки можуть бути інтерпретовані (трактовані) по-різному (двозначно, багатозначно тощо). Найпоширеніші випадки саме двозначності. Загалом двозначність означає, що деякі дані в силу певних обставин позбавлені певної семантичної незалежності та однозначності, що призводить до появи додаткових інтерпретацій.

– двозначні дані, спричинені використанням аббревіатур

Нерідко буває, що використання аббревіатур призводить до плутанини під час трактування даних. У такому випадку слід явно розшифровувати збережене значення, а не скорочувати його аббревіатурою.

– двозначні дані, спричинені неповним контекстом

Розглянемо базу даних, що зберігає певні погодні дані [2]. У ній можуть зберігатися середньоденні значення температур для того чи іншого міста (табл. 2).

Вибірка з відношень бази даних “Погодні умови”

Місто	Дата	Температура
Львів	09.09.2013	20
Львів	10.09.2013	22

Як видно з табл. 2, температурні значення представлені цілочисловими значеннями. Проте не вказано, чи температуру подано у градусах Цельсія чи Фаренгейта.

У тій же табл. 2 наведено дати, які можуть зберігатися у будь-якому дозволеному форматі. Проте чи відомо, з чого починаються дати, з днів чи місяців? З вказаних даних можна зробити два

висновки: або представлені дати відповідають 9 вересню 2013 року, та 10 вересню 2013 року, або 9 вересню 2013 року та 9 жовтню 2013 року. В Україні прийнято першим числом позначати день, а другим – місяць. У США – навпаки.

– двозначні дані, спричинені різним слідуванням слів

Такі дані трапляються в тому випадку, коли дане представлене семантично коректним способом, проте різними шляхами. Якщо подивитися на значення атрибуту *Власник* із табл. 1, то можна побачити, що вони однакові, проте записані по-різному. В цьому випадку це не має великого значення (хіба що проявляється певна неузгодженість та брак вищої нормальної форми для відношення), тому що ім'я *Гнатів* не існує, і ми однозначно розуміємо, що це прізвище людини. Проте трапляються такі повні імена, як *Богдан Борис*. Отут вже однозначно висловитися не можна, що є ім'я, а що прізвище.

5. Нечіткі, або розмиті дані

Нечіткий (або розмитий) означає певний ступінь вираження, коли не можна чітко та ясно сказати про сенс та значення даних. Для того, щоб перевести нечіткість (розмитість) у площину чіткості та ясності, застосовують певні математичні механізми – такі, як нечітка логіка.

Нечіткі дані містять деякий нечіткий предикат (наприклад, “старий”, “молодий”, “низький”, “високий”). Зазвичай такі нечіткі предикати моделюються лінгвістичними змінними.

Лінгвістичною називається змінна, яка набуває значення з нечіткої множини слів або словосполучень деякої природної або штучної мови. Множина допустимих значень лінгвістичної змінної називається терм-множиною. Термом називається будь-який елемент терм-множини. В теорії нечітких множин терм формалізується нечіткою множиною за допомогою функції належності M_F .

Якщо змоделювати будь-яку лінгвістичну змінну як нечітку підмножину значень на інтервалі $[0, \infty)$ з функцією належності M_F на проміжку $[0, 1]$, то проєкція $[0, \infty) \rightarrow [0, 1]$ є математичним описом значення лінгвістичної змінної.

Наприклад, нехай лінгвістична змінна $L = \{\text{низький, високий, дуже високий}\}$. Тоді графічно значення лінгвістичної змінної L за допомогою нечіткої логіки зображено на рис. 1.

Інтервали та проміжки на рисунку вказують лише на те, що кожна людина розуміє та інтерпретує значення *низький, високий, дуже високий* по-різному. Варто зазначити, що при заданні значень лінгвістичної змінної останні повинні відповідати деяким обмеженням, серед яких впорядкованість та обмеженість, повнота та узгодженість, нормальність.

6. Неточні дані

Неточні дані не є помилковими та не порушують цілісності інформаційної системи. Неточність є продуктом існування значення, яке не можна виміряти з належною точністю.

– неточні дані, спричинені наявністю вибору (вибіркові дані)

- Завжди істинні

Як такі розуміють дані, які можуть набувати альтернативних значень з певної множини, проте достовірність значень таких даних завжди дорівнюватиме одиниці, оскільки їхня ймовірність є рівнозначною та завжди істинною.

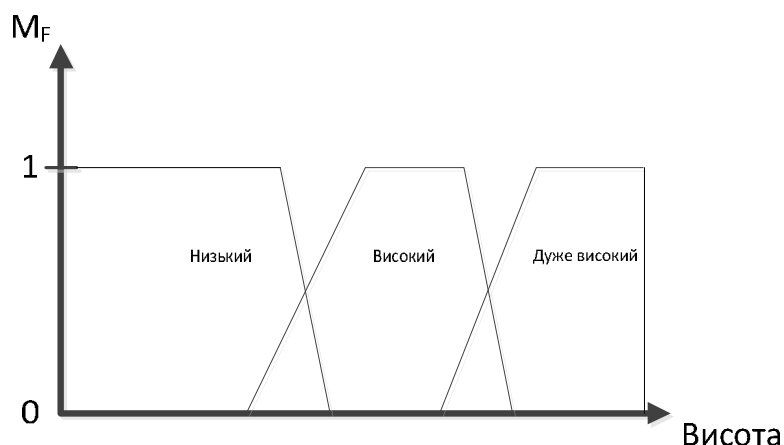


Рис. 1. Графічне моделювання значень лінгвістичної змінної L за допомогою нечіткої логіки

Якнайкраще для прикладу таких даних підходять маршрутні дані (табл. 3).

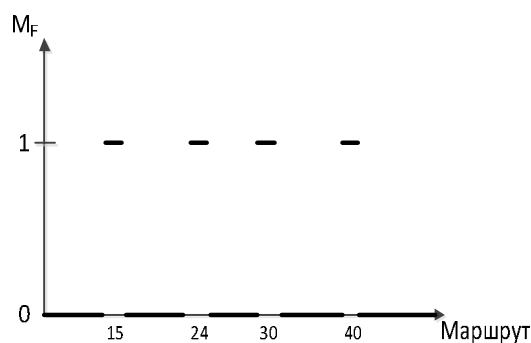
Таблиця 3

Маршрутні дані

Місце призначення	Маршрутне таксі, №
Шевченківський гай	15 або 24 або 30 або 40

Щоб дістатися Шевченківського гаю, необхідно скористатися одним із вищеперелічених маршрутних таксі. Різниця у виборі маршруту немає, оскільки усі вони є рівноцінними. Але це породжує неточність, тому що необхідно зробити вибір із множини значень, щоб отримати точне значення. Цю ситуацію можна змоделювати за допомогою нечіткої логіки (рис. 2).

Рис. 2. Представлення неточних вибіркових даних за допомогою нечіткої логіки



- **Ймовірнісні**

Ці дані схожі на завжди істинні дані, оскільки значення необхідно вибрати із представленої дискретної множини (цілочислового інтервалу). Різниця полягає у тому, що достовірність значень таких даних зовсім не дорівнює одиниці (причому може не дорівнювати ніколи), а, отже, породжує різну математичну ймовірність. Можна сказати, що значення таких даних коливаються у певних межах, причому кожна позначка у таких межах має свою ймовірність. Нехай реальна тривалість існування планети Землі коливається від 3 до 6 мільярдів років (табл. 4). Кожен варіант має свою ймовірність. Грунтуючись на ймовірностях, необхідно вибрати один з варіантів.

Таблиця 4

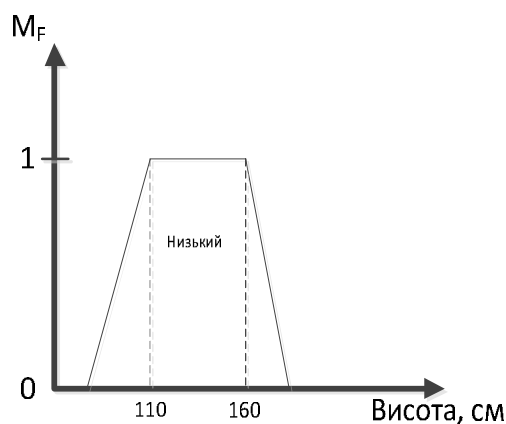
Тестові дані

Тривалість існування Землі, млрд. років	Ймовірність
3	70%
4	90%
5	92%
6	73%

– **неточні інтервальні дані**

Неточні інтервальні дані означають, що їхні значення є істинними на певному інтервалі, а не у конкретній точці. Графічно такі дані зображено на рис. 3.

Рис. 3. Представлення неточних інтервальних даних за допомогою нечіткої логіки



– неточні дані, спричинені можливою похибкою

Подібно до неточних інтервальних даних, значення неточних даних з можливою похибкою можуть набуватися лише із певного інтервалу. Але різниця полягає у тому, що інтервал має бути нечітким синглтоном. Нечіткий синглтон означає, що лише одне скінченне значення із діапазону значень є істинним та чітким. Завдяки нечіткій логіці можна сказати, що функція приналежності такого значення дорівнює одиниці, тоді як приналежність інших значень завжди < 1 . Математично такі дані представляються як $D \pm \delta$, де D – це значення, а δ – допустима (можлива) похибка.

Нехай існує блок живлення із напругою 400 ± 10 В. Тоді графічно це можна зобразити, як на рис. 4.

– дані із *NULL* значеннями

Такі дані є критичним випадком неточних даних. *NULL* значення переважно позначає відсутність інформації. Головною проблемою таких даних є те, що їх можна інтерпретувати по-різному. Найпоширенішими інтерпретаціями є такі:

- значення є невідомим (воно існує, але невідоме);
- значення не існує.

Дослідивши та проаналізувавши форми неklasичних даних, підведемо підсумки (табл. 5).

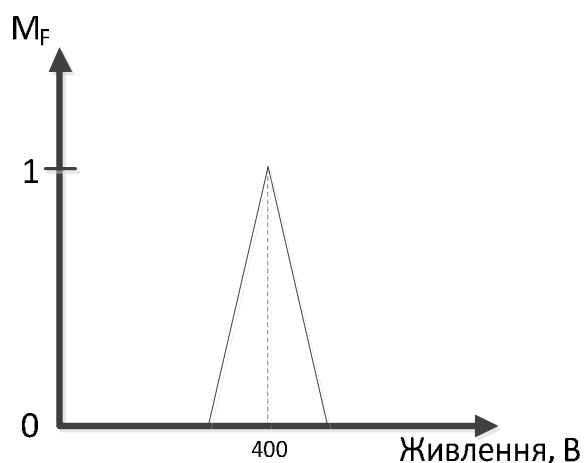


Рис. 4. Графічне представлення неточних даних, спричинених похибкою за допомогою нечіткої логіки

Таблиця 5

Форми неklasичних даних та їхні характеристики

№ з/п	Форма неklasичних даних	Прояв у даних	Найхарактерніші ознаки	Причини виникнення
1	2	3	4	5
1	Неповні дані	Неповнота	Відсутність або нестача значення	<ul style="list-style-type: none"> • Дані отримано невчасно • Втрачене оновлення • Відсутність значення у джерелі даних
2	Нецілісні дані	Нецілісність	Одне і те саме дане має різне значення в одному й тому самому, або в різних джерелах даних	<ul style="list-style-type: none"> • Об'єднання інформації з різних джерел • Помилковість занесених значень
3	Невизначені (сумнівні) дані	Невизначеність	Ймовірність істинності (хибності) даного є невідомою	<ul style="list-style-type: none"> • Суб'єктивність думки, оцінки тощо • Статистичні похибки або похибки вимірювань
4	Неоднозначні (дво- та багатозначні) дані	Неоднозначність	Певна кількість можливих інтерпретацій	<ul style="list-style-type: none"> • Використання абrevіатур • Відсутність деякої інформації у контексті • Різне слідування слів

1	2	3	4	5
5	Нечіткі (розмиті) дані	Нечіткість	Смисл, значення даних є неясним, розмитим	<ul style="list-style-type: none"> • Певні суб'єктивні представлення та оцінки того чи іншого факту, явища тощо
6	Неточні дані	Неточність	Значення не можна виміряти із певною точністю	<ul style="list-style-type: none"> • Наявність вибору із певної рівноймовірної або ймовірнісної (нерівноймовірної) множини значень • Значення представлене інтервалом • Наявність похибки

Більшість з форм неklasичних даних можуть вкраплюватися у нечіткі знання. Нечіткі знання – це насамперед інформація, отримана від експертів. До нечіткості знань призводять неточність та розмитість експертних оцінок та суджень, недовизначеність понять та термінів, можлива неформальність природної мови.

Прикладом нечітких знань є оцінки експертів: “строк придатності близько 5 років”, “старий прилад”, “температура повітря від 10 до 15” тощо. У цих оцінках легко побачити вкраплення вищенаведених форм неklasичних даних. Ще очевиднішими є такі прояви у нечітких продукційних правилах, завдяки яким можуть представлятися знання у базі нечітких знань. Використовувати нечіткі продукційні правила зручно для експертів для подання нечітких знань – це спрощує їх аналіз обчислювальними пристроями.

Нечіткі продукційні правила мають такий абстрактний вигляд:

$$\text{ЯКЩО } \langle \text{нечітке судження} \rangle, \text{ ТО } \langle \text{нечітке судження} \rangle \quad (1)$$

У (1) нечітке судження в умовній частині (ЯКЩО) має назву антецедента, а нечітке судження в результуючій частині (ТО) – консеквента.

Нечіткі судження можуть бути двох типів, а саме:

1) атомарні нечіткі судження, які містять лише одну умову. Наприклад, температура є низькою;

2) складені нечіткі судження, які містять n умов, сполучених сполучниками “І (AND)”, “АБО (OR)”, “НЕ (NOT)”. Наприклад, температура НЕ є низькою І вологість є середньою.

Найпоширенішими формами нечітких продукційних правил є правила, які у антецеденті містять атомарні або складені нечіткі судження, а у консеквенті – лише атомарні судження. Якщо у консеквенті нечіткого правила міститься складене нечітке судження, то до такого правила можна застосувати прийоми декомпозиції та розкласти його на множину правил з консеквентами, які містять лише атомарні нечіткі судження.

До найуживаніших форм належать:

1) канонічні нечіткі продукційні правила такого вигляду:

$$\text{Правило } R_i : \text{ЯКЩО } x_1 \in A_1 \text{ I } \dots \text{ I } x_n \in A_n, \text{ ТО } y \in B, \quad (2)$$

де R_i – мітка (назва) i -го правила у базі нечітких знань (тут та надалі), $x = (x_1, x_2, \dots, x_n)$ – n -вимірний вхідний вектор лінгвістичних змінних (тут та надалі), A, B – нечіткі множини відповідно (тут та надалі), y – вихідна лінгвістична змінна.

Цю форму правил ввів М. Мамдані. Головною перевагою (2) є зрозуміла та прозора лінгвістична інтерпретація.

Приклад: ЯКЩО дорога є слизькою І дорога є крутою, ТО водіння є небезпечним.

Існують й інші форми представлення канонічних нечітких продукційних правил. У середині 80-х рр. ХХ ст. Такагі та Сугено запропонували використовувати нечіткі продукційні правила, які містять лінійні функції у консеквентній частині правила [9]. Тобто вихідний результат визначався у не вигляді лінгвістичної змінної, а деякої лінійної функції.

Правило R_i : *ЯКЩО* $x_1 \in A_1 \text{ I } \dots \text{ I } x_n \in A_n$, *ТО* $y = f_i(x)$,

де $f_i(x) = b_{i0} + b_{i1}x_1 + \dots + b_{in}x_n$, b_{iq} – дійсне число (тут i в подальшому).

Нечіткі правила, представлені у такому форматі, мають властивість високої наближеності до бажаного результату. Проте лінгвістична інтерпретація таких правил є дещо гіршою порівняно з (2).

Приклад: *ЯКЩО* x_1 є маленьким *I* x_2 є маленьким, *ТО* $y = 0.7 - 0.6x_1 + 0.8x_2$

Також активно використовується у задачах класифікації та розпізнавання образів така форма канонічних нечітких правил [10]:

Правило R_i : *ЯКЩО* $x_1 \in A_1 \text{ I } \dots \text{ I } x_n \in A_n$, *ТО* $y = \text{Class}_q$

де Class_q – мітка класу, до якого належить об'єкт, що класифікується чи розпізнається.

Приклад: *ЯКЩО* x_1 є маленьким *I* x_2 є великим, *ТО* $y = \text{Class } 1$

2) *нечіткі продукційні правила зі сполучником "АБО"*

Вигляд таких правил описано нижче.

Правило R_i : *ЯКЩО* $x_1 \in A_1 \text{ I } \dots \text{ I } x_m \in A_m$ *АБО* $x_{m+1} \in A_{m+1} \text{ I } \dots \text{ I } x_n \in A_n$, *ТО* $y \in B$ (3)

(3) можна розписати так:

ЯКЩО $x_1 \in A_1 \text{ I } \dots \text{ I } x_m \in A_m$, *ТО* $y \in B$

ЯКЩО $x_{m+1} \in A_{m+1} \text{ I } \dots \text{ I } x_n \in A_n$, *ТО* $y \in B$

Загалом такий формат використовують для того, аби скоротити кількість правил у базі нечітких знань, оскільки консеквент в обох випадках є спільним.

Висновки

Проілюструвавши множину форм неklasичних даних та нечітких знань, зробимо висновок:

- кожна форма неklasичних даних має свій підхід до інтерпретації та опрацювання, на що варто зважати при таких задачах, як інтелектуальний аналіз даних;
- наведена множина форм неklasичних даних не є остаточною та може бути розширена у майбутньому;
- нечіткі знання тісно переплітаються з неklasичними даними, а саме містять їх, що впливає на трактування перших, а також на можливі сфери застосування нечітких продукційних правил.

Наведена таксономія форм нечітких даних та знань є основою для аналізу та узагальнення методів та задач видобування даних з баз нечітких знань, що згодом буде досліджено.

1. Zadeh L.A. *Fuzzy Sets // Information and Control, Vol.8, 1965. – P. 338-353.* 2. Popat D. *Classification of Fuzzy Data in Database Management System / Popat, D., Sherda, H., Taniar, D. // Proceedings of 8th International KES Conference, Wellington, New Zealand, 2004. - P. 691-697.* 3. Blanco I.J. *About the Use of Ontologies for Fuzzy Knowledge Representation / Blanco, I.J., Marin, N., Martinez-Cruz, C., Vila, M.A. // Proceedings of the Joint 4th Conference of the European Society for Fuzzy Logic and Technology, Barcelona, Spain, 2005. – P. 106-111.* 4. Гнатчук Є.Г. *Моделювання нечіткого логічного висновку процесу діагностування комп'ютерних засобів // Вісник Вінницького політехнічного інституту. – Вінниця: ВНТУ. – 2005. - №6(63). – С. 220-224.* 5. Ma Z.M. *A Literature Overview of Fuzzy Database Models / Ma, Z.M., Yan, L. // J. Inf. Sci. Eng. №24, 2008. – P. 189-202.* 6. Кравець П. *Системи прийняття рішень з нечіткою логікою / П. Кравець, Р. Куркало // Вісник Національного університету "Львівська політехніка". – Львів. – 2009. - №650. – С. 116-123.* 7. Motro A. *Uncertainty Management in Information Systems: From Needs to Solutions / Motro, A., Smets, P. – Springer, 1997. – 464 p.* 8. Parsons S. *Current Approaches to Handling Imperfect Information in Data and Knowledge Bases // Knowledge and Data Engineering IEEE, Vol.8, №3, 1996. - P. 483-488.* 9. Sugeno M. *Fuzzy Identification of Systems and It's Applications to Modeling and Control / Sugeno, M., Takagi, T. // IEEE Trans. On Systems, Man, and Cybernetics №15, 1985. – P. 116-132.* 10. Ishibuchi H. *Pattern Classification with Linguistic Rules / Ishibuchi, H., Nojima, Y. // Fuzzy Sets and Their Extensions: Representation, Aggregation and Models Studies in Fuzziness and Soft Computing, Vol. 220, 2008. – P. 377-395.*

¹⁾ Cherkassy state technological university,
Chair of informatics and information security

²⁾ Institute of applied system analysis
of National technical university of Ukraine “KPI”

³⁾ Cherkassy state technological university,
Chair of informatics and information security

MODEL BASED DECISION SUPPORT SYSTEM FOR FORECASTING FINANCIAL PROCESSES

© Kozhukhivska O., Bidyuk P., Kozhukhivskiy A., 2014

Підвищення якості розв’язку задачі прогнозування фінансово-економічних процесів – актуальна задача, яка виникає на виробничих підприємствах, в інвестиційних компаніях, у банківській системі, транспортних компаніях та багатьох інших напрямках діяльності. Високоякісні оцінки прогнозів дають можливість підвищити якість рішень, що приймаються на їх основі. Незважаючи на те, що у цьому напрямі виконано велику кількість досліджень, існує необхідність розв’язання множини задач, спрямованих на прискорення та підвищення якості розв’язання задач такого класу. Зокрема, існує необхідність створення комп’ютерних систем підтримки прийняття рішень (СППР), орієнтованих на побудову високо адекватних математичних моделей та обчислення прийнятних за якістю оцінок коротко- та середньострокових прогнозів.

Ключові слова: модель, прогнозування фінансових та економічних процесів, принципи системного аналізу.

A computer based decision support system is proposed the basic tasks of which are adaptive model constructing and forecasting of financial and economic processes. The system is developed with the use of system analysis principles, i.e. the possibility for taking into consideration of some stochastic and information uncertainties, forming alternatives for models and forecasts, and tracking of the computing procedures correctness during all stages of data processing. A modular architecture is implemented that provides a possibility for the further enhancement and modification of the system functional possibilities with new forecasting and parameter estimation techniques. A high quality of final result is achieved thanks to appropriate tracking of the computing procedures at all stages of data processing: preliminary data processing, model constructing, and forecasts estimation. The tracking is performed with appropriate set of statistical quality parameters. Examples are given for modeling and forecasting of nonlinear and nonstationary financial and economic processes. The examples show that the system developed has good perspectives for the practical use. It is supposed that the system will find its applications as an extra tool for decision making when developing the strategies for enterprises of various types.

Key words: model, forecasting, financial and economic processes, system analysis principles.

Introduction

The forecasting problems are to be solved practically in all areas of human activities. However, the problems of mathematical modeling, estimation and forecasting process dynamics are particularly urgent for micro- and macroeconomics, banking sphere, insurance, investment companies, industrial enterprises that are functioning in conditions of tough competition, and many others kind of activities. There are many ideologically different approaches to mathematical description of processes dynamics and their volatility that are based on known statistical and recently developed techniques of intellectual data analysis.

Volatility forecasts are used widely as a measure of various kinds of financial risks in the frames of Value-at-Risk (VaR) and other methodologies. The market and some other types of risks are estimated with different modifications of VaR methodology that provides a possibility to reach practically acceptable quality of risk estimates [1, 2]. For forecasting financial processes and enterprises bankruptcy the nonlinear classification type models have found wide application, for example, logistic regression as well as support vector machine (SVM), fuzzy logic, neural networks and neuro-fuzzy techniques, Bayesian networks, decision trees, and combinations of the approaches mentioned [3 – 5].

Selection and application of a specific technique for process description and forecasts estimation depends on application area, availability of statistical data, qualification of personnel, who work on the financial analysis problems, and availability of appropriate applied software. Better results for estimation of financial processes forecasts is usually reached with application of ideologically different techniques combined in the frames of one computer based system. Such approach to solving the problems of quality forecasts estimation can be implemented in the frames of modern decision support systems (DSS). DSS is a powerful instrument for supporting decision making as far as it combines a set of appropriately selected data processing procedures aiming to reach final result of high quality – objective high quality alternatives for a decision maker. Development of a DSS is based on modern theories of system analysis, information processing systems, estimation theory, mathematical and statistical modeling and forecasting, decision making theory as well as many other results of theory and practice of processing data and expert estimates [6, 7].

The paper considers the problem of DSS constructing for solving the problems of modeling and forecasting processes evolution in time with the possibility for application of alternative data processing techniques, modeling and estimation of parameters and states for the processes under study.

Problem formulation. The purpose of the study is as follows: 1) analysis and development of requirements to the modern decision support systems; 2) development of the system architecture; 4) selection of mathematical modeling and forecasting techniques for financial and economic processes; 3) illustration of the system application to solving the problem of financial and economic processes forecasting with statistical data.

General requirements to modern DSS

Modern DSS are rather complex multifunctional computing systems with architecture of hierarchical type. Define DSS formally as follows:

$$DSS = \{DKB, PDP, DT, SE, PE, FG, DQ, MQ, FQ, AQ\},$$

where *DKB* – data and knowledge base; *PDP* – a set of procedures for preliminary data processing; *DT* – a set of statistical tests for determining possible effects contained in data; *SE* – a set of procedures for estimation of mathematical model structure; *PE* – a set of procedures for estimation of mathematical model parameters; *FG* – forecasts generating procedures; *DQ, MQ, FQ, AQ* – the sets of statistical quality criteria for estimating quality of data, models, forecasts, and alternatives, accordingly.

Such systems should satisfy the following general requirements: 1) – contain highly developed bases of data, mathematical models, quality criteria and rules, as well as necessary computational procedures; 2) – their interface should be user friendly, convenient and simple for use, as well as adaptive for the users of various levels (e.g., engineering and managerial staff); 3) – the hierarchy of a system functioning should correspond to the hierarchic process of human decision making; 4) – the system should possess an ability for learning in the process of its functioning, i.e. accumulate appropriate knowledge regarding possibilities of solving the problems of definite (selected) class; 5) – the organization and techniques for computing procedures should provide for appropriate rate of computing that corresponds to the human requirements with regard to the rate of alternatives generation and reaching the final result; 6) – computing (precision) quality should satisfy preliminary established requirements; 7) – intermediate and final results of computations should be controlled with appropriate sets of analytic quality criteria, what will enhance significantly quality and reliability of the final result; 8) – DSS should generate all necessary for a user forms and types of intermediate and final results representations with taking into consideration

the users of various levels; 9) – the system should contain the means for exchange with data and knowledge with other information processing systems via local or global computer nets; 10) – DSS should be easily expandable with new functions.

Satisfaction of all the requirements mentioned above provides a possibility for effective practical application of a system developed and enhancing general behavioristic effect of the DSS as a whole for a specific company or an enterprise.

General and special purpose mathematical tools for DSS

All mathematical methods that are hired for development and implementation of DSS could be divided into two following groups: – general purpose methods that provide for implementation of system functions, and special purpose methods that are necessary for solving specific problems regarding data processing, model constructing, alternatives generating, selecting the best alternative for implementation and forecasting of the implementation consequences.

The group of the general purpose methods includes the following methods: – data and knowledge collecting and editing procedures; – preliminary data processing techniques such as digital filtering, normalization, imputation of missing values, detecting special effects (regime switching, seasonal effects, trends etc); – the methods for accumulating information regarding previous applications of DSS to problem solving for the retrospective use; – computer graphics techniques; – techniques for syntactic analysis to be used in command interpreter; – methods for organizing communications with other information processing systems via local and global nets; – logical rules to control the system functioning. The set of the methods mentioned could be modified or expanded depending on specific application.

Selection of the application defined mathematical methods for a DSS depends on the specific system application area, possible problem statements regarding data processing, model building, processes forecasting, and alternatives generation. However, it is possible to state that in most cases of DSS development it is necessary to use the following mathematical methods: – methods and methodologies for mathematical (statistical and probabilistic) modeling using statistical/experimental data; – forecasting techniques on the basis of the models constructed with possibilities for combining the forecasts computed with different techniques; – operations research optimization techniques and dynamic optimization (optimal control) methods; – the methods for forecasting/foresight of decision implementation consequences; – the sets of special analytic criteria to control the processes of computations performed at each stage of data processing and alternative generation aiming to reach high quality of final results.

All the methods and methodologies mentioned are described well in special modern literature. For example, time series modeling and forecasting are presented in many references, more particularly in [8, 9]. The task for a DSS developer is in appropriate selection of model classes, modeling and optimization techniques, quality criteria as well as relevant methodologies for organizing computational processes.

Generation and implementation of alternatives with DSS

Decision making process includes rather sophisticated procedures that could be partially or completely iterative, i.e. executed repeatedly when the alternative found is not satisfactory for a decision making person (DMP). DSS can return automatically (or on DMP initiative) to the previous stages of data and knowledge analysis.

The whole process of making and implementing decision could be considered as consisting of the stages given below.

1 – A thorough analysis of the decision problem using all available sources of information, collection of data and knowledge relevant to the problem. At this stage it is also important to consider and use former solutions to the problem if such are available. The information regarding former solutions of similar problem can be helpful for correcting problem statement, to select appropriate techniques for data analysis, to speed up alternative generation, and to decline the alternatives that turned out to be ineffective in the past.

2 – Selection of a class (classes) of mathematical models for the problem description, and analysis of a possibility for the use of available (previously developed) models. The models could belong to different classes as far as they can be formulated in continuous or discrete time, be linear or nonlinear, they can be

developed according to the structural or functional approach etc. In some cases it is necessary to construct complex simulative model that would include a set of simpler models of various classes.

3 – Development of new models for the problem (process, object, system) under study what includes structure and parameter estimation for candidate models using available data (and possibly expert estimates) and knowledge of various types. The alternative structures of candidate models provide a possibility for selecting the best one of them for generating alternative decisions (forecasts, control actions, risk estimates etc) on their bases.

4 – Analysis of the candidate models constructed and selecting of the best one of them with application of a set of statistical quality criteria and expert opinion (estimation). At this stage again more than one model can be selected for the further use as far as the best model (for a particular application) can be found only after application of the candidates for solving particular problem, i.e. after alternatives generating and estimating possible consequences of their implementation.

5 – Application of the model (models) selected to solving forecasting and/or control problem (when necessary). If the forecasts or controls computed are not satisfactory we should return back to stage one or stage three, and repeat the process of model constructing. At this stage another set of statistical quality criteria should be applied to the analysis of forecasts or controls determined.

6 – Generating of a set of alternatives with the use of the model (models) constructed and various admissible initial conditions and constraints on variables. In a case of controls generating the alternatives could be built with different optimality criteria, utility functions or other criteria.

7 – Analysis of the alternatives generated with the experts of an enterprise or a company, and final selection of the best one for practical implementation. In a case when no alternative is acceptable we should return back to the model constructing or alternative generating stages. New knowledge or data can be required for the next iteration of computing the decision support.

8 – Planning of actions and estimation of financial, material and human resources that are necessary for implementation of the alternative selected. Determining of a time horizon (horizon of control) necessary for implementing the decision made.

9 – Implementation of the decision made: current monitoring of availability and spending the necessary resources, estimation of necessary time frames, registering and quality estimation of intermediate and final results.

10 – Application of possible analytic and expert quality criteria to estimation of final results.

11 – Analysis of the final results achieved by the company experts, and final elucidation of advantages and disadvantages of the alternative implemented; analysis of the decision making and implementing process, and forming forecasts (foresights) for the future.

12 – Writing the final report on the tasks performed.

Architecture of DSS for forecasting of financial and economic processes

DSS architecture is a generalized large-scale representation of system elements with links between them. Architecture gives a notion for the general purpose of system constructing and its basic functions (Fig. 1). DSS functionality is controlled by user commands, correctness of which is monitored by the command interpreter which constitutes a part of user interface. The user commands are implemented by the main operation module that coordinates functioning of all system elements. The specific commands and actions can be the following: expanding and modification of bases available in the system; initiation and starting of data and knowledge processing procedures; model constructing, forecasts estimation and alternative generating; viewing intermediate and final results of computing; retrospective analysis of previous results of decision making; comparing of current results with previous.

The system interface is considered as its basic element from the point of view of its presentation to the user. This is justified by the fact that interface construction influences substantially convenience as well as rate and effectiveness of user interaction with the system [10]. The principles of interface constructing and its implementation create a separate special task that is not considered here.

It is clear that architecture, given in Fig. 1, is highly generalized. It means that practically the same type of architecture could be used to construct DSS for solving rather wide class of problems that require statistical/experimental data processing, mathematical modeling, optimal state and parameter estimation for dynamic systems, forecasting the future process evolution and making decisions on this basis.

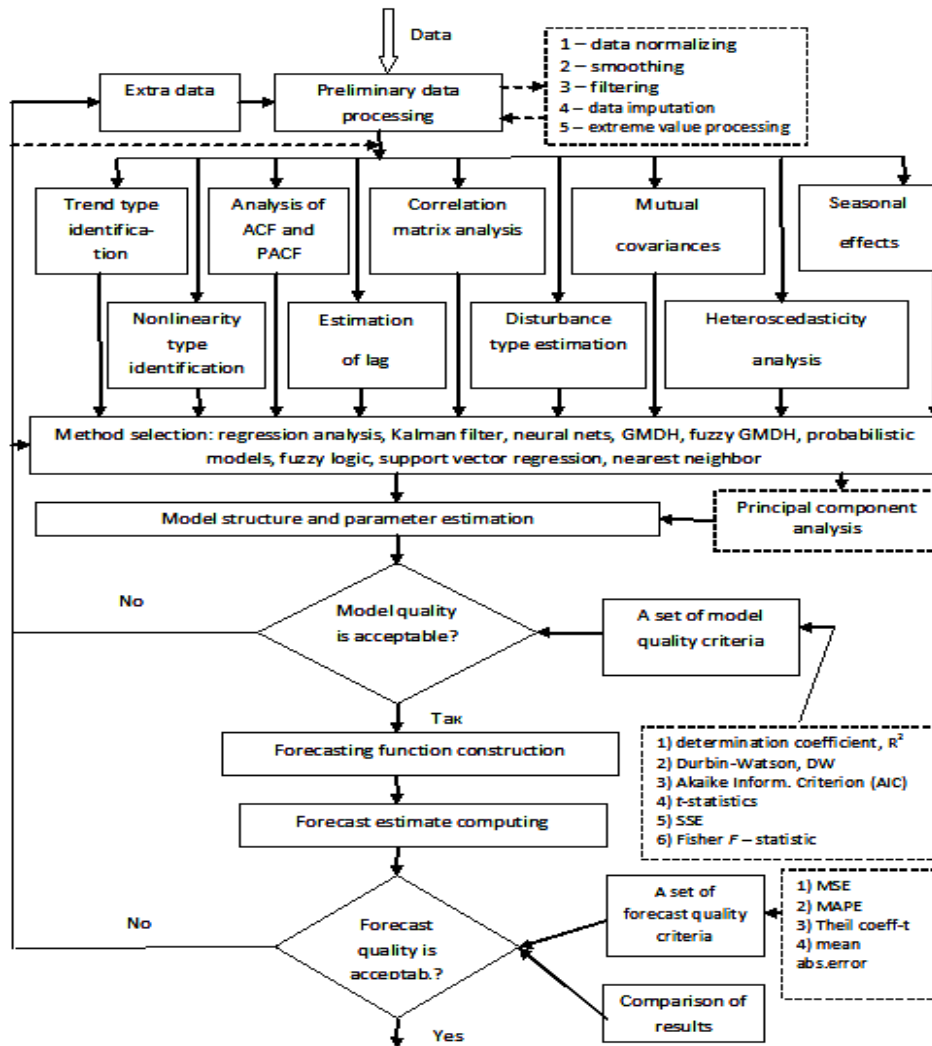


Fig. 1. DSS architecture for forecasting economic and financial process

Coping with uncertainties

The problem of identifying and taking into consideration various uncertainties is a very important one and is practically always present in decision making procedures. Here we do not consider the cases when uncertainties are taken into consideration with expert estimates. Though expert estimates are not excluded from the process of alternative generation. For example, expert estimate can be used for selecting special types of mathematical models that for some reason cannot be chosen automatically due to sophisticated structure or necessity to apply special estimation procedures. Expert judgment can also be useful for final selection of the best alternative from the set of generated decisions.

Statistical data uncertainties such as skipped measurements, extreme values and high level jumps of unknown origin could be processed with appropriately selected statistical procedures. There exist a number of data imputation procedures that help to make collected data complete. For example, very often skipped measurements for time series can be generated with appropriately selected distributions. Processing of jumps and extreme values helps with adjusting data stationarity and to estimate correctly probability distribution.

Application of Kalman filter requires knowledge of covariance for state disturbances and measurement errors. As far as these parameters are often unknown it is useful to apply appropriate adaptive estimation algorithms that provide acceptable estimates for the statistical parameters. An experience of practical application of the filter shows that it better to use separate procedure for covariance estimation to avoid divergence of filtering algorithm.

Fuzzy logic can be hired for coping with the level uncertainties for variables when we consider linguistic variables instead of numerical ones. There are possibilities for transforming fuzzy values into

numerical and vice versa. Thus, there is no problem for processing fuzzy and exact variables in the frames of one computing procedure.

Probabilistic types of uncertainties regarding whether or not some event will happen can be taken into consideration with various probabilistic models. Among them are analysis of distributions, Bayesian networks and other possibilities. From the computational point of view it is easier to process discrete variables as far as they accept a final number of values. In this case probabilities are assigned to outcomes using a probability mass function (PMF). Mass function tells us what “weight” (or a mass) should be assigned to each outcome. The sum for all the masses is 1,0. In a case of dealing with continuous variables, that may accept infinite number of values, we use probability density function (PDF). An integral over the density function should be equal to 1,0. When more than one random variable is considered we have to use joint distribution functions.

Generally speaking the modern instrumentation for coping with uncertainties is very powerful and it should be used in the frames of decision support systems for enhancing their possibilities with respect to reaching the best models and forecasts, and the best possible decisions.

Data, model and forecasts quality criteria

To achieve reliable high quality final result of forecasting at each stage of computational hierarchy separate sets of statistical quality criteria have been used. Data quality control is performed with the following criteria:

- database analysis for missing values using developed logical rules, and imputation of missed values with appropriate techniques;
- analysis of data for availability of outliers with special statistical tests, and processing of outliers to reduce their negative influence on statistical properties of data;
- normalizing of data in a case of necessity;
- application of low-order digital filters (usually that’s low-pass filters) for separation of observations from measurement noise;
- application of principal component method to achieve desirable level of orthogonalization between the variables selected;
- computing of extra indicators for the use in regression models.

It is also useful to test how informative is the data collected. Very formal indicator for data informativeness is sample variance. It is considered formally that the higher is the variance the richer is data with information. Another criterion is based on computing derivatives with a polynomial that describes data in the form a time series. For examples, such polynomial may describe rather complex process trend as follows:

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + c_1 k + c_2 k^2 + \dots + c_m k^m + e(k), \quad (1)$$

where $y(k)$ is the main variable; a_i, c_i are model parameters; $k=0,1,2,\dots$ is discrete time which is linked to real continuous time t via data sampling period T_s as follows: $t=kT_s$; $e(k)$ is a random process that integrates influence of external disturbances to the process being modeled as well as model structure and parameters errors. Autoregressive part of the model (1) describes the deviations that are imposed on a trend, and the trend itself is described with the m -th order polynomial. In this case maximum number of derivatives can be m , though in practice actual number of derivatives is defined by the largest number i of the parameter c_i , that is statistically significant.

To select the best models constructed the following statistical criteria are used: determination coefficient (R^2); Durbin-Watson statistic (DW); Fisher F -statistic; Akaike information criterion (AIC), and residual sum of squares (SSE). The forecasts quality is estimated with hiring the following criteria: mean squared error (MSE); mean absolute percentage error ($MAPE$); and Theil inequality coefficient (U). To perform automatic model selection the following combined criterion is proposed:

$$V_N(q, D_N) = e^{|1-R^2|} + \ln\left(1 + \frac{SSE}{N}\right) + e^{|2-DW|} + \ln(1 + MSE) + \ln(MAPE) + e^U, \quad (2)$$

where q is a vector of model parameters; D_N – is a dataset of power N . The power of the criterion was tested experimentally and proved with a wide set of models and statistical data. Thus, the three sets of quality criteria are used to insure high quality of final result.

Some mathematical models used in DSS

When considering mathematical models it is important to use a unified notion of model structure which we define as follows:

$$S = \{ r, p, m, n, d, z, l \},$$

where r is model dimensionality (number of equations); p is model order (maximum order of differential or difference equation in a model); m is a number of independent variables in the right hand side; n is a nonlinearity and its type; d is a lag or output reaction delay time; z is external disturbance and its type; l are possible restrictions for the variables.

Generalized linear models (GLM). GLM can be considered as further enhancement of multiple linear regression (MLR). It is distinguished from MLR with the following elements:

- distribution of dependent variable can be non-Gaussian and not necessarily continuous, say binomial;
- predicted values of dependent variable are computed as linear combination of predictors that are linked to dependent variable via selected link function.

Generally, GLM create a class of statistical models that includes linear regression, variance analysis relations, nonlinear models like logit and probit, Poisson regression and some others [11]. In a general linear model independent variable is supposed to be normally distributed and the link function is called identity function, i.e. linear combination of independent variables is not subjected to any transform.

In regression analysis the value of $h = \mathbf{X}\mathbf{b}$ is called a linear predictor, where \mathbf{X} is independent variables measurement matrix; \mathbf{b} is a vector of model parameters. When creating GLM instead of description in the form of $m = E[y]$ (where E is mathematical expectation; y is dependent variable as a function of linear predictor) some function $g(m)$ is described, i.e.

$$g(m) = h = \mathbf{X}\mathbf{b},$$

where $g(\cdot)$ is a link function. Thus, GLM is a model of the following type:

$$y = g^{-1} \left(\sum_{i=1}^m \mathbf{b}_i g_i(x) \right),$$

where m is a number of independent (explaining) variables. It is usually supposed that dependent variable y belongs to the class of exponential distributions. Thus, characteristics of GLM suppose the knowledge of dependent variable distribution, characteristics and parameters of the link function $g(\cdot)$, and of linear predictor $\mathbf{X}\mathbf{b}$. The class of exponential distributions includes the following distribution types: normal, gamma, and beta, and the discrete families – binomial, Poisson, and negative binomial. General representation of PDFs or PMFs for them is as follows:

$$f(x|\mathbf{q}) = h(x) c(\mathbf{q}) \exp \left(\sum_{i=1}^k w_i(\mathbf{q}) l_i(x) \right),$$

where $h(x) \geq 0$ and $l_1(x), \dots, l_k(x)$ are real-valued functions of the observation x (they cannot depend on \mathbf{q}); $c(\mathbf{q}) \geq 0$ and $w_1(x), \dots, w_k(x)$ are real-valued functions of the possibly vector-valued parameter \mathbf{q} (they cannot depend on x).

The following three basic functions are used in practice:

- logit:

$$h = \log \left(\frac{m}{1-m} \right);$$

- probit:

$$h = \Phi^{-1}(m),$$

where $\Phi(\cdot)$ – is a cumulative distribution function (CDF) for normal distribution;

– log-log function:

$$h = \log\{-\log(1-m)\}.$$

This type of link is of importance for short samples with positive mean value; it also can be presented as follows:

$$h = \frac{m^l - 1}{l},$$

with limiting value:

$$h = \log m \quad \text{as } l \rightarrow 0;$$

or

$$h = \begin{cases} m^l, & l \neq 0, \\ \log m, & l = 0. \end{cases}$$

Such forms provide a possibility for performing correct computations when $l \neq 0$, and $l = 0$.

Nonlinear forecasting models logit and probit. To solve the problem of classifying credit borrowers into two groups it is quite logically to use appropriately transformed CDF. CDF belongs to the class of monotonous functions that monotonously decrease or increase on some interval. Suppose that for determining probability of crediting a client p_c it is chosen a normal distribution:

$$p_c = \Phi(\mathbf{b}^T \mathbf{x}) = \int_{-\infty}^u j(z) dz,$$

where $j(z)$ is a density for standard normal distribution; $u = \mathbf{b}^T \mathbf{x}$ is upper integration limit. This way so called probit model is constructed.

If the probability for successful crediting is determined with logistic distribution function then logit model is constructed. In this case we have:

$$p_c = \Phi(\mathbf{b}^T \mathbf{x}) = \int_{-\infty}^u j(z) dz = \frac{1}{1 + \exp(-\mathbf{b}^T \mathbf{x})}, \quad (3)$$

or

$$p_c = \frac{\exp(b_1 x_1 + \dots + b_m x_m)}{1 + \exp(b_1 x_1 + \dots + b_m x_m)}.$$

In contrast to the normal distribution logistic function has so called closed form that provides a possibility for simplified computations in comparison to probit. Parameter estimates for both models can be found with maximum likelihood technique. An alternative possibility is Markov chain Monte Carlo (MCMC) approach that is based on correct generation of pseudorandom sequences, that satisfy certain conditions. Due to availability of multiple alternative techniques for generating pseudorandom sequences MCMC has found wide applications [12]. Classification results achieved with logit and probit are usually acceptable in most cases of application.

Bayesian networks (BN). Bayesian networks are probabilistic and statistical models represented in the form of directed acyclic graphs (DAG) with vertices as variables of an object (system) under study, and arcs showing existing causal relations between the variables. Each variable of BN is characterized with complete finite set of mutually excluding states. The relations between the variables are established via expert estimates or applying special statistical and probabilistic tests to statistical data (when available) characterizing variables dynamics. The process of constructing BN is generally the same as for models of other types, say regression models. For example, as model parameters for BN are unconditional and conditional probabilities for specific values of variables, that are stored in respective tables. For parent variables these are unconditional probabilities and for daughter variables – conditional probability tables (CPT). Unconditional and conditional probabilities are determined by experts (in simpler cases), and by special computational algorithms when appropriate sets of statistical (or experimental) data are available. Thus to each node of DAG is assigned CPT that is used for computing probabilistic inference over the BN [5, 13].

The process of constructing a model in the form of BN can be represented with the following steps: 1) – a thorough analysis of the process (object) under study aiming to detecting of its special functioning features and identification of parent and daughter variables; 2) – search and analysis of existing process models and determining the possibility of their usage in DSS; 3) – determining degree of relations between the process variables using special tests and expert estimates; 4) – reduction of the process dimensionality whenever this is possible; 5) – scaling and discretization of the data available when necessary; 6) – determining semantic restrictions on the future model; 7) – estimation of candidate model (directed acyclic graphs) structures using appropriate optimization procedures and score functions; 8) – candidate models analysis and selection of the best one using model quality criteria (including values of score functions); 9) – application of the model(s) constructed to solve the problem stated; 10) – computing inference with the model(s) constructed with regards to the variables selected, quality analysis of the result. In our case the final result of the model application is computing of client default probability with the conditions established by other model variables. According to alternative problem statement BM can be constructed for estimation of operational or other type of financial risks.

Support vector machine. Support vector machine (SVM) belongs to the group of techniques that determines classes with the limits for spaces. It also can be used for constructing SVM based regression models to solve forecasting problem. The support vectors are created with the vectors of data that lay on these limits. The classification result is successful if the space between the limits is empty. Usually SVM is hired to solve the problems of linear classification and regression analysis. The basic idea of SVM is in transformation of input vectors to the space of higher dimension with subsequent search of separating hyperplane with maximum distance in this space. Two parallel hyperplanes are built on both sides of separating hyperplane, and the separating hyperplane is the one that maximizes the distance between the two extra parallel hyperplanes. The algorithm is based on maximization of distance between the parallel hyperplanes what minimizes mean classification error.

The separating hyperplane is described with the equation [14]:

$$\mathbf{w}x - b = 0,$$

where \mathbf{w} is a perpendicular to separating hyperplane; x is a normalized real-valued point of data; b is the shortest distance between the separating hyperplane and coordinate origin. The problem of separating hyperplane construction is in minimizing $\|\mathbf{w}\|$ under condition that $c_i(\mathbf{w}x_i - b) \geq 1$, $1 \leq i \leq n$, where $c_i = 1$, if $(\mathbf{w}x_i - b) \geq 1$, and $c_i = -1$, if $(\mathbf{w}x_i - b) \leq -1$. This is a problem of quadratic optimization of the form:

$$\begin{cases} \|\mathbf{w}\|^2 \rightarrow \min, \\ c_i(\mathbf{w}x_i - b) \geq 1, 1 \leq i \leq n. \end{cases}$$

This problem can be transformed to equivalent quadratic optimization problem that includes only dual variables:

$$\begin{aligned} -L(I) &= -\sum_{i=1}^n I_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n I_i I_j c_i c_j (\mathbf{x}_i \mathbf{x}_j) \rightarrow \min_I, \\ I_i &\geq 0, 1 \leq i \leq n, \\ \sum_{i=1}^n I_i c_i &= 0. \end{aligned}$$

After solving the problem unknown \mathbf{w} and b can be found using the expressions:

$$\mathbf{w} = \sum_{i=1}^n I_i c_i \mathbf{x}_i; \quad b = \mathbf{w} \mathbf{x}_i - c_i, \quad I_i > 0.$$

The classification itself is performed with the expression:

$$a(x) = \text{sign}\left[\left(\sum_{i=1}^n I_i c_i \mathbf{x}_i\right) \mathbf{x} - b\right].$$

The summing is performed over for the support vectors only for which $I_i \neq 0$.

The problem solution considered is a simpler version of SVM. More sophisticated versions consider the cases of linear non-separation (when the two classes cannot be separated linearly) with weaker restrictions for inequalities, and the case of nonlinear classification with application of so called kernel trick.

Results of forecasting selected financial and economic processes

Below we show some results of forecasting economic and financial processes with the DSS proposed using linear regression models, nonlinear logistic regression, Bayesian networks and SVM. Table 1 shows results of forecasting inflation rate (on test sample) in Ukraine with GDP as predictor using autoregressive moving average and other models.

Table 1

Results of forecasting inflation rate

Model type	Model adequacy			Quality of one-step ahead forecasting			
	R^2	$\sum e^2(k)$	DW	MSE	MAE	$MAPE$	U
AR(1)	0,415	141,99	1,931	1,360	1,020	1,008	0,0067
AR(3)	0,317	135,14	1,992	1,360	1,020	1,011	0,0068
AR(7)	0,346	127,24	1,811	1,360	1,012	1,002	0,0067
AR(12)	0,435	97,80	1,941	1,337	1,020	1,013	0,0066
ARMA(1,1)	0,416	141,78	1,996	1,362	1,016	1,005	0,0067
AR(1) + M3	0,419	141,01	1,919	1,340	1,004	0,994	0,0066
AR(1)+GDP	0,419	141,05	1,916	1,335	1,004	0,993	0,0066
SVM	0,438	112,65	1,953	1,341	1,009	1,010	0,0067

Note: here AR is autoregression; $ARMA$ is autoregression with moving average; R^2 is determination coefficient; $e(k)=y(k) - \hat{y}(k)$ is model error; DW is Durbin-Watson statistic; MSE is mean squared forecasting error; MAE is mean absolute error; $MAPE$ is mean absolute percentage error; U is a Theil coefficient; $M3$ is money aggregate; GDP is gross domestic product.

These results show that all the models constructed for inflation rate provide high quality short-term forecasting in spite of the fact that the values of determination coefficient R^2 are rather far from ideal. The last model constructed that includes GDP in the right hand side as a predictor showed the best result with $MAPE = 0,993$.

Table 2

Results of volatility forecasting for Microsoft stocks

Model type	Quality of historical forecasting			Quality of forecasting on test sample		
	MAE	$MAPE$	U	MAE	$MAPE$	U
ARCH	1,2756	2387,7	0,389	1,754	3241,2	0,453
GARCH	0,549	24,73	0,115	0,691	33,35	0,138
E-GARCH	0,436	3,45	0,017	0,073	4,75	0,022
SVM	0,517	5,35	0,031	0,295	7,64	0,039

Note: here $ARCH$ is conditionally heteroskedastic autoregression; $GARCH$ is generalized conditionally heteroskedastic autoregression; $E-GARCH$ is exponential generalized conditionally heteroskedastic autoregression.

Thus, the best result of volatility forecasting for the Microsoft stocks has been reached with the exponential generalized conditionally heteroskedastic autoregression ($MAPE = 4,75\%$ on test sample). Unacceptable quality of forecasting showed ARCH model with simplest structure ($MAPE = 2387,7\%$ on learning sample). It also can be seen that the values of mean absolute error and Theil coefficient are in accordance with the mean absolute percentage error that proves correctness of the experiment.

The methodology developed has also been applied to forecasting direction of moving for the stock prices. In this case we have used linear regression model, nonlinear logistic regression with extra market indicators, classification trees, Bayesian network, and some combinations of the models mentioned. The forecasting results for the price evolution direction are given in the Tables 3 and 4 for the maximum and minimum stock price selected.

Table 3

Results of maximum price direction forecasting for the stocks selected

Model type	Direction forecasting quality
Linear regression with indicators	68,95%
Logistic regression with indicators	69,76%
Classification tree with indicators	68,95%
Combination of linear and logistic regression	74,19%
Combination of linear regression and classification tree	69,48%
Bayesian network	72,38%

Table 4

Results of minimum price direction forecasting for the stocks selected

Model type	Direction forecasting quality
Linear regression with indicators	73,79%
Logistic regression with indicators	66,13%
Classification tree with indicators	64,66%
Combination of linear and logistic regression	74,73%
Combination of linear regression and classification tree	74,62%
Bayesian network	72,65%

Thus, in both cases (forecasting direction of movement for minimum and maximum price value) the best result was achieved with the combination of linear and logistic regression: probability of correct direction forecasting is 74,7 % and 74,19 %, respectively. The role of the linear regression model was to provide the price forecasts, and logistic regression predicted direction of the price evolution. The quality of the forecasts achieved is quite acceptable for their use in the rules of trading. High quality results have also been achieved with the Bayesian networks constructed on the statistical data available with the probability of correct direction forecasting 72,65 % and 72,38 %.

Conclusions

The methodology for constructing of decision support system for mathematical modeling of economic and financial processes that is based on the system analysis principles: hierarchical system structure, taking into consideration of probabilistic and statistical uncertainties, generating of decision alternatives, and tracking of computational processes for all the stages of data processing.

The system proposed has a modular architecture that provides a possibility for the easy extension of its functional possibilities with new model parameter estimation methods, forecasting techniques, and alternative generating. High quality of the final result is achieved thanks to appropriate tracking of the computational processes for all data processing stages: preliminary data processing, model structure and parameter estimation, computing of short- and middle-term forecasts, as well as thanks to convenient for a user intermediate and final results representation. The system is based on different (ideologically different) techniques of modeling and forecasting what creates a good base for combination of various approaches to achieve the best results. The examples of the system application show that it can be used successfully for solving practical forecasting problems.

The DSS can be used for decision making support in various areas of human activities including strategy development for industrial enterprises, transportation and investment companies etc. Further extension of the system functions is planned with new forecasting techniques based on probabilistic technologies.

1. McNeil A.J. *Quantitative Risk Management* / A.J. McNeil, R. Frey, P. Embrechts. – Princeton (New Jersey): Princeton University Press, 2005. – 538 p. 2. *International Convergence of Capital Measurement and Capital Standards. A Revised Framework. Comprehensive Version.* – Basel Committee on Banking Supervision, Bank for International Settlements. – Basel, 2006. – 158 p. 3. Mays E. (Ed.) *Handbook of Credit Scoring* / E. (Ed.) Mays. – Chicago: Glenlake Publishing Company, Ltd., 2001. – 460 p. 4. Neil M. *Using Bayesian networks to model expected and unexpected operational losses* / Neil M., Fenton N.E., Tailor M. // *Risk Analysis*. – 2005. – P. 34-57. 5. Zgurovsky M.Z. *Method of constructing*

Bayesian networks based on scoring functions / M.Z. Zgurovsky, P.I. Bidyuk, O.M. Terentyev // Cybernetics and System Analysis, 2008.- Vol. 44.- No.2.- P. 219-224. 6. *Polovcev O.V. A System Approach to Modeling, Forecasting, and Management of Financial and Economic Processes / O.V. Polovcev, P.I. Bidyuk, L.O. Korshevnyuk. – Donetsk: Oriental Publishing House, 2009. – 286 p.* 7. *Hollsapple C.W. Decision support systems / C.W. Hollisapple, A.B. Winston. – Saint Paul: West Publishing Company, 1996. – 860 p.* 8. *Tsay R.S. Analysis of financial time series / R.S. Tsay. – Hoboken: Wiley & Sons, Inc., 2010. – 715 p.* 9. *Bidyuk P.I. Methods of Forecasting / P.I. Bidyuk, O.S. Menyailenko, O.V. Polovcev. – Lugansk: Alma Mater, 2008. – 608 p.* 10. *Burstein F. Handbook of Decision Support Systems / F. Burstein, C.W. Hollisapple. – Berlin: Springer-Verlag, 2008. – 908 p.* 11. *De Jong P. Generalized Linear Models for Insurance Data / P. De Jong, G.Z. Heller. – New York: Cambridge University Press, 2008. – 197 p.* 12. *Gilks W.R. Markov chain Monte Carlo in practice / W.R. Gilks, S. Richardson, D.J. Spiegelhalter. – New York: Chapman & Hall/CRC, 2000. – 486 p.* 13. *Jensen F.V. Bayesian Networks and Decision Graphs / F.V. Jensen, Th. D. Nielsen. – New York: Springer, 2007. – 457 p.* 14. *Smola A.J. A tutorial on support vector regression / A.J. Smola, B. Scholkopf // Statistics and computing, 2004.- Vol. 14.- P. 199–222.*

УДК 519.7

V. Lytvynenko

Kherson National Technical University,
Dept. of Informatics & Computer Science

HYBRID SWARM NEGATIVE SELECTION ALGORITHM FOR DNA-MICROARRAY DATA CLASSIFICATION

Ó Lytvynenko V., 2014

В роботі запропоновано метод класифікації. Він заснований на комбінованому алгоритмі негативної селекції, який був спочатку розроблений для задач бінарної класифікації. Точність розробленого алгоритму була перевірена експериментальним шляхом з використанням наборів даних мікрочіпів. Експерименти підтвердили, що напрямок змін, внесених в розроблений алгоритм підвищує точність у порівнянні з іншими алгоритмів класифікації.

Ключові слова: алгоритм вибору, класифікатор, мікрочіп даних, аналіз головних компонентів, імпульсне перетворення, відбір ознак.

In the paper, a classification method is proposed. It is based on Combined Swarm Negative Selection Algorithm, which was originally designed for binary classification problems. The accuracy of developed algorithm was tested in an experimental way with the use of microarray data sets. The experiments confirmed that direction of changes introduced in developed algorithm improves its accuracy in comparison to other classification algorithms.

Key words: Negative Selection Algorithm, Swarm Selection Algorithm, Classifier, DNA-Microarray Data, Principal Component Analysis, Wavelet transformation, Feature reduction, Feature selection.

1. Introduction

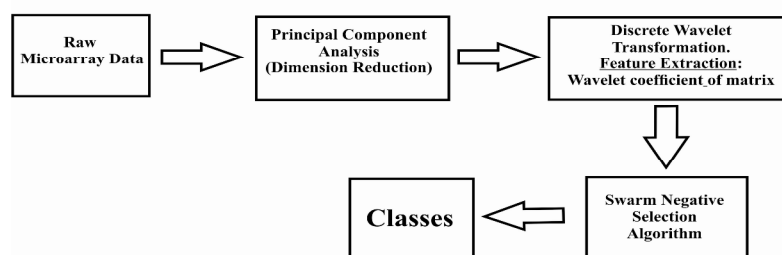
DNA microarray technology, introduced in 1995–1996, allows the measurement of thousands of gene expression values simultaneously, providing insight into the global gene expression patterns of cells (tissues) being studied [1,2,3]. Despite the need for further technological developments with microarray assays [4], the approach remains powerful for studying the myriad of transcription-related pathways involved in cellular growth, differentiation, and transformation in various organisms. In particular, the ability to measure thousands of gene expressions simultaneously using DNA microarrays has made it possible to investigate genome-wide objective approaches to molecular cancer classification[5]. Empirical

microarray data produce large datasets having expression levels of thousands of genes with a very few numbers (upto hundreds) of samples which leads to a problem of “curse of dimensionality”. Due to this high dimension the accuracy of the classifier decreases as it attains the risk of overfitting. As the microarray data contains thousands of genes, hence a large number of genes are not informative for classification because they are either irrelevant or redundant. Hence to derive a subset of informative or discriminative genes from the entire gene set is necessary and a challenging task in microarray data analysis. The purpose of gene selection or dimension reduction is to simplify the classifier by retaining small set of relevant genes and to improve the accuracy of the classifier. For this purpose, researchers have applied a number of test statistics or discriminant criteria to find genes that are differentially expressed between the investigated classes [15]. A typical DNA microarray data set in tumor tissue c classification studies consists of expression measurements on thousands of genes over a small number of known tumor tissue samples ($p \gg N$). However, many standard statistical methodologies for classification and prediction require more samples than predictors. For example, in regression, $N < p$ leads to an ill-posed problem because the ordinary least squares (OLS) solution is not unique. Another example is Fisher’s discriminant analysis, where the covariance matrix is singular when $N < p$ [5]. It is challenging to use gene expression data for cancer classification because of the following two special aspects of gene expression data. First, gene expression data are usually very high dimensional. The dimensionality ranges from several thousands to over ten thousands. Second, gene expression data sets usually contain relatively small numbers of samples, e.g., a few tens. If we treat this pattern recognition problem with supervised machine learning approaches, we need to deal with the shortage of training samples and high dimensional input features. Recent approaches to solve this problem include artificial neural networks [7], an evolutionary algorithm[8], nearest shrunken centroids [9], and a graphical method [10]. A number of recent publications report on the successful application of support vector machines (SVMs) to the classification of high-dimensional microarray data [11-13]. Therefore, high-dimensional microarray data present a major challenge for these classifiers. However, the algorithms of Artificial Immune System (AIS) have not been widely explored for cancer classification with microarray data. Yet there exist in literature only very few studies in which AIS were applied to microarray classification. Therefore, this study introduced an artificial immune system approach for cancer detection based on negative selection algorithm (NSA) and Particle Swarm optimization (PSA) named Hybrid Swarm Negative Algorithm (HSNA).

2. MATERIALS AND METHODS

In this study, the microarray data classification was performed in three stages: dimensionality reduction using the Principal Component analysis, Feature extraction using the discrete wavelet transform and classification using hybrid swarm negative algorithm (HSNA).

Fig. 1. Structure of the HCNA Classifier



2.1. Dataset

Microarray datasets take the form of expression data matrix where rows represent the genes and columns represent the samples. Each cell in this data matrix is a gene expression value which expresses the gene intensity in the corresponding sample. The expression data matrix will be finally dealt with in the form X_{ij} where; $0 < i \leq n_g, 0 < j \leq n_s$ and n_s, n_g are the total number of genes, total number of samples respectively as in figure 2. Each expression data matrix will be further divided into two matrices; training data matrix Y_{ik} and test data matrix Z_{ip} where k, p are the number of samples used in the training process,

test process respectively and $p + k = n_s$. The training data matrix will be used to train all the used classifiers and their performance will be evaluated using the test data matrix only [16].

$$x_{ij} = \begin{pmatrix} x_{11} & x_{12} & \mathbf{K} & x_{1n_s} \\ x_{21} & x_{22} & & x_{2n_s} \\ & & & \\ x_{n_s 1} & & & x_{n_s n_s} \end{pmatrix}$$

Fig. 2. Expression data matrix

In this section, the cancer gene expression data sets used for the study are described. These datasets are also summarized below.

ALL/AML Leukemia Dataset. The dataset consists of two distinctive acute leukemias, namely AML and ALL bone marrow samples with 7129 probes from 6817 human genes. The training dataset consists of 3B8 samples (27 ALL and 11 AML) and the test dataset consists of 34 samples (20 ALL and 14 AML).

Colon Dataset. The dataset consists of 62 samples from 2000 genes. The training dataset consists of 42 samples where (30 class1, 12 class2) and the test data set consists of 20 samples (10 class1, 10 class2).

Prostate cancer. Prostate cancer data contains training set of 52 prostate tumor samples and 50 nontumor (labeled as “Normal”) prostate samples with 12 600 genes. An independent set of test samples is also prepared, which is from a different experiment. The test set has 25 tumor and 9 normal samples.

2.2. Dimensionality reduction using the Principal Component analysis

Principal component analysis (PCA) is used to search new abstract orthogonal principal components (eigenvectors) which explain most of the data variation in a new coordinate system [17]. Classical PCA is based on the decomposition of a covariance/ correlation matrix by eigenvalue (spectral) decomposition (EVD) or by the decomposition of real data matrixes using SVD [18].

PCA is a multivariate procedure aimed at reducing the dimensionality of multivariate data while accounting for as much of the variation in the original data set as possible.

This technique is especially useful when the variables within the data set are highly correlated and when there is a higher than normal ratio of explanatory variables to the number of observation. Principal components seeks to transform the original variable to a new set of variables that are (1) linear combinations of the variables in the data set, (2) uncorrelated with each other, and (3) ordered according to the amount of variation of the original variables that they explain [17,19]

PCA is a well-known method of dimension reduction [20]. The basic idea of PCA is to reduce the dimensionality of a data set, while retaining as much as possible the variation present in the original predictor variables. This is achieved by transforming the p original variables $X = [x_1, x_2, \mathbf{K}, x_p]$ to a new set of K predictor variables, $T[t_1, t_2, \mathbf{K}, t_K]$, which are linear combinations of the original variables. In mathematical terms, PCA sequentially maximizes the variance of a linear combination of the original predictor variables,

$$u_k = \arg \max_{u'} \text{Var}(Xu) \quad (1)$$

subject to the constraint $u_i' S_x u_j = 0$, for all $1 \leq i \leq j$. The orthogonal constraint ensures that the linear combinations are uncorrelated, i.e. $\text{Cov}(Xu_i, Xu_j) = 0, i \neq j$. These linear combinations

$$t_i = Xu_i \quad (2)$$

are known as the principal components (PCs) [21]. Geometrically, these linear combinations represent the selection of a new coordinate system obtained by rotating the original system. The new axes represent the directions with maximum variability and are ordered in terms of the amount of variation of the original data they account for. The first PC accounts for as much of the variability as possible, and each succeeding component

accounts for as much of the remaining variability as possible. Computation of the principal components reduces to the solution of an eigenvalue-eigenvector problem. The projection vectors (or called the weighting vectors) u can be obtained by eigenvalue decomposition on the covariance matrix S_X ,

$$S_X u_i = I_i u_i \quad (3)$$

where I_i is the i -th eigenvalue in the descending order for $i=1, \mathbf{K}, K$, and u_i is the corresponding eigenvector. The eigenvalue I_i measures the variance of the i -th PC and the eigenvector u_i provides the weights (loadings) for the linear transformation (projection). The maximum number of components K is determined by the number of nonzero eigenvalues, which is the rank of S_X , and $K \leq \min(n, p)$. The computational cost of PCA, determined by the number of original predictor variables p and the number of samples n , is in the order of $\min(np^2 + p^3, pn^2 + n^3)$. In other words, the cost is $O(pn^2 + n^3)$ when $p > n$ [22].

2.3 Discrete wavelet transform-feature extraction

Suppose that the vector \bar{X}_1 has a sequence consisting of the 2^n points, for some integer $n > 0$. This sequence can be identified with the next function in the space V^n of piecewise constant functions at equidistant intervals of length $1/2^n$:

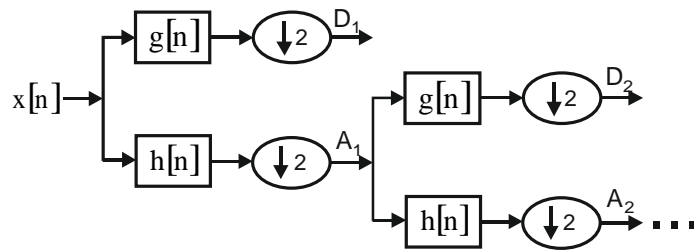
$$f(t) = x_1 f_{n,0}(t) + \dots + x_{2^n} f_{n,2^n-1}(t) \quad (4)$$

where $f(t)$ - scaling functions of space V^n . The first step in calculating the wavelet decomposition of the sequence $\{x_1, x_2, \dots, x_{2^n}\}$ is the decomposition of $f(t)$ on the alternative basis of the space V^n , which constitute half of the wavelets $y(t)$:

$$f(t) = A_{n-1,0} f_{n-1,0}(t) + \dots + A_{n-1,2^{n-1}-1} f_{n-1,2^{n-1}-1}(t) + D_{n-1,0} y_{n-1,0}(t) + \dots + D_{n-1,2^{n-1}-1} y_{n-1,2^{n-1}-1}(t) \quad (5)$$

where A - approximation coefficients, defining coarse low-frequency component of the original signal, D - detail coefficients, defining the high-frequency component of the original signal. The next step of the conversion process is the use of the same basic conversion the members of (2), containing the approximation coefficients. Detail coefficients at the same time remain unchanged. Block diagram of the wavelet decomposition is presented in Figure 3.:

Fig. 3. Structural diagram of a discrete wavelet decomposition of the signal; $g[n]$ – high frequency transmit filter, $h[n]$ – low frequency filter transmitting



The Data used in this research were analyzed into the details D1-D2 and one final approximation, A2. Our previous studies [23] have shown that the smoothing feature of the Daubechies wavelet of order 13 (db13) made it more suitable to detect changes of the microarray data. Hence, in our research, we used the db13 to compute the wavelet coefficients of the microarray data. A detailed description of the method of obtaining the approximation coefficients described by us in [32].

The computed discrete wavelet coefficients provide a compact representation. In order to further decrease the dimensionality of the extracted feature vector, statistics over the set of the wavelet coefficients are used. The following statistical features were used to represent the time-frequency distribution of the microarray data:

- Maximum of the wavelet coefficients in each subband
- Minimum of the wavelet coefficients in each subband
- Mean of the wavelet coefficients in each sub-band
- Standard deviation of the wavelet coefficients in each sub-band.

2.4 Artificial immune algorithms

In the 1990s, Artificial Immune System (AIS) emerged as a new computational research field inspired by simulation of biological behavior of Natural Immune System (NIS). The NIS is a very complex biological network with rapid and effective mechanisms for defending the body against a specific foreign body material or pathogenic material called antigen .

The Artificial Immune Systems, as defined by de Castro and Timmis [24] are: “Adaptive systems inspired by theoretical immunology and observed immune functions, principles and models, which are applied to problem solving”. However AIS are one of many types of algorithms inspired by biological systems, such as neural networks, evolutionary algorithms and swarm intelligence. There are many different types of algorithms within AIS and research to date has focused primarily on the theories of immune networks, clonal selection and negative selection. These theories have been abstracted into various algorithms and applied to a wide variety of application areas such as anomaly detection, pattern recognition, learning and robotics [25].

Negative selection algorithm. The negative selection of T-cells is responsible for eliminating the T-cells whose receptors are capable of binding with self-peptides presented by self-MHC molecules. This process guarantees that the T-cells that leave the thymus do not recognize any self-cell or molecule. Forrest et al. [26] proposed a change detection algorithm inspired by the negative selection of T-cells within the thymus. This procedure was named as negative selection algorithm and was originally applied in computational security. A single type of immune cell was modelled: T-cells were represented as bit strings of length L . The negative selection algorithm of Forrest and collaborators is simple [26]. Given a set of self-peptides, named self-set S , the T-cell receptors will have to be tested for their capability of binding the self-peptides. If a T-cell recognizes a self-peptide – it is discarded, else it is selected as an immune-competent cell and enters the available repertoire A .

The idea of negative selection algorithm is to generate a set of detectors in a complementary set of N and then to use these detectors for binary classification as “Self” or “Non-Self”. Formally, the negative selection algorithm can be represented as [27-28]:

$$NegAlg = (\Sigma^L, L, \mathbf{S}, \mathbf{N}, r, n, s, pr) \quad (6)$$

where Σ^L denotes shape-space; L is receptor length; \mathbf{S} is “Self” detector set; \mathbf{N} is “Non-Self” detector set; r denotes cross-reactive threshold; n is total number of appointed detectors; s is detector set size; pr denotes rule matching rows in adjacent positions.

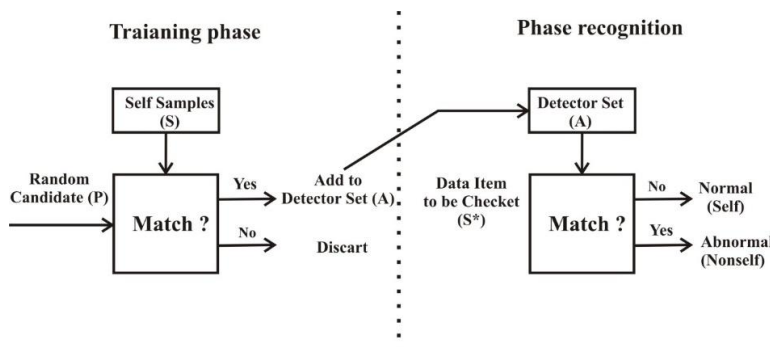


Fig. 4. Negative selection algorithm [29]

The negative selection algorithm can be summarized as follows:

- *Initialization:* randomly generate strings and place them in a set \mathbf{P} of immature T-cells, assuming all the molecules (receptors and self-peptides) are represented as binary strings of the same length L .
- *Affinity evaluation:* determine the affinity of all T-cells in V with all elements of the self set \mathbf{S} .
- *Generation of the available repertoire:* if the affinity of an immature T-cell with at least one self-peptide is greater than or equal to a given cross reactive threshold, then the T-cell recognizes this self-peptide and has to be eliminated (negative selection); else the T-cell is introduced into the available repertoire \mathbf{A} .

The process of generating the available repertoire in the negative selection algorithm was termed learning phase. The algorithm is also composed of a monitoring phase. In the monitoring phase, a set \mathbf{S}^* of protected strings is matched against the elements of the available repertoire \mathbf{A} . The set \mathbf{S}^* might be the own set \mathbf{S} , a completely new set, or composed of elements of \mathbf{S} . If recognition occurs, then a non-self pattern (string) is detected.

It is well known, that the algorithm of negative selection (NS) has the some restrictions and limitations [29]. When it is not appropriate, for example, the number of self samples is small and sparse.

Some limitations of the binary-string representation in NS algorithms are as follows:

- binary matching rules are not able to capture the semantics of some complex self/non-self spaces,
- it is not easy to extract meaningful domain knowledge,
- in some cases a large number of detectors are needed to guarantee better coverage (detection rate),
- it is difficult to integrate the NS algorithm with other immune algorithms,
- the crisp boundary of “self” and “non-self” may be very hard to define.

In real-valued representation the detectors are represented by hyper-shapes in n -dimensional space. The algorithms use geometrical spaces and use heuristics to distribute detectors in the non-self space.

Some limitations of the real-valued representation in NS algorithms are:

- the issue of holes in some geometrical shapes, and may need multi-shaped detectors,
- curse of dimensionality,
- the estimation of coverage,
- the selection of distance measure.

During our experiments it has been established that generation of set of detectors in at training phase occurs casually owing to what it is in advance impossible to define is minimum necessary quantity of detectors which will provide the maximum quality of recognition. The increase in quantity of detectors conducts to delay of a phase of recognition, and its reduction – to deterioration of work of algorithm since the probability of formation of the “cavities” which are areas in space of “Non-self” which are not distinguished by any of detectors increases. Thus, a problem of the given research is working out of an advanced method of generation of the detectors, capable to adaptive selection of their options, quantity and an arrangement.

Particle swarm optimization (PSO) is a computational method that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality. PSO optimizes a problem by having a population of candidate solutions, here dubbed particles, and moving these particles around in the search-space according to simple mathematical formulae over the particle's position and velocity. Each particle's movement is influenced by its local best known position but, is also guided toward the best known positions in the search-space, which are updated as better positions are found by other particles. This is expected to move the swarm toward the best solutions.

PSO is originally attributed to Kennedy, Eberhart and Shi [24,25] and was first intended for simulating social behaviour,[26] as a stylized representation of the movement of organisms in a bird flock or fish school. The algorithm was simplified and it was observed to be performing optimization. The book by Kennedy and Eberhart[33] describes many philosophical aspects of PSO and swarm intelligence. An extensive survey of PSO applications is made by Poli [34,35].

PSO is a metaheuristic as it makes few or no assumptions about the problem being optimized and can search very large spaces of candidate solutions. However, metaheuristics such as PSO do not guarantee

an optimal solution is ever found. More specifically, PSO does not use the gradient of the problem being optimized, which means PSO does not require that the optimization problem be differentiable as is required by classic optimization methods such as gradient descent and quasi-newton methods. PSO can therefore also be used on optimization problems that are partially irregular, noisy, change over time, etc. The PSO algorithm is population-based: a set of potential solutions evolves to approach a convenient solution (or set of solutions) for a problem. Being an optimization method, the aim is finding the global optimum of a real-valued function (fitness function) defined in a given space (search space).

The social metaphor that led to this algorithm can be summarized as follows: the individuals that are part of a society hold an opinion that is part of a “belief space” (the search space) shared by every possible individual. Individuals may modify this “opinion state” based on three factors:

- The knowledge of the environment (its fitness value)
- The individual's previous history of states (its memory)
- The previous history of states of the individual's neighborhood

An individual's neighborhood may be defined in several ways, configuring somehow the “social network” of the individual. Several neighborhood topologies exist (full, ring, star, etc.) depending on whether an individual interacts with all, some, or only one of the rest of the population.

Following certain rules of interaction, the individuals in the population adapt their scheme of belief to the ones that are more successful among their social network. Over the time, a culture arises, in which the individuals hold opinions that are closely related. The particle swarm optimization is based on a set of individuals originally randomly arranged and homogeneous. Therefore we call it particles, which move in the hyperspace of research and are each a potential solution. Each particle has a memory about his best seen as the ability to communicate with the particles forming around it. From this information, the particle will follow a trend made, from one side, willingness to return to its optimal solution, and from the other side, his mimicry in relation to the solutions found in its vicinity. From local optima and empirical, all particles will normally converge to the global optimum solution of the addressed problem.

The process of finding the particles is based on two rules:

- 1) Each particle has a memory that can store the best point by which it has already passed and it tends to return to this point;
- 2) Each particle is informed of the best known point in its neighborhood and it will tend to move towards this point.

Each particle moves according to a compromise between the three following trends:

- a) Repeat its previous motion;
- b) Move towards its best previous position;
- c) Move towards the best position (past) of its group of informants.

Each agent tries to modify its position based on the following information:

- a) Current positions (x, y)
- b) Current velocities (v_x, v_y)
- c) Distance between the current position and **pbest**
- d) Distance between the current position and **gbest**

Thus, the velocity of the particle i is updated using the following equation:

$$v_i^{k+1} = wv_i^k + c_1rand_1(pbest - S_i^k) + c_2rand_2(gbest - S_i^k) \quad (7)$$

Where w the inertia weight, c_1 and c_2 are the acceleration constants ($c_1 + c_2 \leq 4$), $rand_1$ and $rand_2$ are random numbers in the interval $[0, 1]$.

Then the position of the particle S_i^k is modified from the current position and a new speed is calculated v_i^{k+1} :

$$S_i^{k+1} = S_i^k + v_i^{k+1} \quad (8)$$

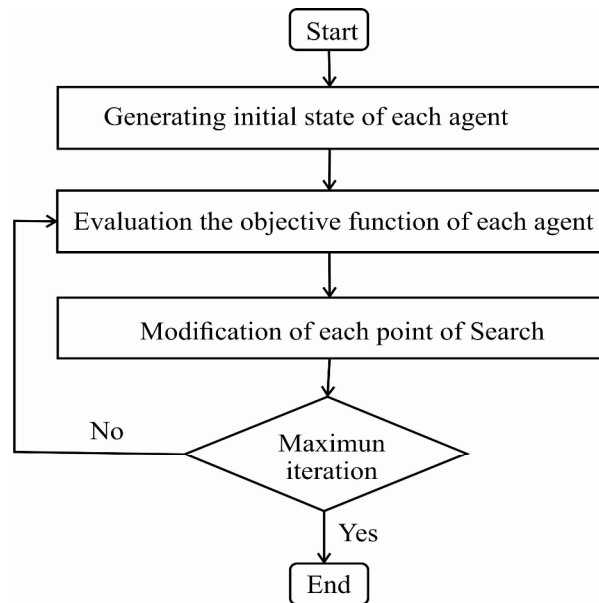
The weight w is given by the following equation:

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{iter_{\max}} iter \quad (9)$$

$w = [0.4 - 0.9]$ during the search procedure gives better results.

The right choice of parameters will allow the rapid convergence and minimizes the computation time. The first rule stipulates that c_1 must have an absolute value less than 1 in practice, this coefficient should be neither too small on PSO recommended that it be equalized to 0.7 or 0.8. The parameter c_2 should not be too large, a value of about 1.5 to 1.7 being regarded as effective in the majority of cases. The pairs of values (0.7 1.47) and (0.8 1.62) are indeed correct. The following figure shows the general flowchart of PSO. The steps involved in the optimization algorithmic of the particle swarm are as follows:

Fig. 5. General flowchart of the PSO algorithm [36]



Step 1: Select several parameters of PSO

Step 2: Initialize a population of particles with random position and velocities in the problem search space;

Step 3: Evaluate the ability of optimization for a desired personal touch to each particle;

Step 4: For each individual particle, compare the value of the particle with its ability *pbest*, If the current value is better than the evaluated *pbest* for the agent i ;

Step 5: Identify the particle that has the best value of its fitness function which will be identified as *gbest*;

Step 6: Calculate the new speed and position of particles using equations 15 and 16.

Step 7: Repeat steps 3-6 until the stopping criterion is met.

The generalized flowchart of the PSO algorithm is presented on a figure 5.

Combined particle swarm optimization and negative selection algorithm. The classifier presented in this paper is based on the hybridization process of negative selection with clonal selection, and was designed to solve problems of classification to many classes. Concept of classification is used in terms of supervised learning, which allows categorizing objects into known groups using training set prepared beforehand. The main task of every classifier based on supervised learning is to create an internal representation of classes (in the form of a function, set of rules or any other). It acquires it during training. When the training is completed the classifier is ready to produce an answer to any (known or unknown) pattern given subsequently.

In this study the efficiency of immune classifiers is researched, when as the classifier, in general, is a function that for attributes vector of object shall decide to which class it belongs [27]:

$$F : \mathfrak{R}^n \rightarrow Y . \quad (10)$$

The function F represents the space of sign vectors in the space of the class labels Y . In the case of two classes $Y = \{0,1\}$, '1' corresponding case of the detection event, '0' - the event is not detected. We consider the variant of training with a teacher (supervised learning), when the classifier training available to us a set of vectors $\{x\}$ for which is known their valid membership in one of the classes.

Negative selection algorithms and swarm optimization particles, described in detail in [27-29]. In this work we give an algorithm to modify the learning phase in the negative selection algorithm. The main idea is a modification of the learning phase in the negative selection algorithm. Generalized flowchart of this algorithm is shown in Figure 6. Each particle D_i has a velocity vector and the coordinates, and is the immune detector. Cross-reactive threshold of this detector is used to separate the population of antigens into two subsets – “Self” and “Non-self”. The learning process is to optimize of this algorithm by using PSO detector location, resizing, changing their cross-reactive thresholds.

For evaluation criteria use values of affinity to antigens, and the detectors together. The flowchart of the learning phase the proposed algorithm is shown in Figure 9.

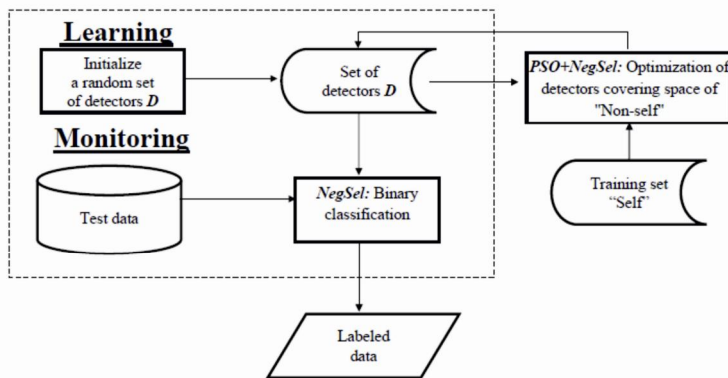


Fig. 6. Flowchart learning phase, the proposed algorithm

Step by step description of the algorithm that implements the proposed method:

Step 1. At this step, initializes the population of particles (detectors) with randomly chosen positions and random velocity vectors. The size of this set remains constant. In the learning process only change the properties of its elements.

Step 2. A loop is created for all elements the learning set “Self” (set of antigens). Further actions are performed sequentially for each antigen.

Step 3. A loop is created for each detector from the population.

Step 4. To calculate the affinity values of the relationship “antigen-antibody” is used the ratio:

$$f_{Ab-Ag} = \frac{k_e}{e} + D_{E(Ab-Ag)} \rightarrow \min \quad (11)$$

where r is the cross-reactive antibodies (detector) threshold, k_r - coefficient of importance of cross-reactive threshold (parameter settings). Parameter k_r is very important study parameter. He manages the robustness of the resulting solution. Increasing this parameter causes the immune network to maintain a larger radius detectors, what gives a coarser, but more sustainable solutions. However, excessive increase of k_r affects negative to the accuracy of solutions. The figure 7 demonstrates the impact parameter k_r on the way of detector generation.

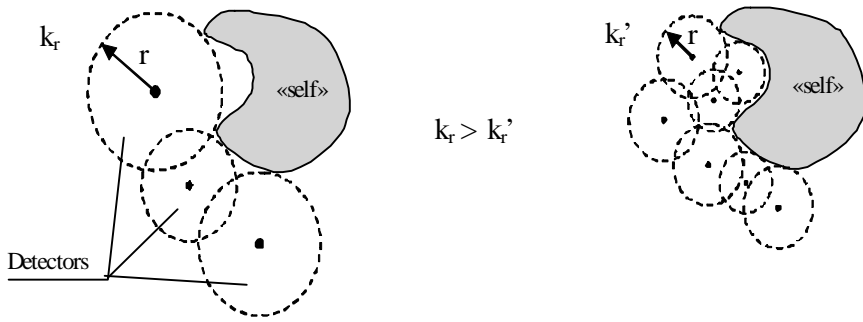


Fig. 7. The influence of parameter value k_r on the way of detector generation

Steps 5,6. Based on the calculated values of the affinity, the best particles are calculated in the Step 7. According to estimates computed values is refreshed values of the velocity vectors and the coordinates of the particles in the swarm. In this case, the ratio used, the algorithm used swarms of particles.

To calculate the values of affinity relationship “antibody-antibody” proposed the following formula:

$$f_{Ab-Ab} = -\frac{D_{E(Ab_1-Ab_2)} - (e_{Ab_1} + e_{Ab_2})}{2 \cdot \min(e_{Ab_1}, e_{Ab_2})}, \quad (12)$$

This possible interpretation of such values $f_{Ab-Ab} : \leq 0$ - recognizing hyper sphere detectors do not overlapped. This option does not require compression, because the antibodies do not recognize each other; $(0, 1)$ - hyper sphere overlapped by shells, namely the value is the degree of overlap. This compression is dependent on the value of the parameter of threshold compression S_s , which is a parameter of study; ≥ 1 - hyper sphere smaller radius (r) is completely inside hyper sphere larger radius. In this case, compression is definitely needed, because there is redundancy recognizing elements.

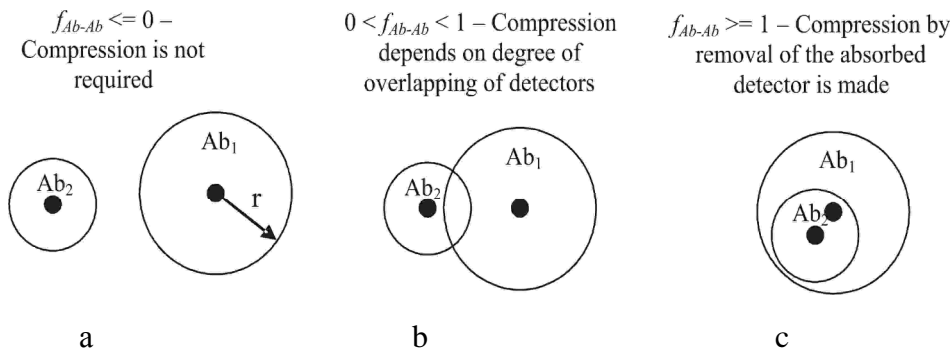


Fig. 8. Different mutual arrangement of the recognizing hyper sphere of detectors depending on the value f_{Ab-Ab}

Step 10. Based on the values calculated estimates relations "detector-detector" is corrected values of the velocity vectors and the coordinates of each particle in the population.

Step 11. Found the best particle returns to a population of memory cells, replacing a disposal corresponding in order to preserve a permanent composition of the population.

Steps 12, 13, 14. This process repeated as long until certain conditions are fulfilled stop (for example, the quality of classification). At the same time output of this phase is learning the population of memory cells (detectors).

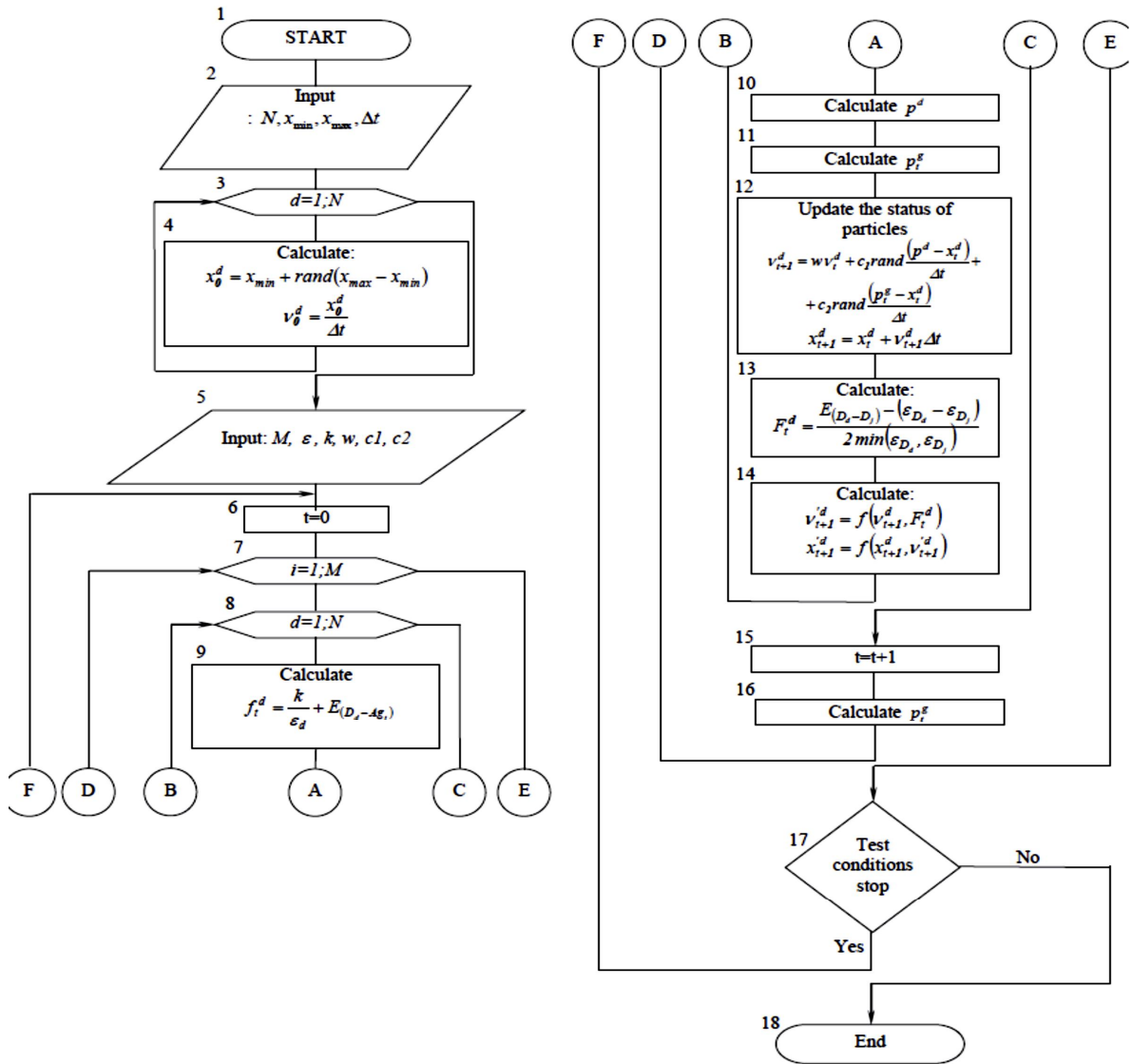


Fig. 9. Flowchart of the learning phase the hybrid algorithm developed

3. RESULTS AND DISCUSSION

It is common practice in machine learning and data mining to perform k-fold cross-validation to assess the performance of a classification algorithm. K-fold cross validation is used among the researchers, to evaluate the behavior of the algorithm in the bias associated with the random sampling of the training data. In k-fold cross-validation, the data is partitioned into k subsets of approximately equal size. Training and testing the algorithm is performed k times. Each time, one of the k subsets is used as the test set and the other k-1 subsets are put together to form a training set. Thus, k different test results exist for the algorithm. However, these k results are used to estimate performance measures for the classification system.

The common performance measures used in medical diagnosis tasks are accuracy, sensitivity and specificity. Accuracy measured the ability of the classifier to produce accurate diagnosis. The measure of the ability of the model to identify the occurrence of a target class accurately is determined by sensitivity. Specificity is determined the measure of the ability of the algorithm to separate the target class. The classification accuracies for the datasets are calculated as in Eq. 15:

$$Accuracy(Z) = \frac{\sum_{i=1}^{|Z|} Assess(z_i)}{|Z|} \quad (13)$$

while

$$Assess(z) = \begin{cases} 1, & \text{if } classify(z) = z.c \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

where z denotes the patterns in testing set to be classified, $z.c$ is the class of pattern z , $classify(z)$ returns the classification of z by classification algorithm.

For sensitivity and specificity analysis, the following equations can be used:

$$Sensitivity = \frac{TP}{TP + FN} \quad (15)$$

$$Specificity = \frac{TN}{TN + FP} \quad (16)$$

where TP, TN, FP i FN denote respectively true positive, true negative, false positive and false negative classification.

In order to compare the efficiency of the proposed method in predicting the class of the cancer microarray data we have used three standard datasets such as All/AML Leukemia, Colon Dataset, Prostate cancer. All the datasets is binary class datasets. The feature selection process proposed in this paper has two steps. First the microarray data is decomposed by factor analysis optimally choose the discriminate feature set then using Discrete wavelet transform into level 4 using db13 wavelet to get the approximation coefficients as the extracted feature set. The performance of the proposed feature extraction method is analyzed with the well studied neural network classifiers such as MLP and RBFNN. The leave one out cross validation (LOOCV) test is conducted by combining all the training and test samples for both the classifiers with all the three datasets and the results are listed in Table 1. For this Data the performance of HSNA is comparable to RBFN and MLP.

Table 1

Comparison study of classification accuracy, sensitivity and specificity of HCNA with MLP and RBFN classifiers

Dataset	Method	Classification Accuracy	Sensitivity	Specificity
ALL/AML Leukemia	MLN	91.3%	98.21	97.01
	RBFN	98.4%	98.55	97.25
	HSNA	100%	99.16	99.20
Colon Dataset	MLN	94.5%	97.44	96.25
	RBFN	97.6%	98.10	97.10
	HSNA	99.7%	98.80	99.45
Prostate cancer	MLN	98.8%	98.03	97.21
	RBFN	97.2%	98.80	98.00
	HSNA	100%	99.10	99.60

The performance of the proposed method is also compared with those obtained by the recently reported methods and the results are listed in Table 2-4. The existing methods also used the cross validation test on the datasets. From Tables 2-4 it reveals that our method is equivalent to the counterparts with the advantage of reduced computational load. Weka [31]. Table 5 shows the decomposition stages upto 4th level by using db13 in discrete wavelet transform.

Table 2

Comparison study of accuracy of Colon Dataset

Methods	Classification accuracy
Bayes Network	85.3 %
Naive Bayes classifier	60.1 %
Multinomial logistic regression model	74.2 %
Support Vector Classifier	94.3 %
Class for doing classification using regression methods	91.8 %
Simple Decision Table Majority Classifier	89.3 %

Table 2 continued

Methods	Classification accuracy
1R classifier	73.9 %
C4.5 decision tree	94.5 %
Forest of Random Trees	97.4 %
Factor Analysis + Wavelet + HCNA	99.7 %
Factor Analysis + Wavelet + HSNA	100 %

Table 3

Comparison study of accuracy of ALL/AML Leukemia dataset

Methods	Classification accuracy
Bayes Network	87.6 %
Naive Bayes classifier	64.2 %
Multinomial logistic regression model	77.4 %
Support Vector Classifier	91.3 %
Class for doing classification using regression methods	97.6 %
Simple Decision Table Majority Classifier.	92.5 %
1R classifier	70.9 %
C4.5 decision tree	96.5 %
Forest of Random Trees	97.6 %
Factor Analysis + Wavelet + HCNA	100 %
Factor Analysis + Wavelet + HSNA	100 %

Table 4

Comparison study of accuracy of Prostate cancer

Methods	Classification accuracy
Bayes Network	91.7%
Naive Bayes classifier	69.2 %
Multinomial logistic regression model	80.5%
Support Vector Classifier	99.4%
Class for doing classification using regression methods	91.7%
Simple Decision Table Majority Classifier.	89.4%
1R classifier	69.2%
C4.5 decision tree	97.1%
Forest of Random Trees	98.8%
Factor Analysis + Wavelet + HCNA	100%
Factor Analysis + Wavelet + HSNA	100%

Table 5

Reduction details of the dataset

Dataset	Original Dimension	Factor Analysis	DWT Db 13 Level 4
Colon	62×2000	62×700	62×180
ALL/AML Leukemia	72×7129	72×700	72×180
Prostate cancer	136×12600	136×700	136×180

4. Conclusion

In this paper we have presented a hybrid feature extraction method using the Factor analysis in conjunction with wavelet transform to effectively select the discriminative genes on microarray data. A

simple HSNA based classifier has also been introduced to classify the microarray samples efficiently. The comparison results elucidated that the proposed approach is an efficient method which performs better than the existing methods. Besides it has reduced computational complexity.

1. Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., Brown, E.L., 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* 14, 1675–1680.; 2. Schena, M., Shalon, D., Davis, R.W., Brown, P.O., 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467–470.; 3. Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P.O., Davis, R.W., 1996. Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. USA* 93, 10614–10619. 4. Nguyen, D.V., Arpat, A.B., Wang, N., Carroll, R.J., 2002d. DNA microarray experiments: biological and technological aspects. *Biometrics* 58, 701–717.; 5. D.V. Nguyen, D.M. Rocke On partial least squares dimension reduction for microarray-based classification: a simulation study / *Computational Statistics & Data Analysis* 46 (2004) 407 – 425.; 6 Jayakishan M., Barik R. C., Ranjani P. M., S. K. Pradhan C.F. G. Dash Analysis and Wavelet Transform Method for Tumor Classification Using Gene Expression Data/ *I.J. Information Technology and Computer Science*, 2012, 9, 73-79.; 7. J. Khan, J. S. Wei, M. Ringner, L. H. Saal and M. Ladanyi et al., Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks, *Nature Medicine* 7 (2001) 673–679.; 8. J. M. Deutsch, Evolutionary algorithms for finding optimal gene sets in microarray prediction, *Bioinformatics* 19 (2003) 45–52.; 9. R. Tibshirani, T. Hastie, B. Narashiman and G. Chu, Diagnosis of multiple cancer types by shrunken centroids of gene expression, in *Proc. Natl. Acad. Sci. USA*, Vol. 99 (2002), pp. 6567–6572.; 10. E. Bura and R. M. Pfeiffer, Graphical methods for class prediction using dimension reduction techniques on DNA microarray data, *Bioinformatics* 19 (2003) 1252–1258.; 11. Ramaswamy S., Tamayo P., Rifkin R., Mukherjee S., Yeang C.H., Angelo M. Ladd C., Reich M., Latulippe E., Mesirov J.P., Poggio T., Gerald W., Loda M., Lander E.S., Golub T.R., Multiclass cancer diagnosis using tumor gene expression signatures, *Proc. Natl. Acad. Sci. USA*. 98(26):15149-15154, (2001).; 12. Brown M.P.S., Grundy W., Lin D., Cristianini N., Sugnet C., Furey T., Ares M., Jr., Haussler D., Knowledge-based analysis of microarray gene expression data by using support vector machines, *Proc. Natl. Acad. Sci. USA*. 97(1):263-267, (2000). ; 13. Mukherjee S., Classifying microarray data using support vector machines, in Berrar D., Dubitzky W., Granzow M. (eds.), *A Practical Approach to Microarray Data Analysis*, Kluwer Academic Publishers, Boston, pp. 166-185, (2002).; 14. Berrar, D.P., Downes, C.S., Dubitzky, W.: A Probabilistic Neural Network for Gene Selection and Classification of Microarray Data. ;In *IC-AI(2003)*342-352. ; 15. Meher J., Barik R. C., Panigrahi M. R., Pradhan S.K., Dash G. Cascaded Factor Analysis and Wavelet Transform Method for Tumor Classification Using Gene Expression Data / *MECS I.J. Information Technology and Computer Science*, 2012, 9, 73-79; 16. D. A. Salem, R.A. A. Seoud, H. A. Ali MGS-CM: A Multiple Scoring Gene Selection Technique for Cancer Classification using Microarrays /*International Journal of Computer Applications* (0975 – 8887) Volume 36– No.6, December 2011 p. 30 – 37.; 17. Dong Hoon Lim Principal Component Analysis using Singular Value Decomposition of Microarray Data/ *World Academy of Science, Engineering and Technology* 81, 2013, p.56-58.; 18. Wall M.E., Dyck P.A., Brettin T.S.(2001). SVDMAN – singular value decomposition analysis of microarray data. *Bioinformatics* 17:566-68.; 19. Everitt, B. and T Hothorn. (2011). *An Introduction to Applied Multivariate Analysis with R (Use R!)*. Springer, New York, NY.; 20. Jolliffe, I.T. (1986). *Principal Component Analysis*. Springer, New York.; 21. Massey, W.F. (1965) Principal components regression in exploratory statistical research. *Journal of American Statistical Association*, 60, 234-246. ; 22. Jian J. Dai, Linh Lieu, David Rocke Dimension Reduction for Classification with Gene Expression Microarray Data/ *Stat Appl Genet Mol Biol*. 2006;5:Article6. Epub 2006 Feb 24. 23. Babichev S.A., Babenko N.I., Dydik A.A. Lytvynenko V.I., Fefelov A.A., SV Shkurdoda Filtration chromatogram with pomoshchju veyvelet-analysis with Using Criteria entropy (in Russian)/ *Regional mizhvuzovskyy*

collection of scientific papers Issue 6 (71) Dnipropetrovsk 2010 C-17-3224. ; 24. Kennedy, J., Eberhart, R. (1995). "Particle Swarm Optimization". *Proceedings of IEEE International Conference on Neural Networks IV*. pp. 1942–1948. doi:10.1109/ICNN.1995.488968.; 25. Shi, Y.; Eberhart, R.C. (1998). "A modified particle swarm optimizer". *Proceedings of IEEE International Conference on Evolutionary Computation*. pp. 69–73.; 26. Forrest, S., A.S. Perelson, L. Allen and R. Cherukuri, 1994. *Self-nonsel self discrimination in a computer*. *Proceedings of the Symposium IEEE Computer Society on Research in Security and Privacy, May 16-18, IEEE Xplore Press, Oakland, CA., USA., pp: 202-212. DOI: 10.1109/RISP. 1994.296580*; 27. Lytvynenko V. I.: *Comparative experimental study of a modified negative selection algorithm and clonal selection algorithm negative for solving classification (in Russian)*, *Vestnik Kherson National Technical University*, nr 4(33), 2008, pp. 7-14. ;28. Lytvynenko V. I.: *Immune classifier for solving binary classification - Theoretical Aspects (in Russian)*, *System technologies*, nr 1(42), 2006, pp. 32-47. ; 29. D. Dasgupta. *Advances in Artificial Immune Systems. IEEE Computational Intelligence Magazine*. November, 2006. - P.40-49.; 30. Kennedy, J. "The particle swarm: social adaptation of knowledge". *Proceedings of IEEE International Conference on Evolutionary Computation*. (1997). pp. 303–308.; 31. <http://www.cs.waikato.ac.nz/ml/weka/>; Lytvynenko V.I. *Application of Clonal Negative Algorithm to Cancer Classification with DNA-Microarray Data// System technologies*. – №6 (83). – Dnepropetrovsk, 2012. – P. 72 – 90. 32. Kennedy, J., Eberhart, R.C. (2001). *Swarm Intelligence*. Morgan Kaufmann. ISBN 1-55860-595-9. ; 33. Poli, R. (2007). "An analysis of publications on particle swarm optimization applications". *Technical Report CSM-469 (Department of Computer Science, University of Essex, UK)*.; 34. Poli, R. (2008). "Analysis of the publications on the applications of particle swarm optimisation". *Journal of Artificial Evolution and Applications* 2008: 1–10. doi:10.1155/2008/685175.; 35. Mekhanet, Med.; Mokrani, L.; Lahdeb, Med. *Comparison between Three Metaheuristics Applied to Robust Power System Stabilizer Design// Acta Electrotehnica* . 2012, Vol. 53 Issue 1, p41-49. 9p.

¹А. Ковальчук, ²І. Цмоць, ²М. Ступень
Національний університет “Львівська політехніка”,
¹кафедра інформаційних технологій видавничих систем,
²кафедра автоматизованих систем управління

КУБІЧНІ І ЛІНІЙНІ ФРАКТАЛИ З ЕЛЕМЕНТАМИ АЛГОРИТМУ RSA В ШИФРУВАННІ І ДЕШИФРУВАННІ ЗОБРАЖЕНЬ

© Ковальчук А., Цмоць І., Ступень М., 2014

Запропоновано застосування кубічних фрактальних перетворень до шифрування і дешифрування зображень у градаціях сірого кольору з використанням елементів алгоритму RSA.

Ключові слова: шифрування, дешифрування, фрактальний алгоритм, зображення

Proposed application of cubic fractal changes to the encryption and decryption, grayscale color using elements of the RSA algorithm.

Key words: encryption, decryption, fractal algorithm, image

Вступ

Як і у випадку традиційного шифрування, схема шифрування з відкритим ключем уразлива з погляду аналізу з перебором всіх ключів. І контрзаходи тут ті самі – використання довгих ключів. Однак у цьому випадку є й інші варіанти захисту. Криптосистеми з відкритим ключем залежать від деякої оборотної математичної функції зі спеціальними властивостями. Складність обчислення такого роду функцій може залежати не лінійно від числа бітів у ключі, а зростати швидше. Тому довжина ключа повинна бути достатньо великою для того, щоб аналіз із перебором всіх ключів став практично нездійсненним, але достатньо малою для того, щоб на практиці можна було використати операції шифрування й дешифрування. Пропоновані для використання на практиці довжини ключів, звичайно ж, забезпечують практичну неефективність аналізу з перебором всіх ключів, але при цьому виявляються занадто повільними для того, щоб відповідні алгоритми можна було рекомендувати для універсального застосування. Тому, як уже згадувалося вище, шифрування з відкритим ключем у цей час обмежується областями керування ключами й додатками цифрового підпису.

Іншою формою атаки є спроба знайти спосіб обчислення особистого ключа за відомим відкритим ключем. Сьогодні немає математичного доказу неможливості такої форми атаки для жодного алгоритму шифрування з відкритим ключем. Із цього погляду будь-який конкретний алгоритм цього типу, зокрема поширений алгоритм RSA, виявляється таким, що не викликає довіри. А історія криптографії показує, що проблема, яка здається нерозв'язною, може виявитися цілком розв'язною, якщо подивитися на неї з якогось, зовсім нового погляду. Це стосується і зображень.

Важливою характеристикою зображення є наявність в зображенні контурів. Задача виділення контуру вимагає використання операцій над сусідніми елементами, які є чутливими до змін і пригашають області постійних рівнів яскравості, тобто контури – це ті області, де виникають зміни, стаючи світлими, тоді як інші частини зображення залишаються темними [2].

У контурі зображення сконцентрована інформація, яка характеризує його форму, що є важливо для сприйняття і розпізнаванні образів. Контурні точки є незначною частиною всього зображення. За допомогою них можна ефективно і аналітично просто описувати зображення об'єктів, які є інваріантні до основних перетворень (перенесення, масштабування і повороту). Задачі контурного аналізу актуальні в автоматизованих системах обробки зображень різноманітної природи: технічного зору, біологічних, медичних тощо [3].

Математично ідеальний контур – це розрив просторової функції рівнів яскравості в площині зображення. Тому виділення контура означає пошук найрізкіших змін, тобто максимумів модуля вектора градієнта [2]. Це є однією з причин, через яку контури залишаються в зображенні при шифруванні в системі RSA, оскільки шифрування тут оснований на піднесенні до степеня за модулем деякого натурального числа. При цьому на контурі і на сусідніх до контуру пікселях піднесення до степеня значення яскравостей дає ще більший розрив.

Існують різні алгоритми, які виділяють контури, наприклад, відстежуючі алгоритми. Відстежуючі алгоритми ґрунтуються на тому, що на зображенні відшукується об'єкт (точка об'єкта, яка зустрілася першою), і контур об'єкта відстежується і векторизується. Перевагою цього алгоритму є його простота, до недоліків належать їх послідовна реалізація і деяка складність під час пошуку і оброблення внутрішніх контурів.

Для визначення контурів у зображеннях використовують статистичний аналіз фрагментів зображення та їх взаємну кореляцію з метою знаходження стрибкоподібних змін кольору і освітленості. Велика група методів основана на використанні математичних моделей, що відображають певну взаємодію між окремими пікселями або фрагментами зображень. Також для розв'язання задач розпізнавання об'єктів застосовують різні методи фільтрації, наприклад, інверсні фільтри, фільтри Вінера, Байєса. При цьому використовують аналогію між динамікою зображень та фізичними процесами, наприклад, дифузії. Для розв'язання деяких задач використовують стохастичні моделі.

Алгоритми сегментації зображень ґрунтуються на одній з двох характеристик сигналу яскравості – розривності або однорідності. В першому випадку підхід ґрунтується на розбитті зображення на основі різких змін сигналу, таких як перепади яскравості на зображенні. Зазвичай шукають розриви за допомогою ковзних масок. Друга категорія методів ґрунтується на визначенні однорідності зображення за наперед обраними критеріями.



Рис. 1. Виокремлення контурів у зображенні

Надалі приймемо, що зображенню відповідає матриця кольорів [4,5]

$$C = \begin{pmatrix} c_{1,1} & \dots & c_{1,m} \\ \dots & \dots & \dots \\ c_{n,1} & \dots & c_{n,m} \end{pmatrix}. \quad (1)$$

Відносно зображення існують певні проблеми його шифрування, а саме частково зберігаються контури на різко флуктуаційних зображеннях [4]. У [5] для шифрування – дешифрування зображень в градаціях сірого було запропоновано використовувати квадратичні фрактальні перетворення. Тут для шифрування – дешифрування таких зображень пропонується використовувати кубічні фрактальні перетворення.

Шифрування і дешифрування за одним рядком матриці зображення

Нехай P, Q – пара довільних простих чисел і $N = P \cdot Q$, $exd \circ 1 \pmod{j(N)}$, $j(N) = (P-1)(Q-1)$, $F = P^e \pmod{j(N)}$, $G = Q^d \pmod{j(N)}$.

Шифрування відбувається з використанням кубічного фрактального перетворення двох сусідніх елементів рядка в матриці зображення C за такими співвідношеннями:

$$\begin{cases} u_{n,k} = F^3 u_{n,k-1}^3 + G^3 u_{n+1,k-1}^3 \\ u_{n+1,k} = F u_{n,k-1} + G u_{n+1,k-1} \end{cases} \quad (2)$$

$n = 1, 2, \dots, N_0$, N_0 – число елементів у рядку; k – номер фрактальної ітерації, $u_{n,0} = u_n$, $u_{n+1,0} = u_{n+1}$. Дешифрують формулами

$$u_{n,k-1} = \frac{3u_{n,k} \pm \sqrt{D}}{6F}; \quad (3)$$

$$u_{n+1,k-1} = \frac{3u_{n,k} \mp \sqrt{D}}{6G}, \quad (4)$$

де $D = 12u_{n,k} / u_{n+1,k} - 3u_{n+1,k}$.

Результати наведено на рис. 2–4.

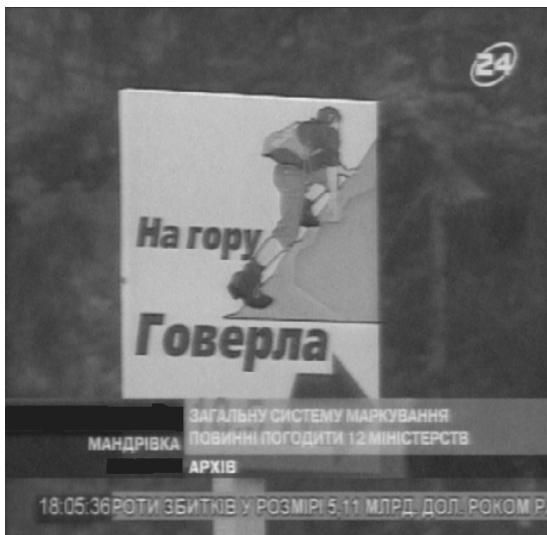


Рис. 2. Початкове зображення

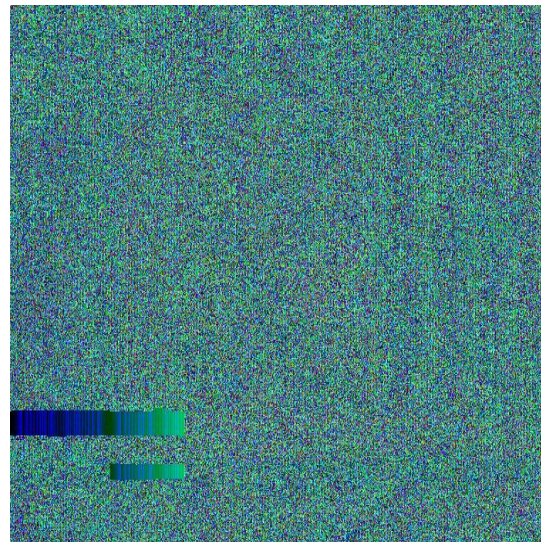


Рис. 3. Зашифроване зображення



Рис. 4. Дешифроване зображення

Шифрування з використанням іншого кубічного фрактального перетворення двох сусідніх елементів рядка в матриці зображення C реалізується за такими співвідношеннями:

$$\begin{cases} u_{n,k} = F^3 u_{n,k-1} - G^3 u_{n+1,k-1}, \\ u_{n+1,k} = F u_{n,k-1} - G u_{n+1,k-1} \end{cases}, \quad (5)$$

де $n = 1, 2, \dots, N_0$, N_0 – число елементів у рядку; k – номер фрактальної ітерації; $u_{n,0} = u_n$, $u_{n+1,0} = u_{n+1}$.

Для дешифрування використовуються формули

$$u_{n,k-1} = \frac{3u_{n,k} \pm \sqrt{D}}{6F}; \quad (6)$$

$$u_{n+1,k-1} = \frac{-3u_{n,k} \mp \sqrt{D}}{6G}, \quad (7)$$

Результати наведено на рис. 5–7. Очевидно, що шифрування і дешифрування суттєво залежать від вибору простих P , Q , R , T , а також від числа фрактальних ітерацій.



Рис. 5. Початкове зображення

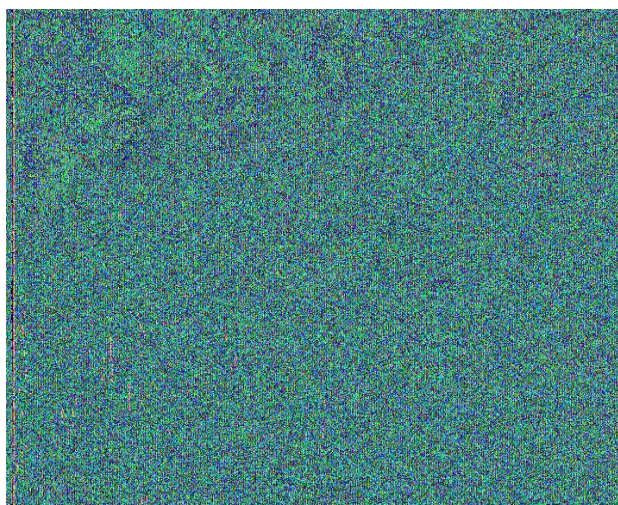


Рис. 6. Зашифроване зображення



Рис. 7. Дешифроване зображення

Загальний вигляд форми програмного ужитку шифрування-дешифрування за формулами (2)–(7) наведено на рис. 8.



Рис. 8. Загальний вигляд форми програмного ужитку

Висновки

З порівняння рис. 3 і рис. 6 видно, що шифрування за формулами (2) дещо структурно відрізняється від шифрування за формулами (5). Контури в обох зашифрованих зображеннях відсутні. Дешифровані зображення в обох випадках є візуально еквівалентними, хоча при дешифруванні за формулами (5) спостерігається незначне затемнення зображення. Вказані алгоритми можна використати для оперативного передавання графічних зображень для отримання задовільного результату стосовно будь-якого типу зображень, але найбільших переваг досягають у випадку використання зображень, які дають змогу чітко виділяти контури. Підвищується стійкість шифрування, оскільки для шифрування і дешифрування використовуються довільні прості числа, які можуть бути доволі великими, і елементи алгоритму RSA. А від цього залежить стійкість криптографічного алгоритму. Криптографічна стійкість запропонованих алгоритмів вища, ніж алгоритму RSA.

Обидва типи запропонованих алгоритмів шифрування – дешифрування можна використовувати і стосовно кольорових зображень. Однак незалежно від типу зображення можуть виникати проблеми під час розв’язування відповідних алгебраїчних рівнянь.

1. Шнайер Б. Прикладная криптография. – М.: Триумф, 2003. – 815 с. 2. Яне Б. Цифровая обработка изображений. – М.: Техносфера, 2007. – 583 с. 3. Прэтт У. Цифровая обработка изображений. Кн. 1,2. – М.: Мир, 1982. 4. Фабрі Л., Ковальчук А., Ступень М. Шифрування і дешифрування зображень з використанням квадратичних фрактальних алгоритмів // Вісник Нац. ун-ту “Львівська політехніка”, “Комп’ютерні науки та інформаційні технології”. – № 694. – С. 180–184. 5. Цмоць І., Ковальчук А., Ступень М. Системи фрактальних алгоритмів в шифруванні – дешифруванні зображень з додатковим зашумленням // Вісник Нац. ун-ту “Львівська політехніка” “Комп’ютерні науки та інформаційні технології”. – № 732. – С. 288–293.

ANALYTIC SOLUTION FOR THE PROBLEM OF IN-SITU CALIBRATING MAGNETOMETRIC 3D-PROBES BASED ON NON-ORTHOGONALLY PLACED HALL DEVICES

© Marusenkova T., 2014

Запропоновано методику in-situ калібрування 3D-зонтів магнітного поля з лінійними вихідними характеристиками давачів за відомих параметрів кутового розсунення їхніх площин з гранями підкладки зонта. Одержані аналітичні вирази для розрахунку лінійних коефіцієнтів. Експериментально підтверджено ефективність запропонованої методики.

Ключові слова: 3D-зонд магнітного поля, методика калібрування, in-situ калібрування, давачі Холла.

This work presents a method of in-situ calibrating 3D magnetic sensors based on Hall devices that have linear field characteristics and are non-orthogonal, with known position of each Hall device in relation to the 3D sensor's substrate. Formulas for calculation of Hall devices linear coefficients are obtained. The method's efficiency is proved experimentally.

Key words: 3D magnetic sensors, calibration method, in-situ calibration, Hall devices

Introduction

Hall devices are widely used as magnetic sensors. They have a number of advantages including a wide measurement range ($10^{-5} \dots 10^2$ T), low power consumption (<10 mW) and small size (1..3 mm). The output signal of an ordinary Hall device is linearly dependent on the projection of the magnetic-field vector \vec{B} onto the normal N to its sensitive semiconductor layer. So, in order to obtain a measurement device for measuring three orthogonal projections B_x , B_y , B_z of the magnetic-field vector simultaneously, one integrates three Hall devices into a 3D-sensor placing them on the adjacent sides of the sensor's cubic substrate [1, 2]. At best three Hall devices in a 3D sensor are completely orthogonal, so in that case 3D-sensor's calibration comes to finding of the linear coefficient of each Hall device included in this sensor. However, the condition of the Hall devices complete perpendicularity is not practically met because of processing limits of 3D sensor's manufacturing. In case if the condition of strict perpendicularity is not fulfilled, calibration gets more complicated: in this case the parameters of Hall devices angular displacements are to found in addition to the above-mentioned linear coefficients.

A number of important magnetometry problems set requirements that sensors have to be stable when working for a long time in a harsh environment. Such problems include magnetic field measurement in magnetic deflection systems of charged particle accelerators, magnetic confinement systems of thermonuclear reactors, systems of magnetic diagnostics of hardware and pipelines etc. The main difficulty of producing magnetic field measurement devices in such a harsh environment is thermal and radiation instability of a Hall device sensitive layer. The described problem can be solved in two ways: using very stable semiconductors and recalibrating sensors.

Problems of improving stability of semiconductors used in Hall devices intended for operation in a harsh environment including radiation conditions of charged particle accelerators and thermonuclear reactors are covered in a number of publications [3, 4, 5]. A number of publications [6] cover in-situ calibration of semiconductor Hall devices by means of a miniature coil integrated into a magnetometric sensor. The coil periodically establishes a test magnetic field. The drift of Hall coefficient can be

determined by measuring the change of the Hall device output signal caused by test magnetic field established by the coil.

However in some cases test coils are not applicable. First reason for it is the fact that test coils can be impacted by low or high temperatures, so the field that they produce should be considered unstable. The second reason is that the magnitude of test magnetic flux density is rather small (does not exceed several mT), so at high outside magnetic or electromagnetic fields the Hall coefficients cannot be defined with sufficient accuracy. The third reason is the fact that test coils increase the magnetic sensor size significantly. This hamper Hall devices to be placed tightly thus restricting the spatial resolution of a 3D sensor.

That's why ordinary periodical calibration of 3D sensors in a magnetic measure or an electromagnet is still urgent in a number of magnetometry problems. The first calibration may be done by the manufacturer using high-precision mechanical devices for exact positioning the calibrated sensors in the uniform magnetic field. However, in-situ calibration (calibration in work environment) sets the condition of involving few if any additional devices, so use of expensive high-precision positioning devices is undesirable for in-situ calibration.

Problem statement and the work purpose

In this work we introduce a new method of periodical refinement of linear coefficients in the output signals of Hall devices included in the 3D sensor being calibrated in the case if the strict perpendicularity condition is not met and the parameters of angular displacements are known. The problem of finding the angular displacements is solved in [7] using numerical methods with high computational complexity. The same solution can be used for refinement of the linear coefficients.

However, the growing significance of intelligent sensors cause the problem of finding such a 3d-sensor's in-situ calibration method that would possess the minimal computational complexity and thus can be implemented by microconverters included in intelligent sensors based on magnetometric 3D-sensors.

We've obtained analytic formulas for the linear coefficients of Hall devices combined into a magnetic 3D-sensor. The derived formulas use the parameters of the mentioned angular displacements that are defined once during the first calibration process and sets of Hall devices output signals measured when rotating a 3D sensor in a uniform magnetic field in arbitrary manner. The proposed method requires no high-precision mechanisms that are usually used for positioning 3D-probes in a magnetic field.

Method description

The set problem has concern to in-situ calibrating magnetometric 3D-sensors, but the mathematical tool for its solution can be used for solving the problems of measuring vector quantities in general. That's why all analytic dependencies given below are formulated abstractedly with undisclosed physical meaning of the formulas terms.

Let the vector quantity $\vec{M} = (x \ y \ z)$ being measured have the invariant module:

$$M = \sqrt{x^2 + y^2 + z^2} \quad (1)$$

Let's suppose that the vector quantity is to be measured by a 3D-probe with three sensors on adjacent sides of its cubic substrate and each sensor's output signal linearly depends on the vector quantity's projection normal to its plane:

$$\begin{aligned} S_1 &= R_1 M_{N1} \\ S_2 &= R_2 M_{N2} \\ S_3 &= R_3 M_{N3} \end{aligned} \quad (2)$$

there S_1, S_2, S_3 are output signals, R_1, R_2, R_3 are the linear coefficients of three sensors in a 3D-probe, M_{N1}, M_{N2}, M_{N3} are the projections of the vector \vec{M} which are normal to the planes of three sensors.

If the planes of all sensors combined into a 3D-probe are completely perpendicular then one can calibrate such a 3D-probe (and thus find the linear coefficients R_1, R_2, R_3) rotating it arbitrarily in a uniform test vector field and reading its output signals. This calibration method is given in [8].

Let the condition of strict perpendicularity be broken. Having the direction cosines of the perpendiculars to the planes of sensors combined into a 3D-probe marked as $(t_{11}, t_{12}, t_{13}), (t_{21}, t_{22}, t_{23}), (t_{31}, t_{32}, t_{33})$ in some Cartesian coordinates xyz assigned to this 3D-probe, we obtain the equations:

$$\begin{aligned} R_1 t_{11} x + R_1 t_{12} y + R_1 t_{13} z &= S_1 \\ R_2 t_{21} x + R_2 t_{22} y + R_2 t_{23} z &= S_2 \\ R_3 t_{31} x + R_3 t_{32} y + R_3 t_{33} z &= S_3 \end{aligned} \quad (3)$$

The work [7] introduces the method of calibrating magnetic 3D-sensor with output signals that are described by equations (3). The parameters of sensor's angular displacements (i.e. the direction cosines t_{ij} , $i = \overline{1,3}$, $j = \overline{1,3}$) remain invariant during sensor's operation but the linear coefficients R_1, R_2, R_3 can change.

Let's find the linear coefficients R_1, R_2, R_3 by arbitrary rotation of a 3D-sensor in a space where the vector quantity has the invariant module M (values t_{ij} ($i = \overline{1,3}$, $j = \overline{1,3}$) are assumed to be known).

If the set of equations (3) possesses a solution then its matrix A has an inverse matrix C :

$$A = \begin{pmatrix} R_1 t_{11} & R_1 t_{12} & R_1 t_{13} \\ R_2 t_{21} & R_2 t_{22} & R_2 t_{23} \\ R_3 t_{31} & R_3 t_{32} & R_3 t_{33} \end{pmatrix}; \quad C = A^{-1} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} \quad (4)$$

where c_{ij} are some unknown coefficients.

Using the inverse matrix (4) we can write unknown components of the vector \overline{M} in the way:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} \cdot \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} \quad (5)$$

So, the components of the vector being measured can be written in the following way:

$$x = c_{11} \cdot S_1 + c_{12} \cdot S_2 + c_{13} \cdot S_3, \quad y = c_{21} \cdot S_1 + c_{22} \cdot S_2 + c_{23} \cdot S_3, \quad z = c_{31} \cdot S_1 + c_{32} \cdot S_2 + c_{33} \cdot S_3.$$

Since the measured vector's module (1) is invariant in a uniform field, for each 3D-sensor's position in this field we can write:

$$(c_{11} S_1 + c_{12} S_2 + c_{13} S_3)^2 + (c_{21} S_1 + c_{22} S_2 + c_{23} S_3)^2 + (c_{31} S_1 + c_{32} S_2 + c_{33} S_3)^2 = M^2 \quad (6)$$

Having opened the brackets and combined similar terms we obtain:

$$\begin{aligned} &(c_{11}^2 + c_{21}^2 + c_{31}^2) S_1^2 + (c_{12}^2 + c_{22}^2 + c_{32}^2) S_2^2 + (c_{13}^2 + c_{23}^2 + c_{33}^2) S_3^2 + \\ &2(c_{11} c_{12} + c_{21} c_{22} + c_{31} c_{32}) S_1 S_2 + 2(c_{11} c_{13} + c_{21} c_{23} + c_{31} c_{33}) S_1 S_3 + \\ &+ 2(c_{12} c_{13} + c_{22} c_{23} + c_{32} c_{33}) S_2 S_3 = M^2 \end{aligned} \quad (7)$$

Having measured the output signals S_1, S_2, S_3 for different sensor's position in a uniform field (the measured vector's components x, y, z are unknown) we compose a set of equations of form (7). This set is a set of linear algebraic equations in 6 unknowns: $(c_{11}^2 + c_{21}^2 + c_{31}^2), (c_{12}^2 + c_{22}^2 + c_{32}^2), (c_{13}^2 + c_{23}^2 + c_{33}^2), (c_{11} c_{12} + c_{21} c_{22} + c_{31} c_{32}), (c_{11} c_{13} + c_{21} c_{23} + c_{31} c_{33}), (c_{12} c_{13} + c_{22} c_{23} + c_{32} c_{33})$. The least number of the output signals set is 6, as follows from the amount of unknowns. The solutions of the set of at least 6 equations of form (7) are some numbers $c_1, c_2, c_3, v_1, v_2, v_3$ which are connected with the matrix C 's elements by the relationship:

$$\begin{aligned} c_{11}^2 + c_{21}^2 + c_{31}^2 &= c_1^2; \\ c_{12}^2 + c_{22}^2 + c_{32}^2 &= c_2^2; \\ c_{13}^2 + c_{23}^2 + c_{33}^2 &= c_3^2; \\ 2c_{11} c_{21} + 2c_{12} c_{22} + 2c_{13} c_{23} &= v_1 \\ 2c_{11} c_{31} + 2c_{12} c_{32} + 2c_{13} c_{33} &= v_2 \\ 2c_{21} c_{31} + 2c_{22} c_{32} + 2c_{23} c_{33} &= v_3 \end{aligned} \quad (8)$$

The matrix $A^{-1} = C$ can be written using the matrix A elements:

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} R_2 R_3 (t_{22} t_{33} - t_{23} t_{32}) & R_1 R_3 (t_{13} t_{32} - t_{12} t_{33}) & R_1 R_2 (t_{12} t_{23} - t_{13} t_{22}) \\ R_2 R_3 (t_{23} t_{31} - t_{21} t_{33}) & R_1 R_3 (t_{11} t_{33} - t_{13} t_{31}) & R_1 R_2 (t_{13} t_{21} - t_{11} t_{23}) \\ R_2 R_3 (t_{21} t_{32} - t_{22} t_{31}) & R_1 R_3 (t_{12} t_{31} - t_{11} t_{32}) & R_1 R_2 (t_{11} t_{22} - t_{12} t_{21}) \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} \quad (9)$$

Substituting c_{ij} expressed through relevant elements of the matrix A^{-1} we obtain:

$$\begin{aligned} & 2R_1 R_2 R_3^2 (t_{22} t_{33} - t_{23} t_{32})(t_{13} t_{32} - t_{12} t_{33}) + 2R_1 R_2 R_3^2 (t_{23} t_{31} - t_{21} t_{33})(t_{11} t_{33} - t_{13} t_{31}) + \\ & + 2R_1 R_2 R_3^2 (t_{21} t_{32} - t_{22} t_{31})(t_{12} t_{31} - t_{11} t_{32}) = v_1 (\det A)^2 \\ & 2R_1 R_2^2 R_3 (t_{22} t_{33} - t_{23} t_{32})(t_{12} t_{23} - t_{13} t_{22}) + 2R_1 R_2^2 R_3 (t_{23} t_{31} - t_{21} t_{33})(t_{13} t_{21} - t_{11} t_{23}) + \\ & + 2R_1 R_2^2 R_3 (t_{21} t_{32} - t_{22} t_{31})(t_{11} t_{22} - t_{12} t_{21}) = v_2 (\det A)^2 \\ & 2R_1^2 R_2 R_3 (t_{13} t_{32} - t_{12} t_{33})(t_{12} t_{23} - t_{13} t_{22}) + 2R_1^2 R_2 R_3 (t_{11} t_{33} - t_{13} t_{31})(t_{13} t_{21} - t_{11} t_{23}) + \\ & + 2R_1^2 R_2 R_3 (t_{12} t_{31} - t_{11} t_{32})(t_{11} t_{22} - t_{12} t_{21}) = v_3 (\det A)^2 \end{aligned} \quad (10)$$

The determinant of A is as follows:

$$\begin{aligned} \det A &= R_1 R_2 R_3 (t_{11} t_{22} t_{33} - t_{11} t_{23} t_{32}) + R_1 R_2 R_3 (t_{12} t_{23} t_{31} - t_{12} t_{21} t_{33}) + R_1 R_2 R_3 (t_{13} t_{21} t_{32} - t_{13} t_{22} t_{31}) = \\ &= R_1 R_2 R_3 (t_{11} t_{22} t_{33} - t_{11} t_{23} t_{32} + t_{12} t_{23} t_{31} - t_{12} t_{21} t_{33} + t_{13} t_{21} t_{32} - t_{13} t_{22} t_{31}) \end{aligned} \quad (11)$$

Taking into consideration equations (10) and (11) we obtain:

$$\begin{aligned} R_1 R_2 &= \frac{2((t_{22} t_{33} - t_{23} t_{32})(t_{13} t_{32} - t_{12} t_{33}) + (t_{23} t_{31} - t_{21} t_{33})(t_{11} t_{33} - t_{13} t_{31}) + (t_{21} t_{32} - t_{22} t_{31})(t_{12} t_{31} - t_{11} t_{32}))}{v_1 (t_{11} t_{22} t_{33} - t_{11} t_{23} t_{32} + t_{12} t_{23} t_{31} - t_{12} t_{21} t_{33} + t_{13} t_{21} t_{32} - t_{13} t_{22} t_{31})^2} \\ R_1 R_3 &= \frac{2((t_{22} t_{33} - t_{23} t_{32})(t_{12} t_{23} - t_{13} t_{22}) + (t_{23} t_{31} - t_{21} t_{33})(t_{13} t_{21} - t_{11} t_{23}) + (t_{21} t_{32} - t_{22} t_{31})(t_{11} t_{22} - t_{12} t_{21}))}{v_2 (t_{11} t_{22} t_{33} - t_{11} t_{23} t_{32} + t_{12} t_{23} t_{31} - t_{12} t_{21} t_{33} + t_{13} t_{21} t_{32} - t_{13} t_{22} t_{31})^2} \\ R_2 R_3 &= \frac{2((t_{13} t_{32} - t_{12} t_{33})(t_{12} t_{23} - t_{13} t_{22}) + (t_{11} t_{33} - t_{13} t_{31})(t_{13} t_{21} - t_{11} t_{23}) + (t_{12} t_{31} - t_{11} t_{32})(t_{11} t_{22} - t_{12} t_{21}))}{v_3 (t_{11} t_{22} t_{33} - t_{11} t_{23} t_{32} + t_{12} t_{23} t_{31} - t_{12} t_{21} t_{33} + t_{13} t_{21} t_{32} - t_{13} t_{22} t_{31})^2} \end{aligned} \quad (12)$$

Let's mark the right sides of all equations (12) as P_1 , P_2 , P_3 respectively in order to make them more concise. Using the third equation of the equations set (12) R_3 can be written in this way: $R_3 = \frac{P_3}{R_2}$.

Substituting this expression in the second equation of the set (12) we can express R_2 in the following way:

$R_2 = \frac{P_3 R_1}{P_2}$. After substitution of the obtained value in the first equation in the set (12) we derive the

following expression: $R_1^2 = \frac{P_1 P_2}{P_3}$, i.e. $R_1 = \pm \sqrt{\frac{P_1 P_2}{P_3}}$.

Hence we can know the module of R_1 . Its sign is to be defined by the sign of the sensor's output signal (the vector \overline{M} direction must be known). The values of R_2 and R_3 can be derived from the second and the third equations of the set (12).

Method approbation

The proposed in-situ calibration method was experimentally approved using a 3D probe based on Hall devices with output signals that can be approximately described by functions (3). In the case of Hall devices the sought values R_1 , R_2 , R_3 have the following physical meaning:

$$R_i = R_{Hi} \frac{I_i}{d_i} \quad (13)$$

where R_{Hi} is the Hall coefficient of the Hall sensor numbered i , I_i is the current through its current electrodes, d_i is the thickness of its sensitive layer.

The direct experiments with 3D magnetic sensors calibration in a harsh environment can be hardly conducted because they can be too time-consuming and require a year or more. So in order to perform an approbation of the proposed calibration method we used the approach based on the imitation of the Hall coefficients instability by means of arbitrary changing the feed current of the Hall devices. As follows from (13), the Hall coefficient R_{Hi} of a sensor and its feed current I_i influence the sought coefficient R_i in the same way. So if we change the current I_i within certain limits and consider this change random and unknown, thus we conduct experiments with imitation of instability of the linear coefficient R_i . During the conducted experiments we changed the feed current of each Hall device in the 3D magnetic sensor independently by $\pm 5\%$ of its nominal value (20 mA).

At the first calibration stage the 3D magnetic sensor was placed into a magnetic shield in order to measure the offset voltages of all Hall devices included into it. Then we measured the 3D probe's output signals S_1, S_2, S_3 for its 100 arbitrary positions in a test uniform magnetic field with magnetic flux density $B_{REF} = 127$ mT and subtracted the corresponding offset voltages measured previously from the obtained values. The Fig. 1 shows results of measuring the magnetic-field vector B_{ER} at the following changes of the feed currents: $dI_1 = -2.9\%$, $dI_2 = 4.2\%$, $dI_3 = -0.8\%$ without correction of the linear coefficients (B_{ER}) and with applied correction in accordance to the proposed method (B_{COR}).

It can be seen that with no correction the deviation of the calculated values of the magnetic-field vector B was in range from 123 to 133 mT, which means the error $\pm 4.6\%$ while refinement of the linear coefficients in accordance with the proposed method provided deviation of the calculated magnetic-field vector's module that does not exceed 0.3%. This error up to 0.3% is caused by the fact that Hall devices output signals are described by equations (3) approximately.

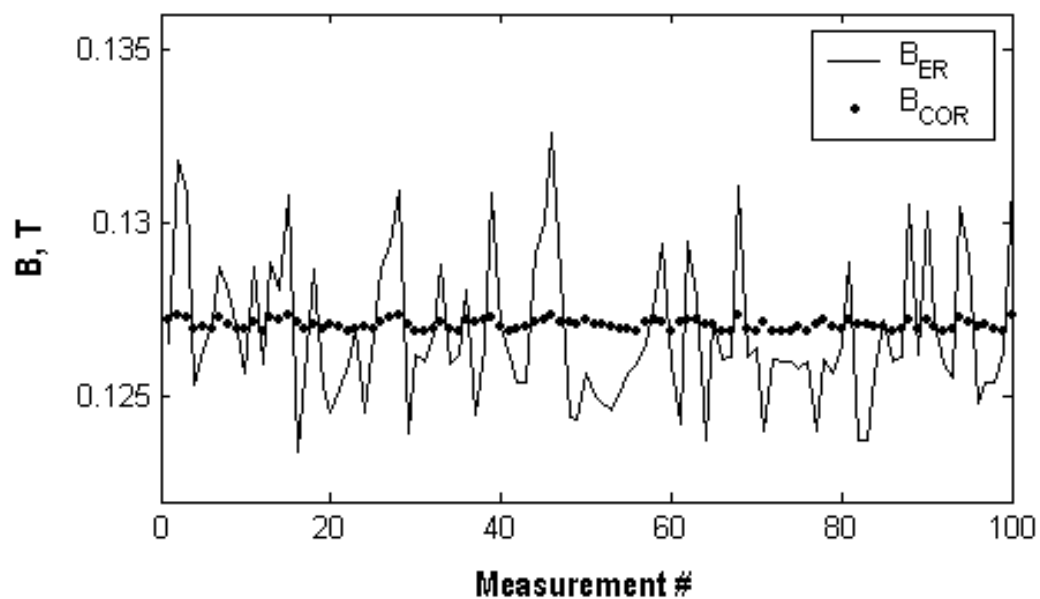


Fig. 2 The results of calculations of the magnetic-field vector module with no correction of linear coefficients (B_{ER}) and after such a correction (B_{COR}).

Conclusion

The work introduces a new method of in-situ calibrating a 3D-probe that is intended for vector quantities measurement and composed from sensors non-orthogonally placed on adjacent sides of its cubic substrate. The parameters of sensors angular displacements are assumed to be known. Each sensor's output signal is linearly dependent on the measured vector's projection normal to its plane. Analytic formulas for the linear coefficients of the 3D-probe's output signals were obtained. The proposed method is based on measurement of the 3D-sensor's output signals by means of its arbitrary rotation in a uniform field (3D-probe's current position is unknown). The method requires no tools for positioning calibrated 3D-sensors in

a test field. In the case if need for development of an intelligent sensor based on 3D-sensors for vector quantities measurement arises the proposed method of refining linear coefficients can be implemented in software of intelligent sensors.

1. Popovic R.S. *Hall Effect Devices* / R.S. Popovic – Adam Hilger, Bristol, Philadelphia and New York, 2002. – 420 p. 2. Hotra Z.Yu. *Microelectronic magnetic sensors* / Z.Yu.Hotra [et al.]. – Lviv: Publishing of Lviv Polytechnic National University, 2001. 3. Bolshakova I. *High precision mapper for cyclotron magnet* / I. Bolshakova, R. Holyaka, V. Erashok, M. Kumada // *IEEE Transactions on Applied Superconductivity*. – 2004. – Vol.14. №2. – P.1818 – 1821. 4. Antokhin E. *Permanent Magnet System for PET Cyclotron* / E. Antokhin, M. Kumada, Y. Iwashita, S. Wakasa, S. Matsumoto, T. Fujisawa, I. Bolshakova, R. Holyaka, V. Erashok, E. Sugiyama, H. Shimizu // *IEEE Transactions on Applied Superconductivity*. – 2008. – Vol.18, №2. – P. 965-968. 5. Bolshakova I. *Stable semiconductor magnetic field sensors under dozes of high radiation* / I. Bolshakova, V. Brudnyi, R. Holyaka [et al.] // *Plasma Physics and Controlled Fusion*. – 2003. – Vol.45. – Num.12A. – P.4.68. 6. Bolshakova I. *Way of providing radiation resistance of magnetic field semiconductor sensor* / I. Bolshakova, S. Krukovskii, R. Holyaka [et al.] // *Radiation Physics and Chemistry*. – 2001.– №61. – PP.743-745. 7. Bolshakova I.A. *Methods of modeling and calibrating 3D magnetic sensors based on splitted Hall structures* / I.A. Bolshakova, R.L. Holyaka, Z.Yu. Hotra, T.A. Marusenkova // *Electronics and communications. Electronics and nanotechnologies*. – 2011. – № 2(61). – PP. 34 – 38. 8. Hotra Z. Yu. *Arbitrary rotation method for 3D magnetic sensors calibration [CAD Systems in Microelectronics (CADSM), 2011 11th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics]*, (Lviv-Polyana, Feb 23-25 2011) / Z.Yu. Hotra, R.L. Holyaka, I.A. Bolshakova [et al.] – 2011. – PP. 413–416.

УДК 517.957:550.837

Л. Журавчак, Ю. Федоришин

Карпатське відділення інституту геофізики ім. С. І. Субботіна НАН України

РОЗПІЗНАВАННЯ ПРОВІДНИХ ТА ВИСОМНИХ ВКЛЮЧЕНЬ У КУСКОВО-ОДНОРІДНОМУ ПІВПРОСТОРИ ПРИ МАТЕМАТИЧНОМУ МОДЕЛЮВАННІ УСТАЛЕНИХ КОЛИВАНЬ ЕЛЕКТРОМАГНІТНОГО ПОЛЯ

© Журавчак Л., Федоришин Ю., 2014

Для кусково-однорідного провідного півпростору розглянуто аналітично-числову методику знаходження компонент векторів напруженості електромагнітного поля, збудженого стороннім джерелом струму. Проведено числові експерименти для повної, з урахуванням струмів зміщення, та квазістаціонарної моделей, досліджено вплив включення з вищою та нижчою, ніж у геосередовищі, провідністю на розподіл поля на денній поверхні.

Ключові слова: система рівнянь Максвелла; система рівнянь Гельмгольца; усталені коливання електромагнітного поля; непрямий метод граничних елементів.

The numerical-analytic technique for finding electric and magnetic components of electromagnetic field in a piecewise homogeneous conductive half-space is suggested. Electromagnetic field is excited by a horizontal contour with current harmonically changing in time. The problem is formulated and solved by means of the boundary element method.

Key words: Maxwell's equation system, the Helmholtz equation system, established EMF fluctuations; indirect boundary element method.

Вступ

Сьогодні гармонічно змінні в часі електромагнітні поля (ЕМП) в однорідних та горизонтально-шаруватих моделях земної кори за дії природних та штучних джерел досліджені за допомогою добре розвинутої теорії спектрального аналізу. Це особливо стосується широкого класу

двовимірних магнітотелуричних задач, коли зовнішнє поле задається у вигляді однорідного магнітного поля або плоскої хвилі. Відомі також аналітичні розв'язки задач розподілу ЕМП в порівняно простих електричних умовах для чужорідних локальних включень канонічної чи близької до неї форми у дво- та тривимірних горизонтально-шаруватих моделях [1, 2]. Щодо об'єктів із включеннями довільної форми, які краще відображають реальну геоелектричну ситуацію, то для математичного моделювання ЕМП, збудженого штучними джерелами, в останні роки на базі сучасних швидкодіючих комп'ютерів все частіше використовують числові та аналітично-числові методи розв'язування прямих задач геоелектрики. Застосування найпоширеніших різницевих методів [3, 4] чи методів скінченних елементів [5, 6] дає хорошу точність результатів, але вимагає покриття сіткою всієї області, яку займає різномірне тіло, що потребує великих обсягів пам'яті й програм для обернення матриць великих розмірів. Використання методу граничних інтегральних рівнянь [7–9] та створених на його основі методів граничних елементів [10] дає можливість дискретизувати тільки граничну поверхню об'єкта та поверхні розподілу середовища, що економить обсяг оперативної пам'яті під час реалізації алгоритму та дає порівняно високу точність розрахунків у внутрішніх точках.

У цій роботі запропоновано і апробовано аналітично-числовий підхід, який ґрунтується на непрямому методі граничних елементів, до розв'язування прямих тривимірних задач теорії ЕМП у випадку усталених коливань для провідного кускового-однорідного півпростору, враховано струми зміщення порівняно з квазістаціонарною моделлю, а також досліджено вплив включень з вищою та нижчою (на прикладі нафто- та газоносних покладів), ніж у геосередовищі, провідністю на розподіл поля.

Формулювання задачі

Розглянемо кусково-однорідний півпростір, що у декартовій системі координат займає область $\Omega = R^{3-} = \{(x_1, x_2, x_3) : -\infty < x_1 < \infty, -\infty < x_2 < \infty, -\infty < x_3 < 0\}$ і містить включення Ω_2 (просторове тіло у формі паралелепіпеда), яке перебуває в ідеальному електромагнітному контакті з геосередовищем $\Omega_1 = \Omega / \Omega_2$ і має відмінні від середовища електромагнітні характеристики. Середовище Ω_1 і включення Ω_2 характеризуються постійними електропровідностями σ_1, σ_2 , магнітними проникностями μ_1, μ_2 і діелектричними проникностями $\varepsilon_1, \varepsilon_2$ відповідно. На поверхні півпростору $\Gamma = \{(x_1, x_2, x_3) : -\infty < x_1 < \infty, -\infty < x_2 < \infty, x_3 = 0\}$ електричне поле (ЕП) відсутнє, а в контурі $C \subset \Omega_1$ діє сторонній струм інтенсивності $\mathbf{y}^{\mathbf{r}}(x, t) = (y_1(x, t), y_2(x, t), y_3(x, t))$, де τ – часова змінна, $x = (x_1, x_2, x_3)$ – точка півпростору. В початковий момент часу спостерігається нульовий розподіл компонент вектора напруженості ЕП $E_i^l(x, \tau)$ у середовищі Ω_1 та у включенні Ω_2 , $i=1,2,3$, $l=1,2$, та швидкості їх зміни.

Поширення ЕМП у провідному кусково-однорідному півпросторі, в зоні Ω_1 якого діють сторонні струми, описується системою рівнянь Максвелла:

$$\begin{array}{ll}
 \text{для геосередовища } ((x, \tau) \in \Omega_1 \times T) & \text{для включення } ((x, \tau) \in \Omega_2 \times T) \\
 \text{rot} \mathbf{H}^1(x, \tau) = \sigma_1 \mathbf{E}^1(x, \tau) + \varepsilon_1 \frac{\partial \mathbf{E}^1(x, \tau)}{\partial \tau} + \mathbf{\Psi}(x, \tau), & \text{rot} \mathbf{H}^2(x, \tau) = \sigma_2 \mathbf{E}^2(x, \tau) + \varepsilon_2 \frac{\partial \mathbf{E}^2(x, \tau)}{\partial \tau}, \\
 \text{div} \mathbf{H}^1(x, \tau) = 0, & \text{div} \mathbf{H}^2(x, \tau) = 0, \\
 \text{rot} \mathbf{E}^1(x, \tau) = -\mu_1 \frac{\partial \mathbf{H}^1(x, \tau)}{\partial \tau}, & \text{rot} \mathbf{E}^2(x, \tau) = -\mu_2 \frac{\partial \mathbf{H}^2(x, \tau)}{\partial \tau}, \\
 \text{div} \mathbf{E}^1(x, \tau) = 0, & \text{div} \mathbf{E}^2(x, \tau) = 0,
 \end{array} \tag{1}$$

де $\mathbf{H}^l(x, \tau) = (H_1^l(x, \tau), H_2^l(x, \tau), H_3^l(x, \tau))$, $\mathbf{E}^l(x, \tau) = (E_1^l(x, \tau), E_2^l(x, \tau), E_3^l(x, \tau))$ – вектори напруженості магнітного та електричного поля в Ω_l , $T = \{\tau : 0 < \tau < \infty\}$.

Для багатьох задач електродинаміки доцільно поділити рівняння Максвелла (1), тобто записати окремо рівняння для електричного та магнітного полів. Використовуючи властивості векторних диференціальних операцій, отримаємо систему телеграфних рівнянь для визначення невідомих компонент ЕП $E_j^1(x,t)$ ($j=1,2,3$) у середовищі Ω_1 та $E_j^2(x,t)$ у включенні Ω_2 :

$$\begin{aligned} \Delta E_j^1(x,t) - s_1 m_1 \frac{\partial E_j^1(x,t)}{\partial t} - e_1 m_1 \frac{\partial^2 E_j^1(x,t)}{\partial t^2} &= m_1 \frac{\partial y_j(x,t)}{t}, (x,t) \in \Omega_1 \times T, \\ \Delta E_j^2(x,t) - s_2 m_2 \frac{\partial E_j^2(x,t)}{\partial t} - e_2 m_2 \frac{\partial^2 E_j^2(x,t)}{\partial t^2} &= 0, (x,t) \in \Omega_2 \times T, \end{aligned} \quad (2)$$

доповнену граничними

$$E_j^l(x,t) = 0, (x,t) \in \Gamma \times T, \quad (3)$$

початковими умовами

$$E_j^l(x,t) = 0, \frac{\partial E_j^l(x,t)}{\partial t} = 0, x \in \Omega_l \text{ при } t = 0, \quad (4)$$

а також умовами ідеального електромагнітного контакту (неперервність тангенціальних компонент електричного і магнітного полів (МП) та стрибок нормальних) на межі поділу середовищ:

$$E_d^1(x,t) = E_d^2(x,t), \quad \frac{1}{m_1} \text{rot} E_d^1(x,t) = \frac{1}{m_2} \text{rot} E_d^2(x,t), \quad (x,t) \in \Gamma_{12} \times T, \quad (5)$$

$$s_1 E_n^1(x,t) = s_2 E_n^2(x,t), \quad \text{rot} E_n^1(x,t) = \text{rot} E_n^2(x,t),$$

Тут Δ – оператор Лапласа, $\Gamma_{12} = \partial\Omega_1 \cap \partial\Omega_2$, $\partial\Omega_l$ – межа зони Ω_l .

Якщо здійснювати спостереження через доволі тривалий час після зародження збурень, то можна припустити, що фізичні величини гармонічно змінюються в часі з кутовою частотою w , тобто ми маємо справу з задачею про усталені коливання. Припускаючи, що $E_j^l(x,t) = \underline{E}_j^l(x,w)e^{-iwt}$, $y_j(x,t) = \underline{y}_j(x,w)e^{-iwt}$, аналіз сильно спрощується, бо часова змінна виключається з початково-крайової задачі (2)–(5):

$$\Delta \underline{E}_j^1(x,w) + m_1 w(e_1 w + i s_1) \underline{E}_j^1(x,w) = -i w m_1 \underline{y}_j(x,w), x \in \Omega_1, \quad (6)$$

$$\Delta \underline{E}_j^2(x,w) + m_2 w(e_2 w + i s_2) \underline{E}_j^2(x,w) = 0, x \in \Omega_2,$$

$$\underline{E}_j^l(x,w) = 0, x \in \Gamma, \quad (7)$$

$$\underline{E}_d^1(x,w) = \underline{E}_d^2(x,w), \quad \frac{1}{m_1} \text{rot} \underline{E}_d^1(x,w) = \frac{1}{m_2} \text{rot} \underline{E}_d^2(x,w), \quad x \in \Gamma_{12} = \partial\Omega_1 \cap \partial\Omega_2. \quad (8)$$

$$s_1 \underline{E}_n^1(x,w) = s_2 \underline{E}_n^2(x,w), \quad \text{rot} \underline{E}_n^1(x,w) = \text{rot} \underline{E}_n^2(x,w),$$

Тут $\underline{E}_j^l(x,w) = \underline{E}_j^{l1}(x,w) + i \underline{E}_j^{l2}(x,w)$, $\underline{y}_j(x,w) = \underline{y}_j^1(x,w) + i \underline{y}_j^2(x,w)$ – комплексні амплітуди компонент вектора напруженості ЕП та сторонніх джерел струму.

Методика розв'язування. Інтегральне подання розв'язку

Для знаходження розв'язку задачі (6)–(8) використаємо непрямий метод граничних елементів. Згідно із цим методом границя розподілу середовища та включення розбивається на V областей (граничних елементів) Γ_v^{12} , таких що $\Gamma_{12} = \cup_{v=1}^V \Gamma_v^{12}$, $\Gamma_v^{12} \cap \Gamma_w^{12} = \emptyset$ при $v \neq w$. На кожному граничному елементі Γ_v^{12} задаються фіктивні джерела струму $j_{jv}^l(x,w) = j_{jv}^{l1}(x,w) + i j_{jv}^{l2}(x,w)$, ($j = 1,2,3$). Відповідно отримуємо нову систему рівнянь:

$$\Delta \underline{E}_j^1(x,w) + m_1 w(e_1 w + i s_1) \underline{E}_j^1(x,w) = - \sum_{v=1}^V j_{jv}^1(x,w) c_v - i w m_1 \underline{y}_j(x,w), \quad (9)$$

$$\Delta \underline{E}_j^2(x,w) + m_2 w(e_2 w + i s_2) \underline{E}_j^2(x,w) = - \sum_{v=1}^V j_{jv}^2(x,w) c_v,$$

де c_v – характеристична функція елемента Γ_v .

Використавши для півпростору без включення спеціальний фундаментальний розв'язок $\Phi_h^1(x, \mathbf{x}, w)$ рівняння Гельмгольца, який автоматично задовольняє граничну умову (7), та фундаментальний розв'язок $\Phi^2(x, \mathbf{x}, w)$ для включення, запишемо інтегральне подання розв'язку задачі (9), (7), (8) для компонент $\mathcal{E}_j^0(x, w)$ та $\mathcal{E}_j^2(x, w)$:

$$\mathcal{E}_j^0 = \sum_{v=1}^V \int_{\Gamma_v} \Phi_h^1(x, \mathbf{x}, w) j_{jv}^1(x, w) d\Gamma_v(x) + I_{cj}(x, w, \Phi_h^1), \quad \mathcal{E}_j^2 = \sum_{v=1}^V \int_{\Gamma_v} \Phi^2(x, \mathbf{x}, w) j_{jv}^2(x, w) d\Gamma_v(x), \quad (10)$$

а також одержані на його основі інтегральні зображення для похідних від цих компонент за координатами x_l :

$$\begin{aligned} \frac{\partial \mathcal{E}_j^0(x, w)}{\partial x_k} &= \sum_{v=1}^V \int_{\Gamma_v} Q_{hk}^1(x, \mathbf{x}, w) j_{jv}^1(x, w) d\Gamma_v(x) + I_{cj}(x, w, Q_{hk}^1), \\ \frac{\partial \mathcal{E}_j^2(x, w)}{\partial x_k} &= \sum_{v=1}^V \int_{\Gamma_v} Q_k^2(x, \mathbf{x}, w) j_{jv}^2(x, w) d\Gamma_v(x), \end{aligned} \quad (11)$$

де $F_h^1(x, \mathbf{x}, w) = F^1(r, w) - F^1(r', w)$, $F \in \{\Phi, Q_k\}$, $F^l(r, w) = F^l(x, \mathbf{x}, w) = F^{l1}(r, w) + iF^{l2}(r, w)$, $r^2 = \sum_{k=1}^3 (x_k - x_k)^2$, $r'^2 = \sum_{k=1}^2 (x_k - x'_k)^2$, $x'_k = x_k$ ($k = 1, 2$), $x'_3 = -x_3$, $\mathbf{x} = (x_1, x_2, x_3) \in R^3$,

$$\Phi^l(r, w) = \frac{\exp(ir\sqrt{m_l w}(e_l w + iS_l))}{4pr}, \quad \Phi^{l1}(r, w) = \frac{1}{4pr} \exp(-rA_l \sin(\frac{q_l}{2})) \cdot \cos(rA_l \cos(\frac{q_l}{2})),$$

$$\Phi^{l2}(r, w) = \frac{1}{4pr} \exp(-rA_l \sin(\frac{q_l}{2})) \cdot \sin(rA_l \cos(\frac{q_l}{2})), \quad Q_k^l(x, \mathbf{x}, w) = \frac{\partial \Phi^l(x, \mathbf{x}, w)}{\partial x_k} = (x_k - x_k) Q^l(r, w),$$

$$Q^{l1}(r, w) = -\frac{1}{r^2} \left(rA_l (\Phi^{l1} \sin(\frac{q_l}{2}) + \Phi^{l2} \cos(\frac{q_l}{2})) + \Phi^{l1} \right), \quad A_l = \sqrt{m_l w} \sqrt[4]{e_l^2 w^2 + S_l^2}, \quad q_l = \arctg\left(\frac{S_l}{e_l w}\right),$$

$$Q^{l2}(r, w) = \frac{1}{r^2} \left(rA_l (\Phi^{l1} \cos(\frac{q_l}{2}) - \Phi^{l2} \sin(\frac{q_l}{2})) - \Phi^{l2} \right),$$

$$I_{cj}(x, w, F_h^1) = iw m_l \int_C F_h^1(x, \mathbf{x}, w) y_j^0(x, w) dC(x) = I_{cj}^1(x, w, F_h^1) + iI_{cj}^2(x, w, F_h^1),$$

$$I_{cj}^1(x, w, F_h^1) = -w m_l \int_C (F_h^{11} y_j^0 + F_h^{12} y_j^1) dC(x), \quad I_{cj}^2(x, w, F_h^1) = w m_l \int_C (F_h^{11} y_j^1 - F_h^{12} y_j^0) dC(x).$$

Для спрощення алгоритму апроксимуємо невідомі функції $j_{jv}^{l1}(x, w)$, $j_{jv}^{l2}(x, w)$ константами d_{jv}^{l1} , d_{jv}^{l2} . Для їх знаходження побудуємо з використанням виразів (10), (11) систему лінійних алгебраїчних рівнянь (СЛАР), вимагаючи задоволення в колокаційному сенсі умов контакту (8).

Після розв'язання СЛАР, використовуючи формули (10), обчислюємо значення компонент ЕП у будь-яких точках x^p півпростору, які нас цікавлять:

$$\mathcal{E}_j^{01}(x^p, w) = \sum_{v=1}^V \left(d_v^{11} I_{vp}(\Phi_h^{11}) - d_v^{12} I_{vp}(\Phi_h^{12}) \right) + I_{cj}^1(x^p, w, \Phi_h^1),$$

$$\mathcal{E}_j^{02}(x^p, w) = \sum_{v=1}^V \left(d_v^{11} I_{vp}(\Phi_h^{12}) + d_v^{12} I_{vp}(\Phi_h^{11}) \right) + I_{cj}^2(x^p, w, \Phi_h^1),$$

$$\mathcal{E}_j^{21}(x^p, w) = \sum_{v=1}^V \left(d_v^{21} I_{vp}(\Phi^{21}) - d_v^{22} I_{vp}(\Phi^{22}) \right), \quad \mathcal{E}_j^{22}(x^p, w) = \sum_{v=1}^V \left(d_v^{21} I_{vp}(\Phi^{22}) + d_v^{22} I_{vp}(\Phi^{21}) \right),$$

де $I_{vp}(F) = \int_{\Gamma_v} F(x, \mathbf{x}, w) d\Gamma_v(x)$, $F \in \{\Phi_h^{11}, \Phi_h^{12}, \Phi^{22}, \Phi^{21}\}$.

Дійсні та уявні частини $\mathcal{H}_j^{01}(x^p, w)$, $\mathcal{H}_j^{02}(x^p, w)$ компонент вектора напруженості МП обчислимо за формулами:

$$\mathcal{H}_k^{01}(x^p, w) = \frac{1}{w m_l} \left(\frac{\partial E_s^{l2}}{\partial x_j} - \frac{\partial E_j^{l2}}{\partial x_s} \right), \quad \mathcal{H}_k^{02}(x^p, w) = -\frac{1}{w m_l} \left(\frac{\partial E_s^{l1}}{\partial x_j} - \frac{\partial E_j^{l1}}{\partial x_s} \right). \quad (12)$$

Тут для різних k використовуються різні пари індексів s та j , так, для $k=1 - s=3, j=2$, для $k=2 - s=1, j=3$, для $k=3 - s=2, j=1$.

Квазістаціонарна модель

У випадку квазістаціонарної моделі рівняння системи (6) мають вигляд:

$$\begin{aligned} \Delta \mathbb{E}_j^{\mathbb{P}^0}(x, w) + im_1 w s_1 \mathbb{E}_j^{\mathbb{P}^0}(x, w) &= -i w m_1 \mathbb{Y}_j^{\mathbb{P}^0}(x, w), \\ \Delta \mathbb{E}_j^{\mathbb{P}^2}(x, w) + im_2 w s_2 \mathbb{E}_j^{\mathbb{P}^2}(x, w) &= 0. \end{aligned} \quad (13)$$

Відповідно дещо спростяться вирази для фундаментальних розв'язків:

$$\begin{aligned} \Phi^{l1}(r, w) &= \frac{1}{4pr} \exp(-rA_l / \sqrt{2}) \cdot \cos(rA_l / \sqrt{2}), \quad \Phi^{l2}(r, w) = \frac{1}{4pr} \exp(-rA_l / \sqrt{2}) \cdot \sin(rA_l / \sqrt{2}), \\ Q^{l1}(r, w) &= -\frac{1}{r^2} (rA_l (\mathbb{P}^{l1} + \mathbb{P}^{l2}) / \sqrt{2} + \mathbb{P}^{l1}), \quad Q^{l2}(r, w) = \frac{1}{r^2} (rA_l (\mathbb{P}^{l1} - \mathbb{P}^{l2}) / \sqrt{2} - \mathbb{P}^{l1}), \end{aligned}$$

тут враховано, що $q_l = p / 2$, а для A_l маємо формулу $A_l = \sqrt{m_1 w s_1}$.

Числова реалізація

Задача 1. Для числових експериментів було розглянуто включення Ω_2 у формі паралелепіпеда розмірами $p_x=100, p_y=100, p_z=30$, розміщеного на глибині h_0 . Стороннім джерелом, яке збуджувало ЕМП, була квадратна рамка зі стороною $h=100$, розміщена на глибині $h_3=-0.001$. Залежність сили струму в ній описувалася функцією $\mathbb{Y}_j^{\mathbb{P}^0}(x, w) = C_j(x)$, де $C_j(x)$ – проекції одиничного вектора, колінеарного дотичній до контура C в точці $x \in C$. У цьому частковому випадку інтеграли, які входять у вирази (10), (11), мали вигляд:

$$\mathbb{Y}_{cj}^{\mathbb{P}^0}(x, w, \mathbb{P}^0) = -w m_1 \int_C \mathbb{P}^0 C_j(x) dC(x), \quad \mathbb{Y}_{cj}^{\mathbb{P}^2}(x, w, \mathbb{P}^0) = w m_2 \int_C \mathbb{P}^0 C_j(x) dC(x).$$

Кожну грань паралелепіпеда розбивали на 27 граничних елементів (рис. 1).

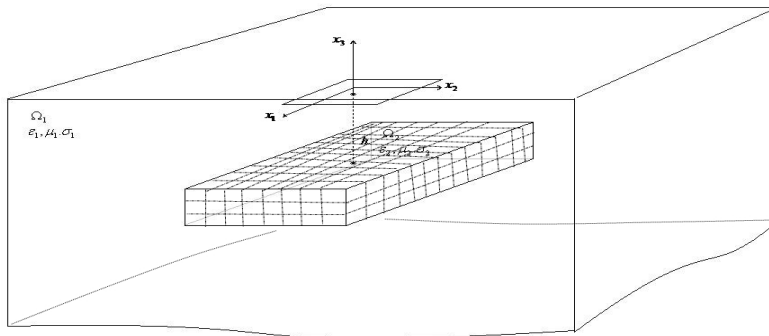


Рис. 1. Геоелектрична модель півпростору з включенням

На рис. 2 зображено залежність значень вертикальної компоненти $H_3^1(0, 0, h_3, w)$, одержаних в центрі рамки, від кутової частоти ω , для повної і квазістаціонарної моделей у випадку нафтового ($\sigma_2=0,25\sigma_1, \mu_2=0,99994\mu_1, \epsilon_2=2\epsilon_1$, ряди 3, 6), газового ($\sigma_2=0,1\sigma_1, \mu_2=1,00008\mu_1, \epsilon_2=\epsilon_1$, ряди 4, 7) та провідного ($\sigma_2=5\sigma_1, \mu_2=\mu_1, \epsilon_2=30\epsilon_1$, ряди 1, 5) включень, розміщених на глибині $h_0=40$; для порівняння наведено розподіл ЕМП в однорідному півпросторі $\sigma_1=1\text{См/м}, \mu_1=4\pi 10^{-7}\text{Гн/м}, \epsilon_1=15\epsilon_0, \epsilon_0=1/36 \pi 10^{-9} \text{ Ф/м}$ для повної моделі (ряд 2). Як бачимо, для високоомних включень значення $H_3^1(0, 0, h_3; w)$ на всьому діапазоні частот є меншими, ніж в однорідному півпросторі, а для провідних – більшими; причому врахування струмів зміщення дає вищі величини, ніж квазістаціонарна модель, що особливо проявляється в області високих частот.

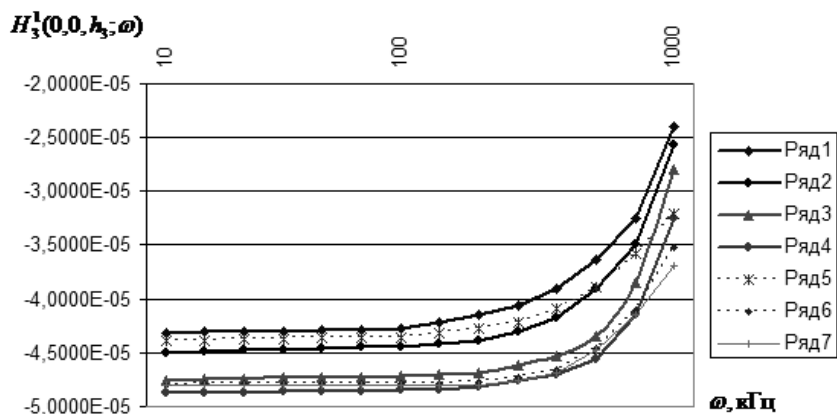


Рис. 2. Вплив електромагнітних параметрів включення на вертикальну компоненту вектора напруженості магнітного поля

Як відомо, під час інтерпретації результатів важливе значення мають теоретичні криві [11], які характеризують амплітуду $|r_w^f|/r_1$ та фазу q_w^f позірнього опору, тут

$$|r_w^f|/r_1 = \frac{\sqrt{(f^1 f^{01} + f^2 f^{02})^2 + (f^2 f^{01} - f^1 f^{02})^2}}{(f^{01})^2 + (f^{02})^2}, \quad q_w^f = \arctg \frac{f^2 f^{01} - f^1 f^{02}}{f^1 f^{01} + f^2 f^{02}}, \quad f^k \in \{H_1^{1k}, H_2^{1k}, H_3^{1k}, E_1^{1k}, E_2^{1k}, E_3^{1k}\}$$

($k=1,2$), f^1, f^{01}, f^2, f^{02} – дійсні та уявні частини компонент векторів напруженості ЕМП, індекс “0” вказує на величини в однорідному півпросторі з питомим опором $r_1 = 1/s_1$.

На рис. 3, 4 зображено графіки залежності теоретичних кривих $|r_w^f|/r_1$ та q_w^f від I_1/h_0 , побудованих на базі вертикальної компоненти МП, у центрі рамки для повної моделі.

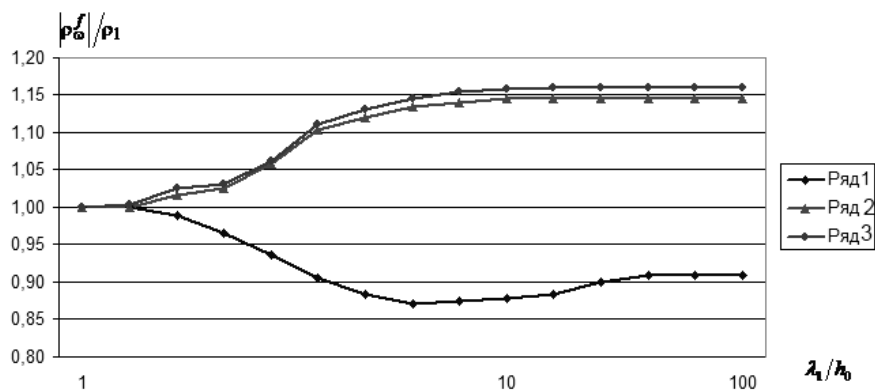


Рис. 3. Залежність амплітудних кривих від довжини хвилі

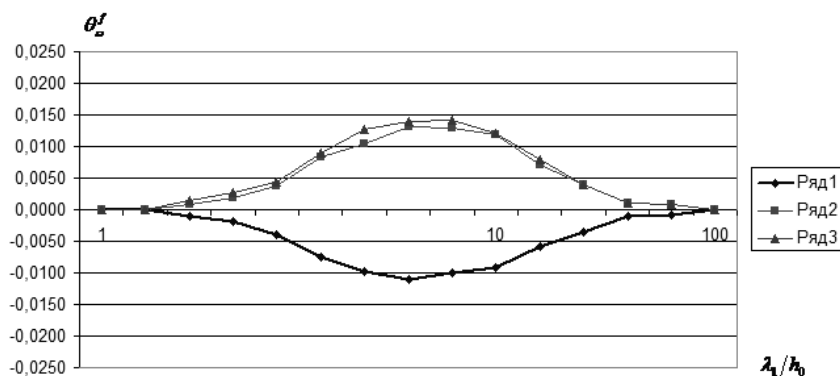


Рис. 4. Залежність фазових кривих від довжини хвилі

Також було розглянуто залежність величин $|r_w^f|/r_1$ та q_w^f , побудованих на базі вертикальної компоненти МП, від глибини залягання провідного включення ($\sigma_2=5\sigma_1$, $\mu_2=\mu_1$, $\epsilon_2=30\epsilon_1$) – $h_0=40$ (ряд 1), $h_0=80$ (ряд 2) і $h_0=100$ (ряд 3) для діапазону частот від 1кГц до 100 кГц. (рис. 5, 6).

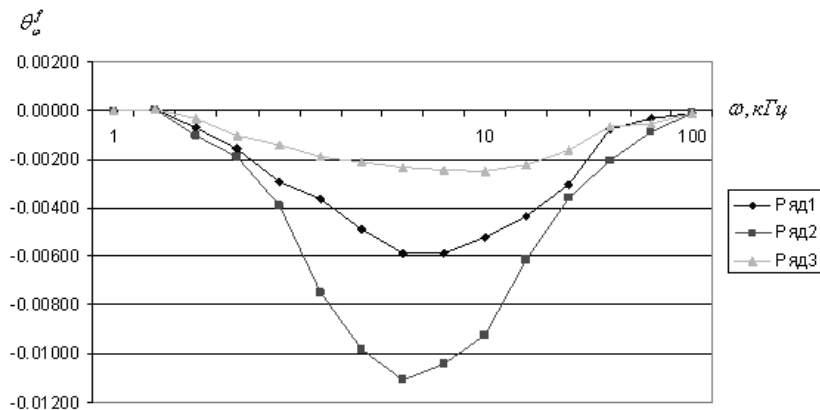


Рис. 5. Графік залежності фазових кривих від довжини хвилі та глибини залягання

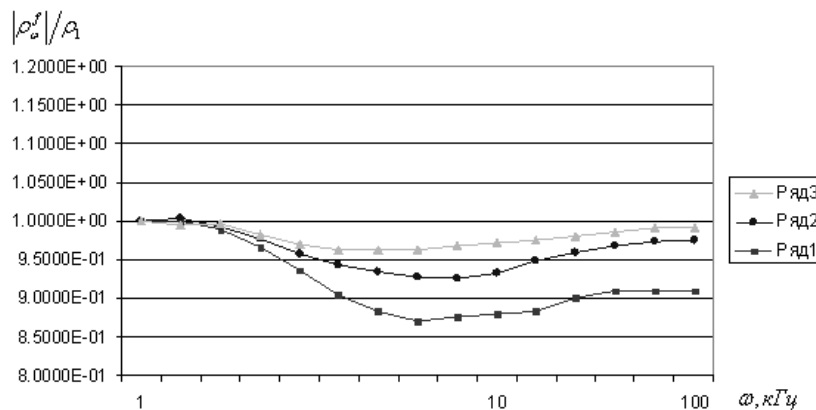


Рис. 6. Графік залежності амплітудних кривих від довжини хвилі та глибини залягання

Випадок однорідного об'єкта з довільною криволінійною межею

Задача знаходження компонент ЕМП для усталених коливань складається з системи рівнянь:

$$\Delta \tilde{E}_j(x, \omega) + \mu\omega(\epsilon\omega + i\sigma)\tilde{E}_j(x, \omega) = -i\omega\mu\tilde{\psi}_j(x, \omega), \quad x \in \Omega,$$

та граничних умов:

$$\tilde{E}_j(x, \omega) = \tilde{E}_{jz}(x, \omega), \quad x \in \Gamma. \quad (14)$$

СЛАР для знаходження інтенсивності невідомих фіктивних джерел струму будемо, задовольняючи граничні умови (14).

Задача 2. Запропонований підхід також було застосовано для моделювання усталених коливань у тривимірному однорідному об'єкті з електромагнітними характеристиками провідного (ряд 1) і нафтового (ряд 2) покладів у формі еліпсоїда з осями $a=3$, $b=2$, $c=1$. Межу тіла розбивали на 74 граничні елементи. Досліджено вплив стороннього джерела струму у вигляді квадратної рамки, аналогічної випадку кусково-однорідного півпростору (див. задачу 1), та крайової умови першого роду ($\tilde{E}_1(x, \omega) = \tilde{E}_2(x, \omega) = 0$, $\tilde{E}_3(x, \omega) = 0.000005 \cdot (x_3/c + 1)/2$) на компоненти ЕМП всередині тіла.

На рис. 7 подано залежність тангенціальної компоненти МП від кутової частоти (в діапазоні від 1 до 20 кГц) у центрі тіла; рис. 8 описує залежність вертикальної компоненти ЕП у точках тіла, рівномірно розподілених по осі Ox_3 на відрізку $[-1, 1]$.

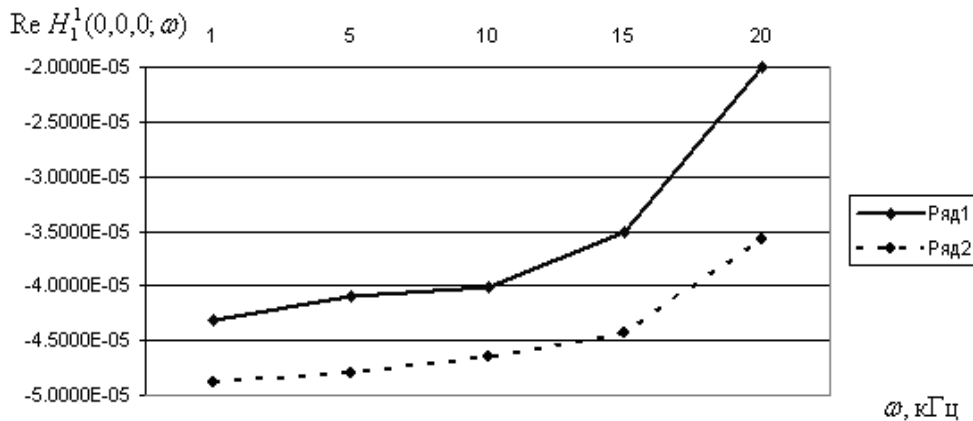


Рис. 7. Залежність дійсної компоненти МП від кутової частоти у центрі тіла

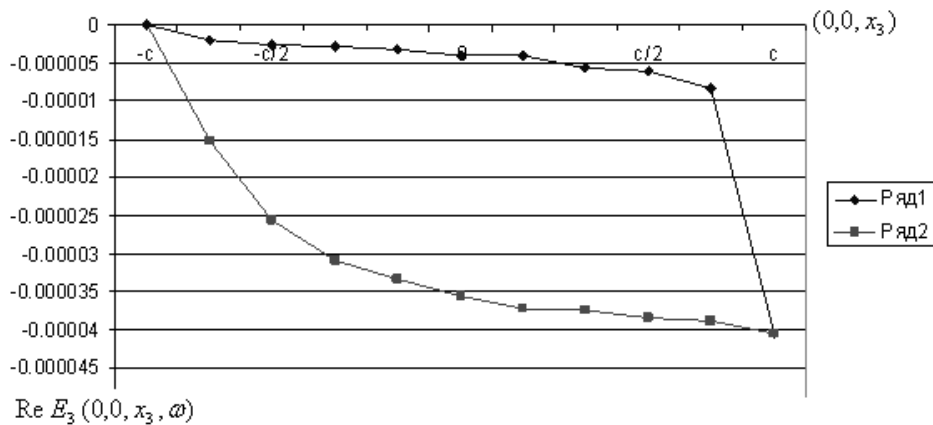


Рис. 8. Залежність дійсної компоненти ЕП від координат

Висновки

З погляду інтерпретаційного процесу пряма задача індукційних зондувань є дуже важливою і визначальною та підлягає постійній оптимізації. Варто зазначити, що інтерпретація даних, які стосуються горизонтально-неоднорідних середовищ, із застосуванням позірних параметрів одновимірних моделей (нехай навіть багат шарових) вже себе вичерпала. Сучасний інтерпретаційний процес вимагає застосування тривимірних математичних моделей. Гостроти ситуації додає й те, що тенденцією ЕМ досліджень є організація збирання й обробки все більших об'ємів даних, і, отже, всі елементи апарату інверсії, включаючи і елементи тривимірного характеру, повинні бути оперативні. Як відомо, обернену задачу розв'язують переважно на базі аналізу чутливості прямої задачі до змін геопараметрів і проводять його найчастіше шляхом багатократного використання процедури прямої задачі. Тому алгоритм прямої тривимірної задачі, який дозволяє в частотній області ефективно і з достатньою точністю обчислювати всі компоненти ЕМП й позірні величини, не вводячи потенціалів електричного чи магнітного типів, є в цій ситуації надзвичайно корисним.

З усього вищенаведеного випливає, що запропонована для розв'язування прямих тривимірних задач у провідному кусково-однорідному півпросторі методика є ефективною, а розвиток створеного на її базі програмного забезпечення може відбуватись із врахуванням таких факторів. Окрім розглянутого кусково-однорідного півпростору варто дослідити ще й шарувату модель з включеннями довільної форми, оскільки вона істотна для сучасних умов застосування індукційної

електророзвідки. Водночас для оперативнішої обробки великих масивів площинних даних під час розв'язування оберненої задачі геоелектрики слід одержати інтерпретаційні формули на базі амплітудних та фазових кривих позірнього опору.

Результати проведеного математичного моделювання свідчать про доцільність застосування індукційних електромагнітних методів геофізики для виявлення високопровідних включень типу рудних родовищ та високоомних включень типу нафтових і газових покладів. Загальними сприятливими умовами для виявлення і дослідження локальних неоднорідностей з провідністю, вищою й нижчою, ніж геосередовище, треба вважати співвимірну з горизонтальними розмірами (чи меншу) глибину їх залягання.

1. Жданов М.С. *Электроразведка*. – М.:Недра, 1986 – 316с.
2. Кауфман А.А. *Введение в теорию геофизических методов. Часть 2. Электромагнитные поля*. – М.:Недра, 2000. – 483с. – С. 190.
3. Newman G.A., and Alumbaugh, D.L. *Frequency-domain modeling of airborne electro-magnetic response of high-contrast prisms in layered earth: Geophys. Prosp.*, 1995, 43, 1021-1042.
4. Fomenko E.Y., and Mogi, T. *A new computation method for a staggered grid of 3D EM field conservative modeling: Earth, Planets and Space*, 2002, 54, 499-509.
5. Badea, E.A., Everett, M.E., Newman, G.A., and Biro, O. *Finite-element analysis of controlled-source electromagnetic induction using Coulomb-gauged potentials: Geophysics*, 2001, 66, 786-799.
6. Mitsuhashi, Y., and Uchida, T. *3D magnetotelluric modeling using the T- Ω finite-element method: Geophysics*, 2004, 69, 108-119.
7. Табаровский Л.А. *Применение метода интегральных уравнений в задачах геоэлектрики*. – Новосибирск: Наука, 1975. – 144 с.
8. Dmitriev, V.I., and Nesmeyanova, N.I. *Integral equation method in three-dimensional problems of low-frequency electrodynamics: Computat. Math. Modeling*, New-York: Plenum Pub. Corp., 1992, 3, 313-317.
9. Avdeev, D.B., Kuvshinov, A.V. Pankratov, O.V., and Newman, G.A. *Three-dimensional induction logging problems, Part I: An integral equation solution and model comparisons: Geophysics*, 67, 413-426.
10. Бенерджи П., Баттерфилд Р. *Метод граничных элементов в прикладных науках*. – М.:Мир, 1984. – 494с.
11. *Электроразведка: Справочник геофизика*. – М.:Недра, 1979. – 517с.

МЕТОД РЕКОНСТРУКЦІЇ ПОВЕРХНІ З ІНТЕРФЕРОГРАМИ БІЛОГО СВІТЛА НА ОСНОВІ ПЕРЕТВОРЕННЯ ГІЛЬБЕРТА

© Хома А., 2014

Опрацьовано теоретичні основи інтерферометрії білого світла та проаналізовано математичну модель інтерферограми. Розроблено метод визначення топології поверхні із фази сигналу інтерферограми за допомогою його квадратурної компоненти, одержаної на основі перетворення Гільберта. Досліджено роботу запропонованого алгоритму на похилій та сферичній поверхнях, а також встановлено та проаналізовано основні джерела та причини виникнення похибок під час їх реконструкції.

Ключові слова: реконструкція топології поверхні, інтерферометрія білого світла, інтерферометр Майкелсона, аналітичний сигнал, перетворення Гільберта.

In this paper theoretical fundamentals of white light interferometry and mathematical model of interferogram were analyzed. The method of topography reconstruction from phase of white-light interferogram based on amplitude demodulation using Hilbert transform was developed. The verification of introduced algorithm on tilted and sphere surface is presented and the main sources of reconstruction errors were analyzed.

Key words: surface topology reconstruction, white light interferometry, Michelson interferometer, analytic signal, Hilbert transform.

Вступ

Інтерферометрія білого світла є безконтактним методом вимірювання, поширеним для визначення механічних величин, як от геометричні розміри, позиція, топографія поверхні досліджуваного об'єкта. Вона дає змогу вимірювати в діапазоні від кількох сантиметрів до кількох мікрометрів. Перевагами інтерферометрії є висока швидкість сканування та велика область вимірювання.

Інтерферометрія білого світла в поєднанні з атомно-силовою мікроскопією та раманівською спектроскопією широко застосовується для вимірювання низки параметрів із нанометровою роздільною здатністю, зокрема геометричних розмірів, форми, товщини, шорсткості, хвилястості, кривизни. Явище інтерференції покладено в основу принципу роботи багатьох вимірвальних приладів, а саме мікроскопів, профілометрів, віброметрів, сенсорів тиску і переміщення. Основними сферами використання інтерферометрів білого світла є мікро- і нанотехнології, біомеханіка, хімія полімерів, напівпровідникова техніка та ін. В нанотехнологіях інтерферометрія білого світла застосовується для нанопозиціонування в 3D-координатних системах, нормування метрологічних характеристик матеріальних еталонів, що надалі використовують для калібрування атомних силових мікроскопів. Цікавою технологією підвищення твердості та механічної стійкості пластикових виробів є покриття їх наноалмазними плівками, для контролю якості яких можуть застосовуватися інтерферометри білого світла [1].

На цей час механічні та електронні вузли інтерферометрів забезпечують реєстрацію інтерферограм з високою точністю і відтворюваністю. Але достовірність відтворення топографії поверхні залежить не лише від точності вимірювання інтерферометричної картини, але й від алгоритмів, за допомогою яких здійснюють її аналіз та реконструкцію. При цьому вкрай важливо забезпечити стійкість алгоритмів до впливу шумів, оптичних нелінійностей та інших небажаних чинників. Іншою важливою вимогою до цих алгоритмів є можливість реконструкції поверхні без

ап'риорних даних про неї. Про актуальність завдань із розроблення таких алгоритмів свідчать численні публікації [2–4].

Оглядом літератури встановлено, що більшість сучасних методів для визначення топології поверхні на основі інтерферограми білого світла ґрунтуються на алгоритмах так званого фазового зсуву. Такий підхід принципово вимагає прецизійного механічного переміщення деяких вузлів інтерферометра (зазвичай опорного дзеркала), що, своєю чергою, ускладнює будову інтерферометра та збільшує часові затрати під час вимірювань. У багатьох випадках, наприклад, коли йдеться про побудову мініатюрних недорогих систем або про опрацювання даних у реальному часі, вказані недоліки є критичними.

Метою роботи є розроблення ефективного та точного алгоритму для визначення топографії поверхні з її інтерферограми, на основі мінімальної кількості інтерферометричних картин.

Теоретичні основи інтерферометрії білого світла

Інтерференція як явище накладання двох чи більше електромагнітних хвиль призводить до підсилення та послаблення результуючої (сумарної) хвилі в кожній конкретній точці простору. Основною умовою виникнення інтерференційної картини є когерентність хвиль, що накладаються. Оскільки когерентні хвилі повинні мати незмінну в часі різницю фаз, то на практиці їх вдається одержати розділенням світлового променя, сформованого одним і тим самим джерелом випромінювання. Наслідком когерентності є стаціонарність в часі інтерферометричної картини, що забезпечує однозначну залежність її характеру від параметрів досліджуваного об'єкта. Слід також зазначити, що інтерференція може спостерігатися як для монохроматичного (лазерне випромінювання), так і білого (галогенні лампи) світла.

Вимірювальний прилад, принцип роботи якого ґрунтується на явищі інтерференції, називається інтерферометром. Одним із найпоширеніших видів інтерферометра є інтерферометр Майкелсона. Будову цього приладу зображено на рис. 1. Основними структурними елементами інтерферометра є джерело світла Q , розділювач світла ST , два дзеркала $S1$ і $S2$ та екран. Також для покращення якості інтерферометричного зображення застосовується низка додаткових елементів, як от поляризатори, лінзи, коліматори, світловоди тощо. Оскільки їх роль у процесі формування інтерферограми не є принциповою, то для спрощення подальшого аналізу їх вплив не враховуватиметься.

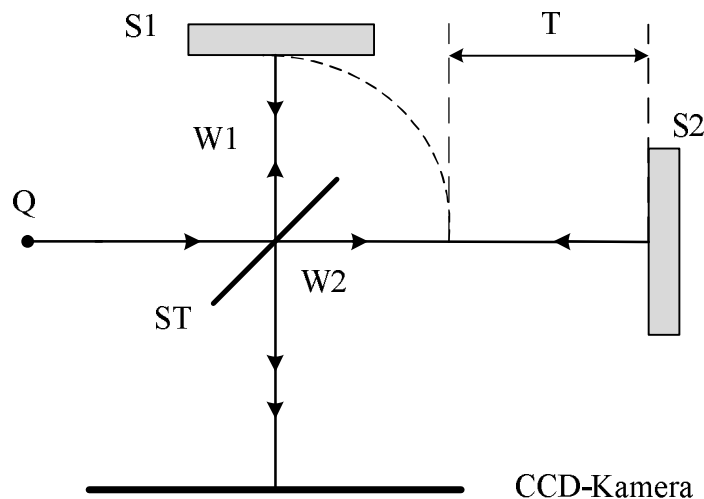


Рис. 1. Оптична схема інтерферометра Майкелсона [5]

Інтерферометр Майкелсона працює так: хвиля із джерела світла поділяється розділювачем на дві хвилі $W1$ і $W2$, що спрямовуються до відповідних дзеркал $S1$ і $S2$. Хвилі відбиваються від дзеркал і повертаються до розділювача хвилі ST , де вони накладаються. Результат накладання хвилі спостерігається на екрані та реєструється CCD-камерою. Явище підсилення і послаблення результуючої хвилі спостерігається через оптичну різницю фаз T . Потрібно зазначити, що максимальна інтенсивність інтерферограми (світлі ділянки) спостерігається в точках, де оптична

різниця фаз T дорівнює нулю і, навпаки, темні ділянки інтерферограми відповідають точкам, коли хвилі знаходяться у протифазі, компенсуючи одна одну.

На рис. 2 представлені інтерферограми білого світла для сферичної поверхні, отримані для джерел світла з різними параметрами. Слід зауважити, що вигляд інтерферограми залежить від спектральної ширини і центральної довжини хвилі. Інтерферограма справа містить кольорові кільця різної товщини, які швидше згасають у міру віддалення від центра, оскільки одержана за використання джерела світла із ширшою спектральною густиною, розташованою в центрі видимого діапазону. Інтерферограму зліва одержано за використання джерела світла із вузкою спектральною густиною, розташованою на краю видимого діапазону, тому ширина інтерференційних кілець змінюється незначно, краще зберігається їх контрастність у міру віддалення від центра, а сама інтерференційна картина більш монохроматична.

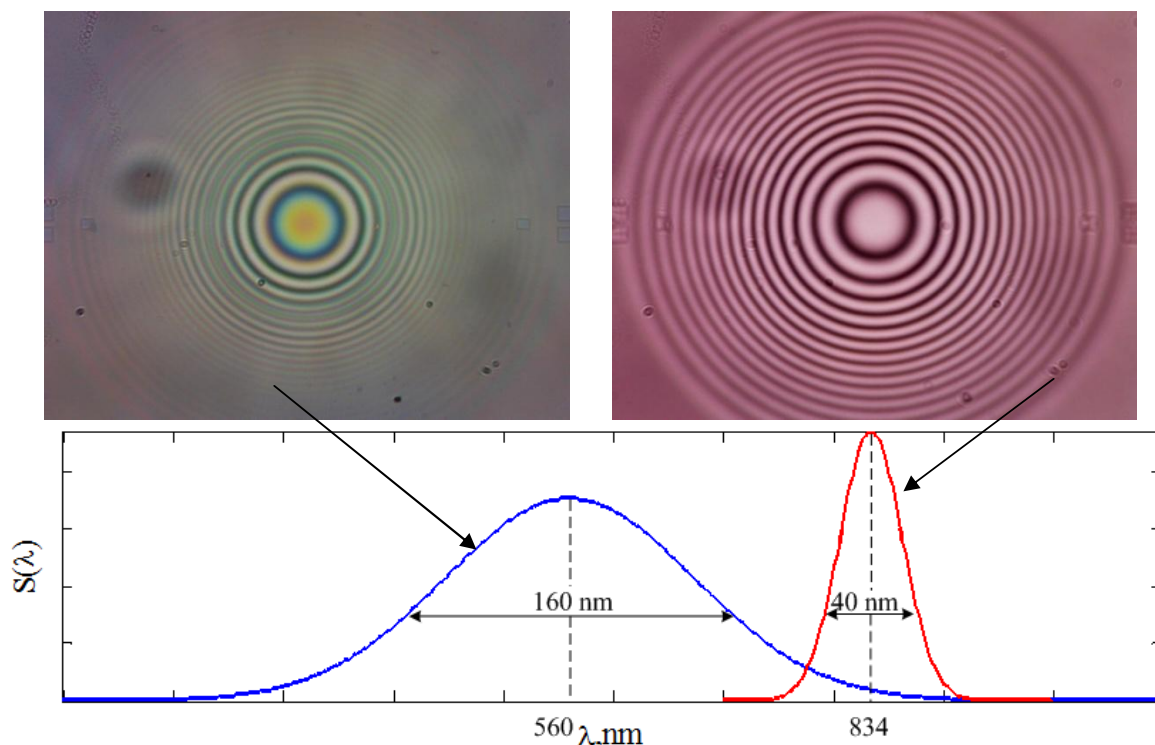


Рис. 2. Реальні інтерферограми сферичної поверхні для різних джерел світла

Математична модель інтерферограми білого світла описується такими формулами [5,6]:

$$I(T) = I_0 + E(T) \cdot C(T); \quad (1a)$$

$$E(T) = I_M \cdot \exp\left(-\frac{4 \cdot \Delta I^2 \cdot T^2}{I_0^4}\right) \text{ (огинаюча);} \quad (1б)$$

$$C(T) = \cos\left(\frac{4 \cdot p}{I_0} \cdot T\right) \text{ (несуча),} \quad (1в)$$

де I_0 і I_M – відповідно стала складова та амплітуда огинаючої сигналу інтенсивності інтерферограми; T – оптична різниця фаз; λ_0 і $\Delta\lambda$ – центральна довжина хвилі та ширина спектральної густини джерела світла.

Як впливає із поданих рівнянь, форма несучої сигналу інтерферограми визначається двома параметрами: оптичною різницею фаз T та центральною довжиною хвилі λ_0 . На характер огинаючої, окрім цих двох параметрів, значною мірою впливає також ширина спектра $\Delta\lambda$. Підтвердженням цього впливу може бути рис. 2. Така математична модель справедлива і для інтерферограми монохроматичного світла, оскільки в цьому випадку ширина спектра $\Delta\lambda$ прямує до нуля і, як наслідок, відсутня амплітудна модуляція.

Розроблення концепції методу реконструкції топології поверхні

Проаналізувавши рівняння (1), можна зробити висновок, що шукана величина T є аргументом одночасно як функції огинаючої $E(T)$, так і несучої $C(T)$. Однак аналізом фази несучої можна досягти вищих чутливості та точності, ніж аналізуючи амплітуду огинаючої. Крім того, частота чи фаза як інформативний параметр порівняно з амплітудою є стійкішою до впливу завад. Отже, доцільно розробляти методи для визначення топології поверхні із фази несучої інтерферограми.

З погляду теорії сигналів інтерферограму можна трактувати як сигнал із амплітудно-фазовою модуляцією, в якому інформативним параметром є оптична різниця фаз T , а коефіцієнтами модуляції є параметри джерела світла. Очевидно, що для поставленого завдання (визначення фази несучої) необхідно насамперед усунути вплив амплітудної модуляції, яка описується рівнянням несучої. Для цього можна застосувати аналітичне представлення сигналу, яке дає змогу поширити поняття амплітуди і фази на негармонічні сигнали [7].

Аналітичним сигналом називається комплексний сигнал, дійсна складова якого є аналізованим сигналом, а уявна – його ортогональним доповненням:

$$X(T) = E(T)\cos(w \cdot T) + Q(T)\sin(w \cdot T), \quad (2)$$

де $E(T)$ та $Q(\omega T)$ – амплітуди синфазної та квадратурної складових; $\omega = 4\pi / \lambda_0$.

Якщо під час опрацювання інтерферометричного сигналу спершу виділити синфазну $s(T) = E(T)\cos(w \cdot T)$ та квадратурну $s_{\perp}(T) = Q(T)\sin(w \cdot T)$ складові, а потім поділити їх одна на одну, то в результаті отримаємо сигнал, пропорційний тангенсу фази несучої:

$$\frac{s_{\perp}(T)}{s(T)} = \frac{Q(T)\sin(w \cdot T)}{E(T)\cos(w \cdot T)} = \frac{Q(T)}{E(T)} \cdot \operatorname{tg}(w \cdot T). \quad (3)$$

У випадку $E(T) = Q(T)$ рівняння набуде вигляду:

$$\frac{s_{\perp}(T)}{s(T)} = \operatorname{tg}(w \cdot T), \quad (4)$$

а фазу можна визначити, застосовуючи функцію арктангенса

$$f = \operatorname{arctg} \frac{s_{\perp}(T)}{s(T)}. \quad (5)$$

Із зіставлення виразів (1) і (2) стає очевидним, що сигнал інтерферограми $I(T)$ можна розглядати як синфазну складову $s(T)$, але для цього необхідно попередньо вилучити адитивне зміщення I_0 . Отже, щоб одержати вираз (5), необхідно сформувати відповідний квадратурний сигнал (з однаковою амплітудою і зсунутою фазою на 90°), після чого вимірювальну величину T можна обчислити за такою формулою:

$$T = \frac{I_0}{4 \cdot p} \cdot \operatorname{arctg} \frac{I_{\perp}(T)}{I(T) - I_0}, \quad (6)$$

де $I_{\perp}(T)$ – квадратурний сигнал, отриманий із сигналу інтерферограми після вилучення сталої складової.

По суті практична реалізація цього методу зводиться до задачі формування квадратурного сигналу інтенсивності інтерферограми. Графічну інтерпретацію запропонованого методу наведено на рис. 3.

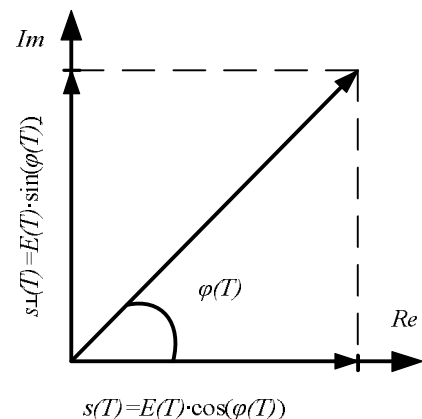


Рис. 3. Графічне представлення фазової демодуляції за допомогою квадратурного сигналу

Теоретично найточніший метод визначення квадратурного сигналу інтерферограми без використання апріорних даних про тип поверхні ґрунтується на перетворенні Гільберта [8]. Перетворення Гільберта для деякого дійсного сигналу $s(T)$ забезпечує формування його ненульового “квадратурного еквівалента” $s_{\perp}(T)$, який задовольняє таку умову:

$$\int_{-\infty}^{\infty} s(T) \cdot s_{\perp}(T) dT = 0. \quad (7a)$$

Аналітичний вираз перетворення Гільберта має вигляд [7]:

$$s_{\perp}(T) = \frac{1}{p} \int_{-\infty}^{\infty} \frac{s(t)}{T-t} dt, \quad (7b)$$

де $s(\tau)$ і $s_{\perp}(T)$ – аналізований вхідний сигнал та його квадратурна компонента (вихідний сигнал); τ і T – аргумент (незалежна змінна) для вхідного і вихідного сигналів.

Наведений вираз можна трактувати як згортку сигналу $s(T)$ із функцією $h(T) = 1/\pi T$, яка, по суті, є ядром гільберт-перетворення. Отже, перетворення Гільберта формує ортогональне (квадратурне) доповнення вхідного сигналу, здійснюючи фазовий зсув на 90° кожної гармоніки, наявної у спектрі аналізованого дійсного сигналу $s(T)$.

У дискретному вигляді перетворення Гільберта описується виразом:

$$s_{\perp}(m) = \frac{1}{p} \sum_{n=1}^{N-1} \frac{s(n)}{n-m}, \quad (8)$$

де $s(n)$ і $s_{\perp}(m)$ – вибірки вхідного та вихідного квадратурного сигналів; n і m – аргументи вхідного та вихідного сигналів.

Тобто, алгоритм реконструкції поверхні за методом фазової демодуляції складається із таких пунктів:

- завантаження або симуляція даних інтерферограми і параметрів інтерферометра;
- вилучення сталої складової I_0 ;
- обчислення квадратурного сигналу інтерферограми за допомогою перетворення Гільберта;
- демодуляція інтерферограми з подальшим обчисленням фази;
- візуалізація реконструйованої поверхні.

Очікуваною перевагою такого алгоритму є його незалежність від форми поверхні досліджуваного об’єкта, типу джерела світла, а також від реальних числових значень параметрів математичної моделі. Крім того, запропонований алгоритм володіє потенційно високою швидкістю через відсутність рекурсивних або ітераційних обчислень. Наступним етапом після розроблення концепції є його практична імплементація на конкретній платформі з подальшою верифікацією та аналізом результатів.

Верифікація розробленого методу та аналіз результатів

Реалізують та верифікують алгоритму у програмному середовищі MATLAB, яке відзначається зручним інтерфейсом користувача, гнучкою скриптовою мовою та численними бібліотеками для опрацювання і візуалізації даних. З метою усунення впливу різного роду потенційних дестабілізуючих факторів (шуми, завади, оптичні спотворення тощо) під час аналізу роботи алгоритму використовуватимуться дані, одержані за допомогою імітаційного моделювання. Для досліджень вибрано два типи поверхонь: похилу (рис. 4, а) та сферичну (рис. 4, б).

Відповідні інтерферограми (рис. 5) було отримано за формулами (1), при цьому вибрано такі параметри джерела світла та CCD-камери:

- стала складова сигналу – $I_0 = 2$;
- модуляційна амплітуда – $I_M = 2$;
- центральна довжина хвилі – $\lambda_0 = 620$ нм;
- відхилення довжин хвиль (ширина спектра) – $\Delta\lambda = 62$ нм;
- кількість даних по осях x та y – 800×800 точок;
- оптична різниця фаз у центрі інтерферограми дорівнює нулю.

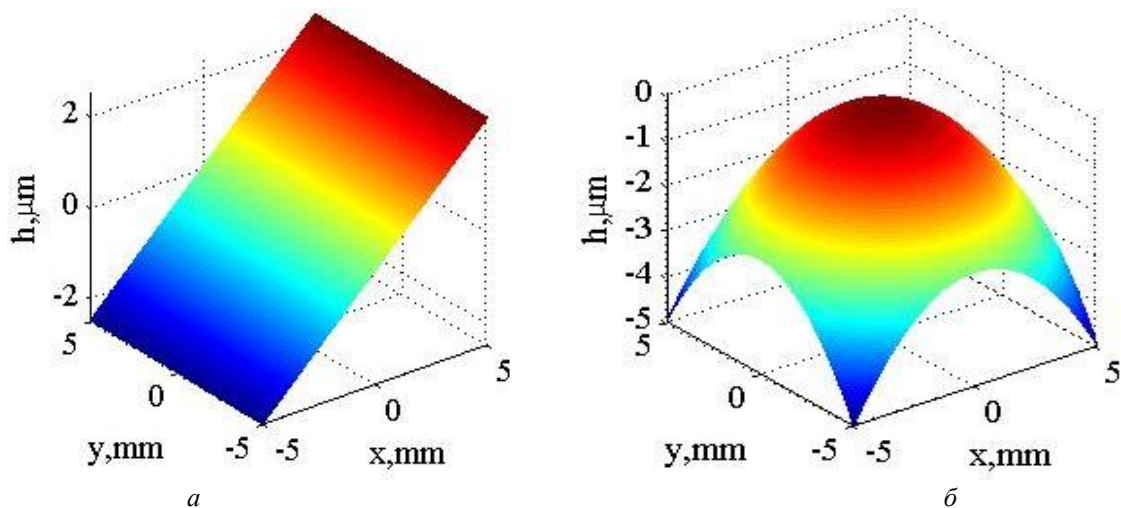


Рис. 4. Вигляд модельованих поверхонь: а – похила, б – сферична

З метою спрощення спочатку розроблення алгоритмів стосуватиметься лише одновимірних даних, тобто однієї лінії інтерферограми. Як зазначалося вище, для правильного визначення квадратурного сигналу необхідно спочатку вибрати сталу складову. Для цього було використано функцію MATLAB *mean(*)* [9]. Слід зазначити, що перетворення Гільберта автоматично усуває сталу складову, тому цю операцію можна виконати як до, так і після формування квадратурного сигналу.

У MATLAB дискретне перетворення Гільберта реалізується вбудованою функцією *hilbert(*)*. Особливістю цієї функції є те, що вона повертає не лише квадратурну $s_{\perp}(T)$, але і дійсну $s(T)$ складові:

$$s(m) + js_{\perp}(m) = \text{hilbert}[s(n)]. \quad (9)$$

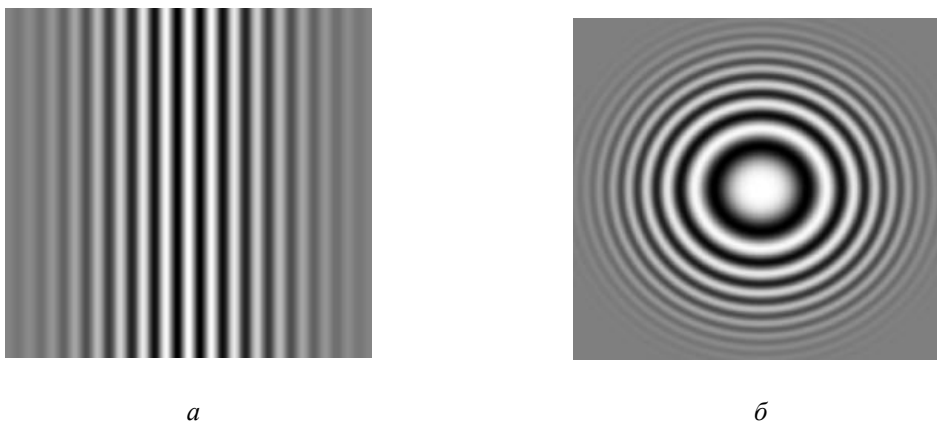


Рис. 5. Вигляд інтерферограм для модельованих поверхонь: а – похила, б – сферична

Тому для визначення фази сигналу за формулою (5) насамперед потрібно виділити дійсну та уявну частини. Це завдання можна реалізувати, використавши відповідно функції *real(*)* та *imag(*)*.

Сигнал, одержаний після обчислення фази інтерферограми, характеризується низкою стрибкоподібних переходів, зумовлених неврахуванням періодичності функції арктангенс. Відтак, наступним етапом є усунення цих розривів за допомогою спеціальної процедури – так званого розгортання фази (англ. – *phase unwrapping*). На практиці розгортання фази можна зреалізувати за допомогою вбудованої функції *unwrap(*)* у пакеті MATLAB [9]. Ця функція працює так: якщо різниця між двома сусідніми значеннями більша $\pm\pi$, то значення доповнюється на $\pm 2\pi$.

Після кроку розгортання фази можна обчислювати топологію поверхні за виразом (6). Однак слід зазначити, що між оригінальною та реконструйованою поверхнями спостерігається систематичне зміщення. У випадку оригінальної поверхні система координат по осі y має абсолютний характер, оскільки нульовий пункт ($T=0$) відповідає такому положенню поверхні, коли різниця

оптичного ходу променів дорівнює нулеві. Для реконструйованої поверхні система координат має відносний характер, оскільки функція арктангенс завжди повертає величини в межах першого періоду $[-\pi; +\pi]$ без врахування члену $2\pi \times k$, тобто інформація про абсолютне значення періоду втрачається. Оскільки згладнення за допомогою функції *unwrap(*)* відбувається зліва направо, то крайні ліві точки завжди лежатимуть у межах першого періоду. Цей ефект не є критичним, оскільки при визначенні топології поверхні важливим є відносне, а не абсолютне положення точок поверхні. Разом з тим це явище можна усунути, аналізуючи функцію огинаючої. Точка, де огинаюча сягає максимуму (точка симетрії інтерферограми), збігається з істинним початком координат (оптична різниця ходу дорівнює нулеві), а значення реконструйованої поверхні в цій точці по суті і є сталим зміщенням. Для усунення систематичного зсуву необхідно відняти від кожної точки реконструйованої поверхні значення в точці, де огинаюча максимальна:

$$h_{corrected}(x, y) = h_{reconstructed}(x, y) - h(x_{Ienv,max}, y_{Ienv,max}). \quad (10)$$

Результати використання методу фазової демодуляції інтерферограми для реконструкції топології поверхні за описаним вище алгоритмом наведено на рис. 6. Як видно з наведених графіків, у випадку похилої площини реконструйована та оригінальна поверхні ідентичні, а для сферичної спостерігаються суттєві відхилення, особливо на краях. Можливою причиною є те, що несуча у випадку сферичної поверхні має явно виражений несинусоїдальний характер. Відтак необхідно провести додаткові дослідження стосовно правомірності застосування формули (5) до сигналів такого типу. Також важливо зауважити, що, застосовуючи метод фазової демодуляції для реконструкції поверхні, бажано використовувати ціле число періодів інтерферограми, оскільки за таких умов результати роботи алгоритму є найкращими.

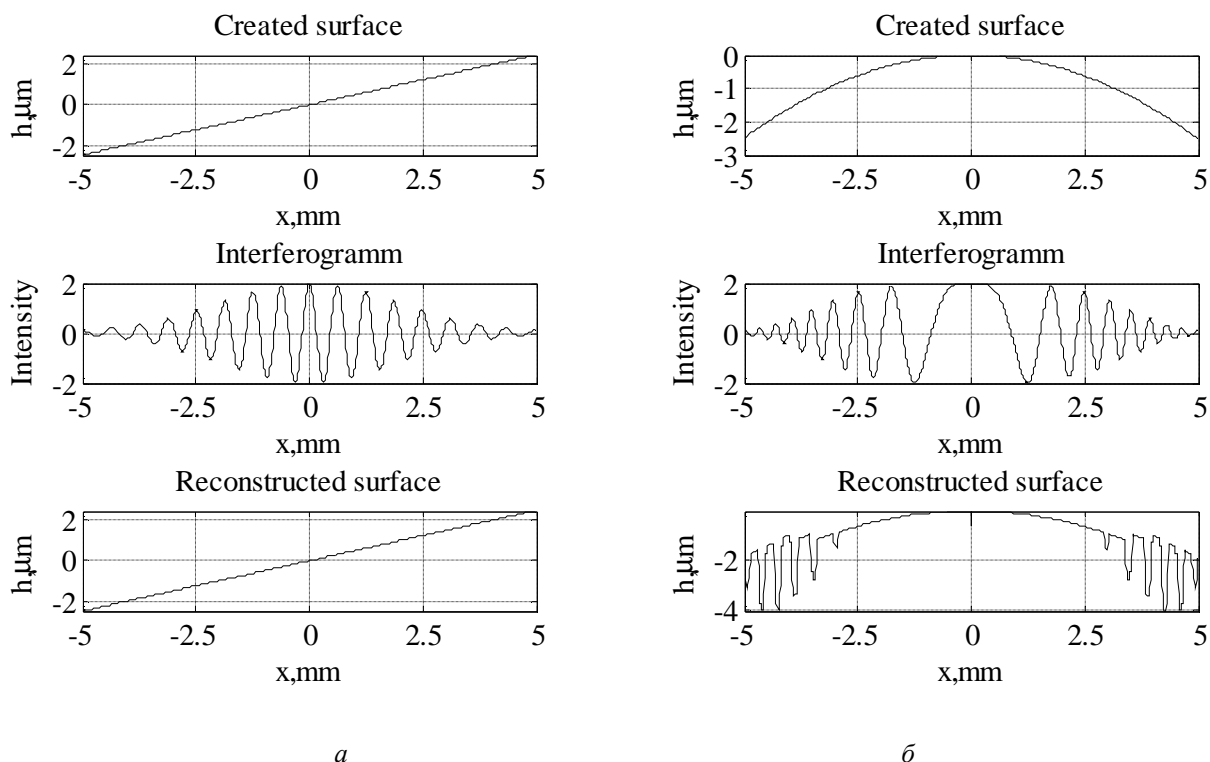


Рис. 6. Результати використання методу квадратурної демодуляції:
а – похила, б – сферична поверхня

Висновки

Розглянуто можливість реконструкції топології поверхні із інтерферограми білого світла за допомогою квадратурного сигналу. Наведено основні теоретичні положення, зокрема будову та принцип роботи інтерферометра білого світла, залежність характеру та форми інтерферограми від параметрів джерела випромінювання. Також проаналізовано математичну модель інтерферограми.

Запропоновано метод визначення топології поверхні, що ґрунтується на фазовій демодуляції за допомогою квадратурного сигналу, одержаного на основі перетворення Гільберта. Здійснено верифікацію на симульованих даних для двох випадків: лінійної та сферичної поверхонь. Одержані результати свідчать, що запропонований метод ефективний лише для лінійних поверхонь. Крім того, існує низка інших обмежень, як от необхідність визначення і усунення сталої складової, використання цілого числа періодів, тощо. Відтак доцільними є подальші дослідження, спрямовані на пошук нових підходів для аналізу сферичних та інших нелінійних поверхонь.

1. Ajay Kumar etc. *Formation of nanodiamonds at near-ambient conditions via microplasma dissociation of ethanolvapour* // *Nature Communications*, 4, 2013. 2. Jirí Novák. *Computer analysis of interference fields using MATLAB* // In: *MATLAB 2002*. Prague : HUMUSOFT, 2002, p. 406-410. 3. Kieran G. Larkin. *Efficient nonlinear algorithm for envelope detection in white light interferometry* // *J. Opt. Soc. Am. A*/ Vol. 13, No. 4/ April 1996, 832-843. 4. Masaaki Adachi. *Phase-Shift Algorithm for White-Light Interferometry. Insensitive to Linear Errors in Phase Shift* // *Kakuma-machi, Kanazawa*, 920-1192, Japan. 5. Кальянов А.Л. и др. *Оптическая низкокогерентная интерферометрия и томография*. Саратов, 2009, 86 с.. 6. Thomas Seiffert: *Schnelle Signalvorverarbeitung in der Weißlichtinterferometrie durch nichtlineare Signalaufnahme* – Stuttgart 2004. 7. А.Б.Сергиенко "Цифровая обработка сигналов" 2-е изд.: Пупер,2006. – 751с. 8. Sang, Wolfram: *Entwicklung und Implementierung eines Verfahrens zur Auswertung von Weißlichtinterferogrammen zur Bestimmung der dreidimensionalen Oberflächentopographie von Mikro- und Nanostrukturen als Anwendung für eine Nanopositionier- und -messmaschine*. Diplomarbeit –Technische Universität Ilmenau,2006. 9. <http://www.mathworks.de/products/matlab/>.

COMPUTER SIMULATIONS OF A STOCHASTIC MODEL FOR THE NON-IMMUNE DISEASE SPREAD

© Ilnytskyi Ja., Holovatch Yu., Kozitsky Yu., Ilnytskyi H., 2014

Наведено модель поширення неімунного захворювання, а також алгоритмічний підхід і відповідні результати свого дослідження за допомогою комп'ютерного моделювання. Модель являє собою узагальнення моделі SIS з рівномірним двовимірним просторовим розподілом особин, що перебувають в еволюції типу Маркова з дискретним часом. Описано авторський підхід і наведено попередні результати, отримані для випадку відомого розподілу особин на сторонах простої квадратної решітки та однопараметричну стохастичу динаміку (синхронно модель SIS на квадратній решітці з різною кількістю сусідів).

Ключові слова: епідеміологія, комп'ютерне моделювання, SIS модель, критичні явища.

We present a model of the non-immune disease spread, as well as an algorithmic approach and the corresponding results of its study by computer simulations. The model is a generalization of the SIS model with the uniform two-dimensional spatial distribution of individuals undergoing a Markov-type evolution with discrete time. In this work, we describe our approach and present a number of the preliminary results obtained for the case of the quenched distribution of individuals on the sites of a simple square lattice and the one-parameter stochastic dynamics (synchronous SIS model on a square lattice with varying number of neighbors). The dynamical properties of the model are studied in terms of the behavior of the fraction of the infected individuals as well as of the maximum size and dimension of the largest cluster formed by them. These properties are found to be affected by the effective range of the local infectivity, which demonstrates the role of the underlying graph of the individual communications on the global disease spread. The presented approach allows for numerous extensions, including the possibility to consider non-homogeneous spatial distributions and various forms of the stochastic dynamics.

Key words: epidemiology, computer simulations, SIS model, critical phenomena

1. Introduction

Since the seminal work by Kermack and McKendrick [1], modeling the disease spread had found much interest in both medical and applied mathematics circles (see, e.g. [2-3]). There are many ways the epidemiology models can be classified, therefore here we will restrict ourselves to discussing the features relevant to the purpose of this study. In deterministic models, the population is split into several large groups, the principal ones would be: fraction of susceptible to illness (S) and fraction of infected (I). If the individual develops an immunity to the disease after being cured (e.g. smallpox), it will move to the group of removed with their fraction R . This model is commonly abbreviated as SIR model. In the case of non-immune illness, cured individuals return back to S group, thus the SIS model. The definition of the model

contains a set of transfer rules for the individuals to move between the groups. Diagrams for *SIS* and *SIR* models are depicted in Fig. 1. As one can see, newborn individuals are joining group *S* with the rate π , whereas dying ones leave groups *S* and *I* with the rates μ and μ_I , respectively, β and γ are infecting and curing rates, respectively, reflecting the rate at which individuals move to and from group *I*. Note that the probability of the individual to be infected is proportional to the product of currently available susceptible and infected individuals SI . Normalization conditions $S + I = 1$ and $S + I + R = 1$ hold for *SIS* and *SIR* models, respectively.

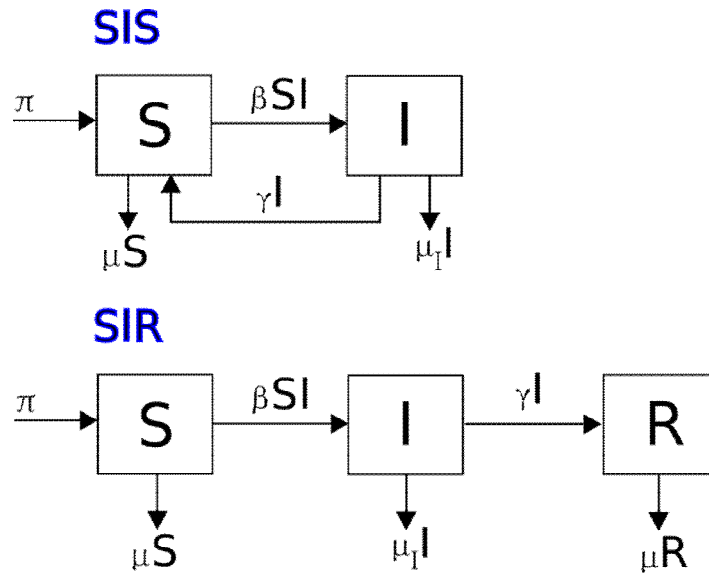


Fig.1. *SIS* and *SIR* deterministic models definitions. *S*, *I* and *R* are fractions of susceptible, infected and removed, respectively, flow rules are shown via arrows; π is birth rate, μ and μ_I are natural death rate and death rate for infected, respectively; β and γ are infecting and curing rates, respectively. $S + I = 1$ for *SIS* model and $S + I + R = 1$ for *SIR* model.

For the *SIS* model, the flow diagram of Fig. 1 is conveniently written as a set of first-order differential equations:

$$\begin{aligned} \frac{dS}{dt} &= \pi - \beta SI + \gamma I - \mu S \\ \frac{dI}{dt} &= \beta SI - \gamma I - \mu_I I \end{aligned} \tag{1}$$

As mentioned by Britton [5], “the type of questions that were addressed were for example: is it possible that there is a big outbreak infecting a positive fraction of the community? How many will get infected if the epidemic takes off? What are the effects of vaccinating a given community fraction prior to the arrival of the disease? What is the endemic level?” The ratio $R_0 = \beta/\gamma$ is of main importance here and it is usually referred to as the basic reproduction number for the infection. When $R_0 > 1$ the epidemic takes off and when $R_0 < 1$ there is no epidemic. The set of differential equations (1) can also be used to obtain a balance equation for the final state at $t \rightarrow \infty$. As these problems were resolved, the simple models were generalised in several ways towards making them more realistic. The examples of such generalisations are: introducing of an additional group of latently infected (*E*) for the case of diseases with long incubation period (e.g. tuberculosis [3]), splitting group *S* into several subgroups (e.g. low-risk and high-risk subgroups [3]), etc. Such extensions enable more flexibility for the simple deterministic models by tuning them to a particular type of the disease and particular environment of its spread.

However, the results obtained via employment of simple deterministic models, rely on the homogeneity of a community and on an assumption that individuals mix uniformly with each other. In this

way, these models are of the mean-field type, using the terminology from the phase transitions [4]. Hence, another way of generalization of simple deterministic model is to study stochastic epidemic models [5]. These have also shown to be advantageous when the contact structure in the community contains small complete graphs: e.g. households and other local social networks. The details of an underlying network of human contacts, as shown by many examples in rapidly developed theory of complex networks [7], may bring the description of the disease spread into another level, as compared to structureless deterministic models. In this way we arrive to stochastic epidemic models which contain certain graph for human communication and stochastic dynamics for the infecting spread and cure defined via certain algorithm. Each individual in such model is treated as a single entity with several possible states (e.g. susceptible and infected) and connected to other individuals by means of a certain graph. The latter can be quenched or dynamically changed by means of bonds creation/destruction, individual jumps, etc. The infection spreads in a stochastic way, the same hold for curing of infected individuals. The dynamics is described as the Markovian process. In this way, the link to the simple deterministic model is provided via evaluation of S , I and R , the total number of susceptible, infected and removed individuals, respectively. However, on a top of that, the spatial arrangement of clusters with infected individuals, distribution of their sizes and other local properties can also be studied. The influence of a local and global graph structure, as well as of possible individuals movement can be studied as well. In this study we consider computer simulation of one particular realization of a stochastic epidemic model, namely: SIS model on a two-dimensional square lattice. The details of the model, algorithmic solutions and results are presented in section 2. Conclusions are provided at the end of the study.

2. Model, algorithms and results

We consider SIS model on a two-dimensional square lattice of linear size L , hence total number of sites is $N = L^2$. Each i -th site contains one individual that is characterized by its state $s(i) = 0,1$, where 0 stays for susceptible and 1 for infected. Individuals are assumed to be immobilized and can only change their state reacting on the state of their neighbours. The stochastic dynamics of the system is defined *via* the rules for changing the state of the individual i :

- if $s(i) = 0$ then it can switch to $s(i) = 1$ with the probability defined locally as $\beta n_I/q$, where n_I is the number of infected neighbours for site i , and q is the total number of its neighbours (coordination number of a lattice)
- if $s(i) = 1$ then it can switch to $s(i) = 0$ with constant probability γ

A link to simple deterministic SIS model, Eq.(1) can be established now. First of all, the infecting rate is proportional to β in both cases, but in the case of stochastic model it is a local quantity, being dependant on a number of infected neighbours n_I . The average value for n_I/q , however, will be equal to βI if one assumes that all individuals communicate one with another, providing a link to the deterministic model. Curing rate is constant in both models and equal to γ . In this study we will restrict ourselves to the case of synchronous SIS model, when $\beta=1-\gamma$. Its behaviour is governed by a single parameter, e.g. curing rate γ or $\lambda = \beta/\gamma = (1-\gamma)/\gamma$ can be used equally. The latter parameter λ has a meaning of local reproduction ratio for the infection. The model is already considered for the case of the nearest neighbors only ($q = 4$ for the case of square lattice): it is called the contact process [8-10]. It exhibits phase transition with the critical behavior similar to that of the directed percolation [8-9].

In this study we consider a family of lattice stochastic SIS models with the condition $\beta=1-\gamma$ considered on a square lattice with various coordination number q . The latter is defined via a neighbours sphere of certain radius R_m . In particular, at $R_m = 1$, one obtains the nearest neighbours model with $q = 4$, at $R_m = \sqrt{2}$ one has $q = 8$ and so on. We focus on critical behaviour of such models, which is governed by a local curing rate γ (λ being defined above can also be used). Several aspects of the critical behavior [4] are of the interest for us. First of all it is the influence of a finite system size L . Then, we are studying the

effect of an increase of infection range, via variation of the R_{mn} value and, hence, the coordination number q . This effectively changes the connectivity of the underlying graph, on which the individuals reside. One expects to reproduce the critical behaviour of simple deterministic model in the limit of $R_{mn} \rightarrow \infty$. Finally, we would exploit the benefits of stochastic model by analysing the average and maximum cluster sizes of infected individuals and study them on a subject of percolation.

The workflow of the simulation algorithm is as follows:

1. The initial configuration is formed by creation of the simulation box of linear size L with $N = L^2$ individuals with their IDs numbered sequentially from 1 to N . The individuals are confined to the sites of a square lattice with lattice spacing equal to 1. Each site acquires initial status 1 with the probability p and status 0 with the probability $1 - p$, in most cases the value $p = 0,1$ was used.

2. The list of neighbors has been built for each individual according to the radius R_{mn} or the neighbors sphere and taking into account periodic boundary conditions [11]. Aiming on future extension of our study to the off-lattice models (with the continuous coordinates space) we used the general linked cells list algorithm used in molecular dynamics simulations of atomic systems [11,12]. To this end the simulation box is split into square domains of linear size larger or equal to R_{mn} and the linked lists are built for each domain out of individuals IDs falling into this particular domain. Then, the pairs of adjacent domains are analyzed for the neighbors within the distance of R_{mn} and the neighbour lists are saved. Computing resources required to build neighbour lists using this algorithm scale as N at large N , in contrary to $N(N-1)/2$ for simple loop over all pairs in the simulation box.

3. The infecting/curing sweep over the system is performed. The sweep contains N attempts to pick an individual randomly. If the individual is infected than it is cured with the probability γ and it transmits the infection to one of its neighbours with the probability $1 - \gamma$. To do so one generates a random number θ evenly distributed within an interval $[0,1]$. If $\theta < \gamma$ then the status of selected individual changes from 1 to 0 (curing takes place), otherwise the status of one of its randomly selected neighbours changes from 0 to 1 (infecting of a neighbour takes place).

4. The clusters of infected individuals are identified using efficient Hoshen-Kopelman algorithm [13]. Initially each infected individual is assigned a cluster label that coincides with its ID. The sweeps over the neighbors pairs using the linked cells lists (introduced in step 2) are performed and cluster labels reduction takes place for the neighboring individuals, the process continue until no reductions occur. At the end each individual is assigned the label that identifies its host cluster.

5. The properties of interest are evaluated such as fraction of infected individuals, average and maximum cluster sizes, search for percolating clusters. The accumulators for average properties are updated.

6. Go back to step 3 until desired number of sweeps is performed.

First we consider the case of $q = 4$ model (square lattice with disease spread over nearest neighbours only). For this model we performed simulations for various linear system sizes ranging from $L = 16$ to 512 with the periodic boundary conditions being applied. For each system size the interval of values for curing rate from $\gamma = 0,005$ to $\gamma = 0,7$ is covered with the step of 0,005, and for each pair $\{L, \gamma\}$ 5000 sweeps were performed. The system was found in a stationary state with saturated value of I pretty fast, typically after up to 100 sweeps. The evolution of I during first 100 sweeps is illustrated in Fig.2 for two system sizes, $L = 32$ and 256, and a few characteristic values of the curing rate γ . The time unit is set equal to one sweep over the system.

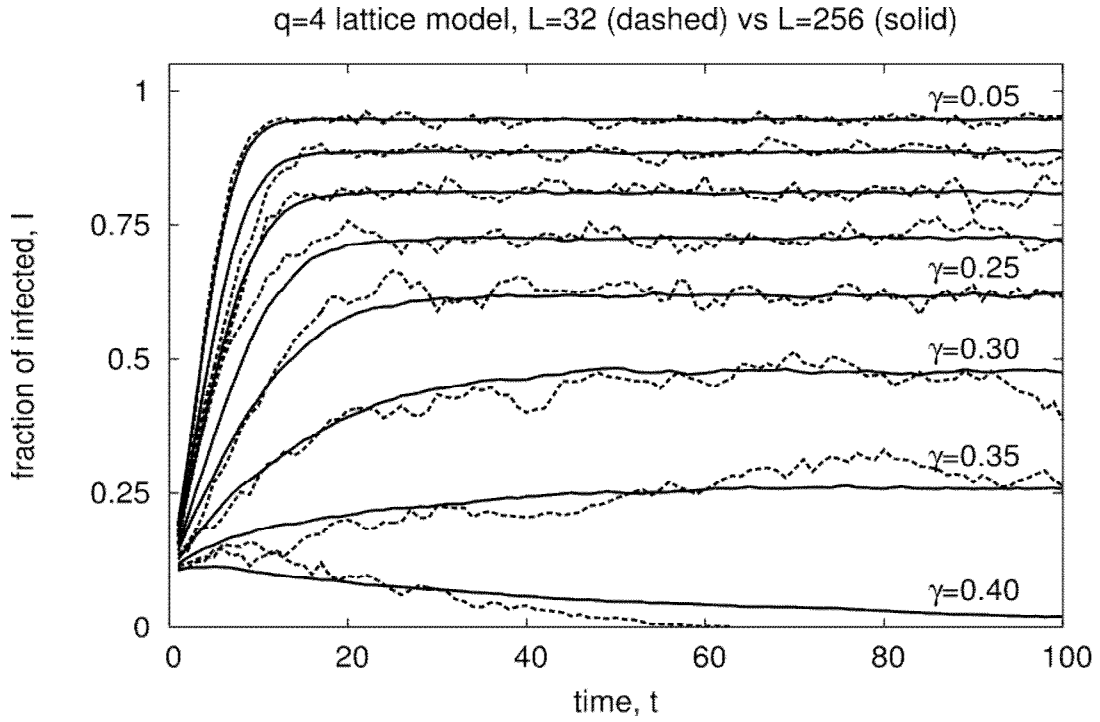


Fig.2. Evolution of the fraction of infected individuals I towards a stationary state for the $q = 4$ model of system sizes $L = 32$ (shown as dashed lines) and $L = 256$ (solid lines) at various curing rate γ .

One can see an essential increase of the time needed to reach stationary state when curing rate approaches the interval of values $\gamma = 0,3 - 0,4$. However, there is no evidence for the system size dependence, as the curves built for $L = 32$ and $L = 256$ at the same γ exhibit essentially the same shape with only more fluctuations being evident for the smaller system size of the two. Obviously, small values of curing rate γ lead to the $I = 1$ stationary state (all are infected) whereas with an increase of γ one observes that fraction of infected individuals vanishes in time starting from the ‘critical’ value of $\gamma \approx 0,4$. This is further outlined in Fig. 3 as described below.

Let us examine the fraction of infected individuals in a stationary state I vs curing rate γ for the same model. The data are shown in Fig.3 and there is a clear evidence for a continuous transition from the state with $I \approx 0$ and $I > 0$ that occurs at certain critical value γ_c (see, e.g. [9]). Therefore the data can be fitted to the following power form:

$$I = A(\gamma_c - \gamma)\beta'$$

where β' is the critical exponent for the order parameter [4], which in this case is I . Fitting of data is performed via the built-in routine of plotting program *gnuplot* that uses the least-squares method. The results of the fit are displayed in Fig.3. First of all let us remark that, again, we found no system size dependence as all the data obtained at various system sizes fit to practically the same master curve. The value for critical curing rate $\gamma_c \approx 0,3762$ is very close to the estimate $\gamma \approx 0,3775$ made for the contact process [9,10], whereas the critical exponent has a mean-field value $\beta' \approx 0,5$.

Now we will switch to the more general case of the lattice model with the values of q other than 4 (this can be achieved by choosing appropriate value for the radius R_{nn} for neighbors sphere, as remarked above). By doing so we increase the effective range of the disease spread from infected individual to susceptible ones that surround it. If R_{nn} is of the range of $L/2$ then the model is reduced to the deterministic *SIS* model, Eq.(1).

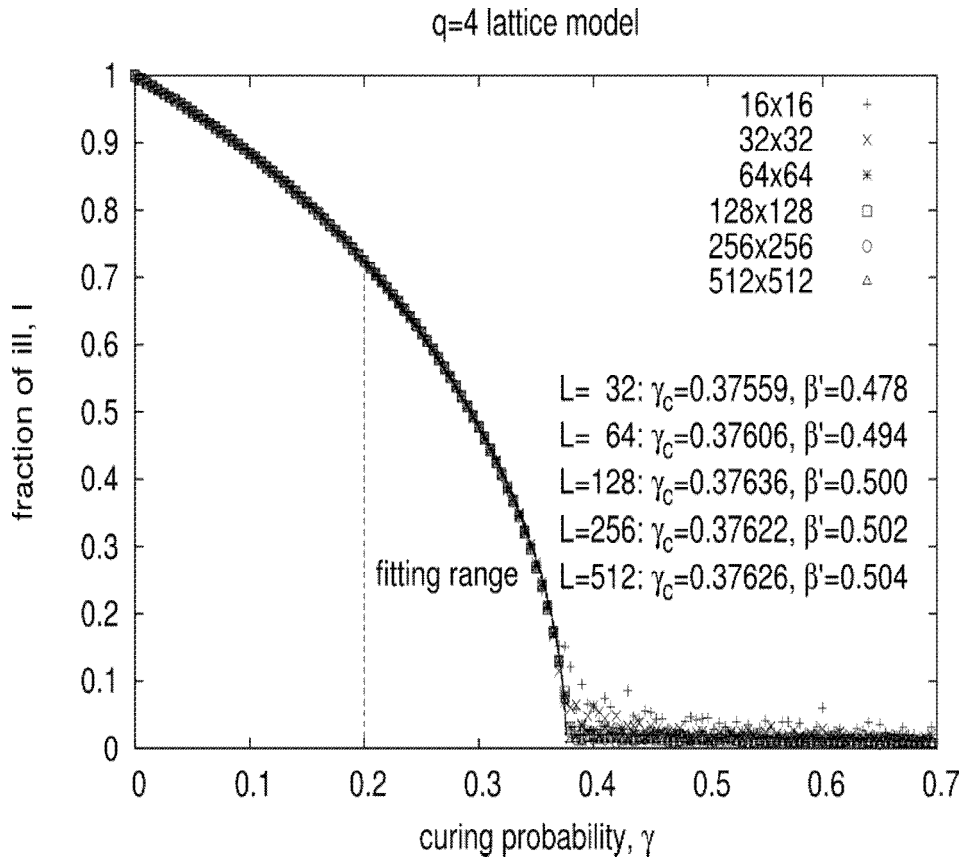


Fig.3. Fraction of infected individuals I in the stationary state vs curing rate γ for the $q = 4$ model of system sizes $L = 16 - 512$. Fitted parameters β' and γ_c are shown in the plot.

In this case all individuals are bonded one with another and the model is reduced to a mean-field type system. The simulation algorithm discussed above is inefficient in this case, as far as the computing time expenditure is proportional to N^2 (due to the fact that the neighbor lists for each individual include all the other individuals). Therefore step 2 of the algorithm is omitted and step 3 is simplified in a way that each infected individual infects with the probability $(1-\gamma)/I$ any other random individual in a system. The local reproduction number $\lambda = \beta/\gamma = (1-\gamma)/\gamma$ reduces itself to its global counterpart R_0 in this case. The latter has its critical value at $R_0 = 1$. Therefore, the critical value for γ_c is expected to be equal to $\gamma_c = 1/(\lambda_c + 1) = 1/(R_0 + 1) = 0.5$. This is found indeed to be a case, as is indicated in Fig.4, where the fraction of infected individuals I vs curing rate γ is shown for the cases of $q = 4, 8, 12, 20$ and for the case when all particles interact one with another, denoted as 'q=all'. In the latter case γ_c is found to be very close to 0.5 (see, Fig.4). The critical exponent β' is found to be dependant on the effective range of interaction q , approaching value close to $3/4$ in the limiting case of q =all.

Let us now turn our attention to the effects of clustering for the infected individuals. Their spatial distribution within the simulation box can be monitored via snapshots, where infected individuals are shown in black, whereas susceptible are not shown. Fig.5 contains such snapshots for the $q = 4$ model of linear size $L = 256$ studied at curing rate $\gamma = 0.3$ and monitored at time instances $t = 0, 20$ and 80 . The initial system (at $t = 0$) is homogeneous, whereas an evident clustering takes place as the system moves towards the stationary state.

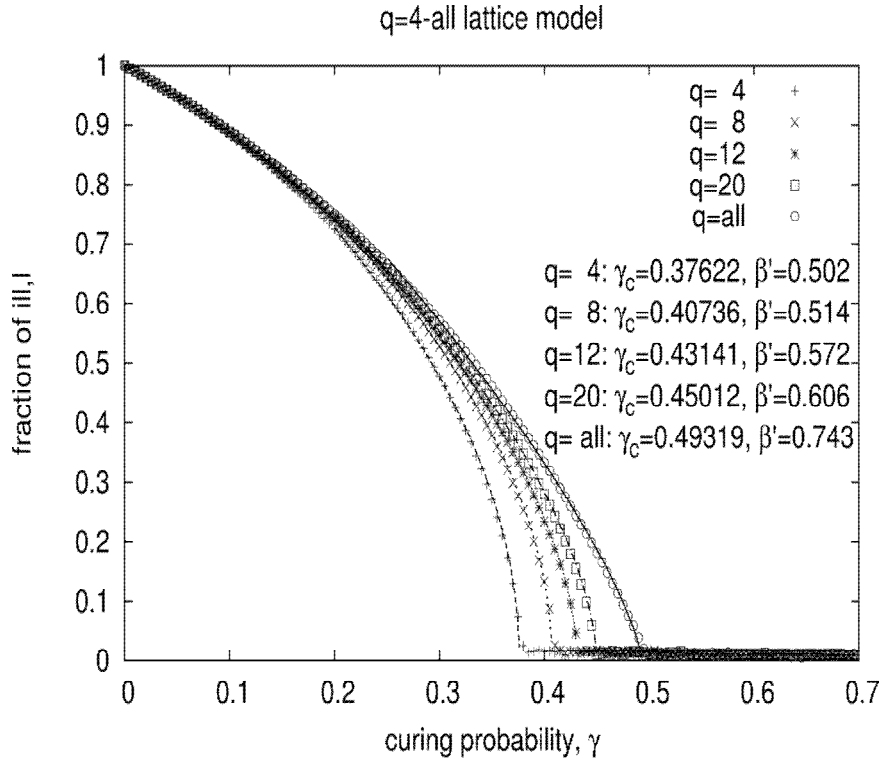


Fig.4. The same as Fig.3 but for the generalized model with various q ($q=all$ is for the limit). Linear system size is $L = 256$

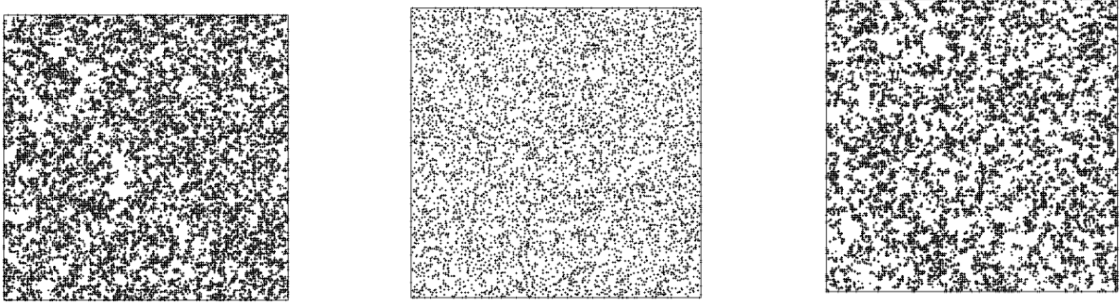


Fig.5. Snapshots illustrating state of the system of linear size $L = 256$ and curing rate $\gamma = 0,35$ while reaching the stationary state. Time instances $t = 0$ (image on the left), $t = 20$ (image in the middle) and $t = 80$ (image on the right) are shown. Infected individuals are shown in black, susceptible ones are not shown. See also Fig.2 for the reference.

To quantify the clustering effects we introduce such properties as scaled average cluster size $\langle n_c \rangle = \langle N_c \rangle / N$ (here N_c denotes cluster sizes in number of individuals, averaging is performed over all clusters in the system), scaled maximum cluster size $n_{c,max} = N_{c,max} / N$ ($N_{c,max}$ is the number of particles in the largest cluster in the system) and scaled maximum cluster dimension $l_{max} = \max\{L_x, L_y\} / L$, where L_x, L_y is maximum span of the cluster along X and Y axis, respectively. The latter property serves as an indicator for the percolation, which takes place if $l_{max} \approx 1$. Behaviour of the average cluster size indicates strong system size, as is seen in Fig.6. Shift of the curve towards left side is a consequence of the fact that a larger number of small clusters appear in larger system, resulting in both lowering their scaled sizes and reducing the average value.

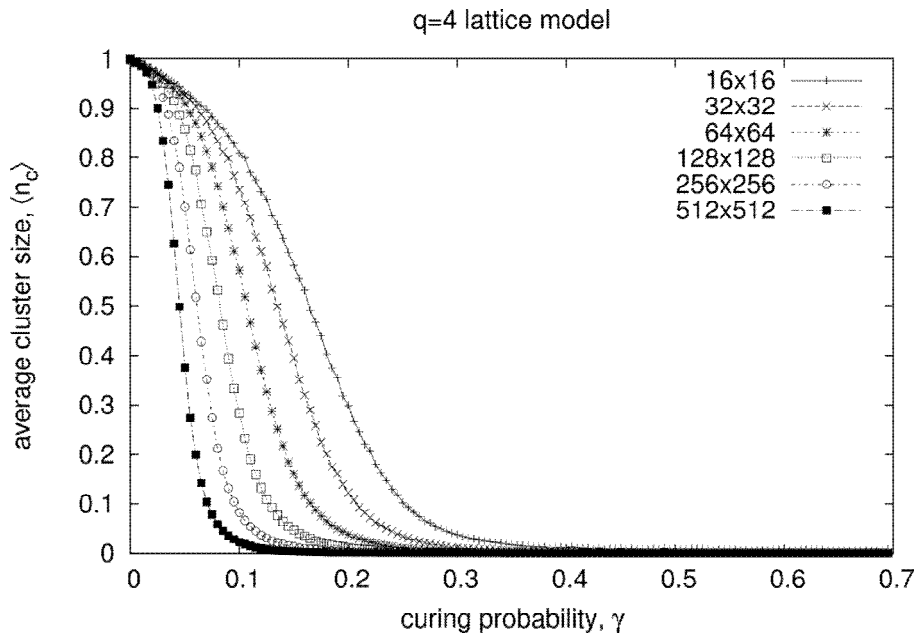


Fig.6. The average cluster size vs γ for $q = 4$ and various system sizes $L = 16 - 512$ indicated in the plot.

The behavior of both maximum cluster size (Fig.7) and maximum cluster dimension (Fig.8) indicate the presence of another transition which takes place at $\gamma = 0,27$. The plot for the maximum cluster dimension (Fig.8) indicate the existence of the percolation transition exactly at the same value $\gamma = 0,27$. The curves drop down abruptly at $\gamma = 0,27$ and the effect is more pronounced for larger system sizes. The details of both transitions requires more thorough analysis and, possibly, application of more sophisticated techniques applicable near the phase transition point. As one knows, in this region the system is subject to slow relaxation and both longer runs and special algorithms should be applied. This, as well as generalized models with various types of connecting graphs, are reserved for future publications.

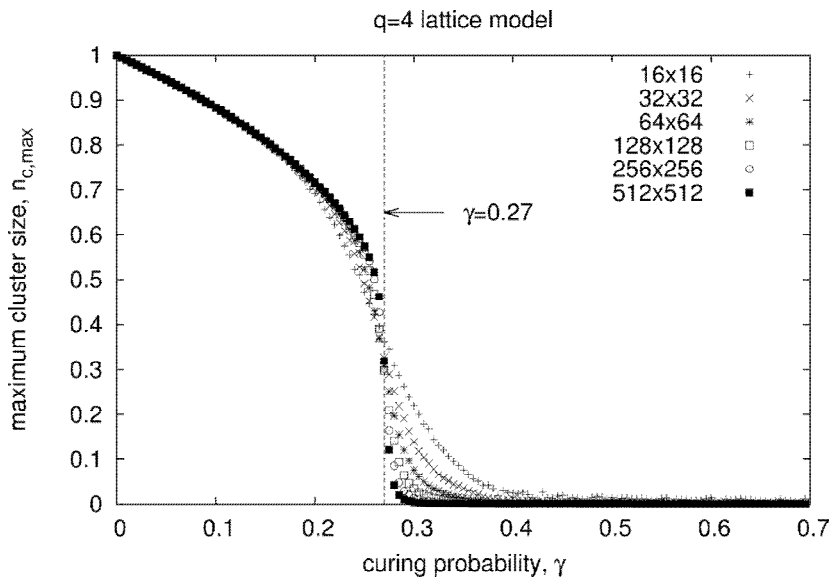


Fig.7. The maximum cluster size vs γ for $q = 4$ and various system sizes $L = 16 - 512$ indicated in the plot

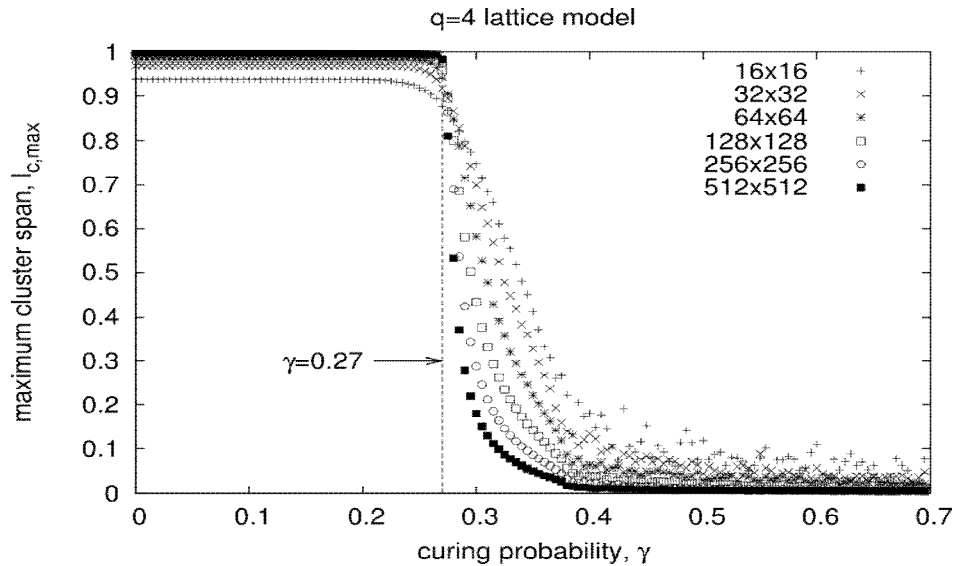


Fig.8. The maximum cluster dimension l_{max} vs γ for $q = 4$ and various system sizes $L = 16 - 512$ indicated in the plot.

3. Conclusions

In this paper, we have presented results of computer simulations of a stochastic model for study of non-immune disease spread. In models of this type, the point at which an epidemic occurs corresponds to the percolation probability at which a giant component of connected nodes first forms. In addition to study static behavior of a system in vicinity of this point (which corresponds to percolation), one also analyzes dynamics of disease spread and a non-equilibrium phase transition which is typical for the disease-spreading models. Our results bring about the influence of the details of microscopic structure and local infective dynamics to the spread of non-immune disease. We reproduced the known critical behavior of contact process and extended the study to the case of variable range of infecting. It is found that the percolation transition is followed by complete dying out of the disease at larger value of curing probability. In-between two transition system is split into clusters of infected individuals with their dimensions much smaller than the system size.

With the above results at hand, we are in position to introduce several features modifying the model and making it closer to the real-world situations. On the one hand, work under progress is targeted on the off-lattice case, on the asymmetric model with adjustable both local curing and infecting rates independently, as well as including stochastic jumps of individuals. On the other hand, the delayed infecting and curing can be studied, matching the real dynamics of tuberculosis. In the latter case the statistical data available for Ukraine will be used. We also plan to consider the model, keeping its geographical embedding effects, on the scale-free small world network [7]. The latter are much closer to human social networks and have already served as a good testing ground for analysis of epidemic spreading [14].

Acknowledgement

This work was supported by the International Research Staff Exchange Scheme grant “Structure and Evolution of Complex Systems with Applications in Physics and Life Sciences” STREVCOMS-612669 within the 7th European Community Framework Program.

1. W.O. Kermack, A.G. McKendrick, *A contribution to the mathematical theory of epidemics*, *Proc. Roy. Soc. Lond. A* 115 (1927) 700. 2. F. Brauer, *The Kermack–McKendrick epidemic model revisited*, *Math. Biosci.* 198 (2005) 119–131. 3. C. Ozcaglara, A. Shabbeera, S.L. Vandenbergc, B. Yenera, and K.P. Bennetta, *Epidemiological models of Mycobacterium tuberculosis complex infections*, *Math. Biosci.* 236(2) (2012) 77–96; Z. Hu, Z. Teng, H. Jiang, *Stability analysis in a class of discrete SIRS epidemic models*, *Nonlinear Analysis: Real World Applications* 13 (2012) 2017–2033. 4. See e.g. C. Domb. *The Critical Point*. (Taylor & Francis, London, 1996); Yu. Holovatch (Ed.). *Order, Disorder and Criticality*.

Advanced Problems of Phase Transition Theory, vol. 1 (World Scientific, Singapore, 2004); vol. 2 (World Scientific, Singapore, 2007); vol. 3 (World Scientific, Singapore, 2012). 5. T. Britton, *Stochastic epidemic models: A survey*, *Mathematical Biosciences* 225 (2010) 24–35. 6. R.I. Hickson, G.N. Mercer, K.M. Lokuge, *A Metapopulation Model of Tuberculosis Transmission with a Case Study from High to Low Burden Areas*, *PLoS ONE* 7(4) (2012) e34411. 7. See e.g. R. Albert, A.-L. Barabási, *Statistical mechanics of complex networks*. *Rev. Mod. Phys.* 74 (2002) 47-97; M. E. J. Newman, *The structure and function of complex networks*. *SIAM Review* 45 (2003) 167-256; S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, *Complex networks: Structure and dynamics*. *Physics Reports* 424 (2006) 175-308; Yu. Holovatch, O. Olemskoi, C. von Ferber, T. Holovatch, O. Mryglod, I. Olemskoi, V. Palchykov, *Complex networks*. *J.Phys.Stud.* 10 (2006) 247-289. 8. P. Grassberger and A. de la Torre, *Reggeon field theory (Schlöggl's first model) on a lattice: Monte Carlo calculations of critical behaviour*, *Ann. Phys. (N.Y.)* 122 (1979) 373. 9. A.G. Moreira, R. Dickman, *Critical dynamics of the contact process with quenched disorder*, *Phys. Rev. E* 54 (1996) R3090-3093. 10. M.M.S. Sabag and M.J. de Oliveira, *Conserved contact process in one to five dimensions*. *Phys. Rev. E* 66 (2002) 036115. 11. D.C. Rapaport, *The Art of Molecular Dynamics Simulation*, Cambridge University Press; 2 edition (2004), 564p. 12. J.M. Ilnytskyi, M.R. Wilson, *A domain decomposition molecular dynamics program for the simulation of flexible molecules with an arbitrary topology of Lennard–Jones and/or Gay–Berne sites*, *Comp. Phys. Commun.* 134 (2001) 23–32. 13. J. Hoshen and R. Kopelman, *Percolation and cluster distribution. I. Cluster multiple labeling technique and critical concentration algorithm*, *Phys. Rev. B* 14, (1976) 3438. 14. R. Pastor-Satorras, A. Vespignani, *Epidemic Spreading in Scale-Free Networks*. *Phys. Rev. Lett.* 86 (2001) 3200-3203; 15. M. E. J. Newman, D. J. Watts, *Scaling and percolation in the small-world network model*. *Phys. Rev. E* 60 (1999) 7332-7342; 16. C. Moore, M. E. J. Newman, *Epidemics and percolation in small-world networks*. *Phys. Rev. E* 61 (2000) 5678-5682; 17. Y. Moreno, R. Pastor-Satorras, A. Vespignani, *Epidemic outbreaks in complex heterogeneous networks* *Eur. Phys. Journ. B* 26 (2002) 521; 18. M. E. J. Newman, *Spread of epidemic disease on networks*. *Phys. Rev. E* 66 (2002) 016128.

УДК 004.773.2

А. Пелецишин, Р. Корж

Національний університет “Львівська політехніка”,
кафедра соціальних комунікацій та інформаційної діяльності

ФОРМУВАННЯ МОДЕЛІ КАТАЛОГУ ГЕНЕРАТОРІВ ІНФОРМАЦІЙНОГО ОБРАЗУ ВНЗ

© Пелецишин А., Корж Р., 2014

Уведено поняття коректно визначеного генератора, побудовано формальну теоретико-множинну модель генератора, запропоновано формалізацію структури віртуальної спільноти як середовища інформаційної діяльності ВНЗ.

Ключові слова: вищий навчальний заклад (ВНЗ), генератор інформаційного образу, віртуальна спільнота.

In the paper the notion of correctly defined generator is introduced, formal set-theoretic model of generator is built and formalization of structure of virtual community as a platform of university information activity is suggested.

Key words: university, generator of information image, virtual community.

Вступ

Перспективні задачі з організації, автоматизації та оптимізації інформаційної діяльності у віртуальних спільнотах охоплюють широкий спектр теоретичного та практичного інструментарію зі сфери як програмно-технічних рішень, так і соціальних комунікацій. Тому модель віртуальної спільноти повинна максимально повно відображати характеристики у цих сферах.

Відповідно до моделі віртуальних спільнот введено такі групи додаткових характеристик:

- технологічні показники;
- кількісні показники (показники обсягів та інтенсивності);
- соціально-демографічні характеристики;
- комунікативні характеристики;
- спеціальні показники освітньої спільноти.

Указані характеристики та показники є базовими, оскільки отримуються в результаті безпосереднього аналізу віртуальних спільнот і є фактично первинними даними про спільноти. Проте не в усіх випадках описують спільноти як генератори інформаційного образу, адже генератор може бути як окремою спільнотою, так і лише її фрагментом (набором дискусій або розділом) або “квазіспільнотою” на основі матеріалів онлайн-ЗМІ.

Місце взаємодії та спільноту, що взаємодіє, називатимемо генератором інформаційного образу. Як видно з цього означення генератора, для великих форумів або сайтів, що мають активну аудиторію, яка створює контент (“спільноту”), генератором можна вважати певний тематичний сегмент такого сайту.

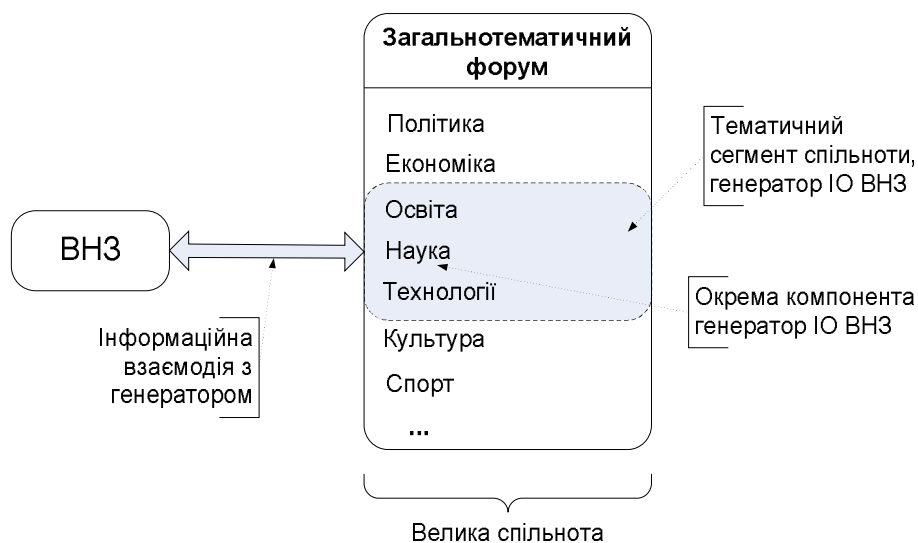


Рис. 1. Інформаційний генератор як частина віртуальної спільноти

Виділення генератора як частини спільноти (рис. 1) є виправданим кроком, який допомагає усунути колізію між необхідністю захисту інтересів ВНЗ у великих спільнотах та нерелевантністю окремих частин наповнення інтересам ВНЗ.

Зазначимо, що на основі однієї окремої спільноти можуть формуватись кілька окремих генераторів, значною мірою не залежних один від одного [1].

Постановка проблеми

Розглянемо питання достатності базової моделі каталогу для процесу формування інформаційного образу ВНЗ. Теоретично за умови вузького набору спільнот, до якого не входять загальнотематичні спільноти, професійні спільноти, проблемні спільноти, “квазіспільноти” і який обмежується лише освітньо-організаційними спільнотами, зіставлення “спільнота–генератор” було би коректним. Проте у цьому разі виникають такі критичні для великих ВНЗ недоліки:

- наукові та проблемно-орієнтовані підрозділи відсторонюються від інформаційної діяльності, оскільки не мають доступу до спільнот, що відповідають їхній сфері компетенції та інформаційному потенціалу;
- більша частина спільнот з частковою релевантністю або релевантністю окремим напрямкам ВНЗ за лишається поза сферою інформаційної діяльності.

Аналіз досліджень та публікацій

Тематику формалізації ВНЗ як суб'єкта інформаційної діяльності в соціальних середовищах Інтернету започатковано в попередніх дослідженнях авторів. Так, у статті [2] визначено комплекс показників віртуальної спільноти для вищих навчальних закладів, а саме: технологічні, кількісні, соціальної релевантності, комунікативної цінності, комунікативної складності та тематичної релевантності. Розглянуто та проаналізовано соціально-демографічні характеристики, показники обсягів та інтенсивності, технологічні показники та комунікативні характеристики віртуальних спільнот, які є основою для визначення окремих інтегрованих показників.

У статті [3] подано роль веб-сайту кафедри, її функції та інформаційне наповнення, значення веб-сайту кафедри як ключового структурного підрозділу формування порталу ВНЗ, що дає змогу при формуванні моделі каталогу генератора інформаційного образу врахувати всі підрозділи ВНЗ.

Згадаємо працю [4], у якій запропоновано методичку розроблення інформаційної моделі контенту університетського веб-сайту, що функціонує в інтегрованому інформаційно-комунікаційному просторі ВНЗ.

Формулювання мети статті

У статті вирішується завдання формування моделі каталогу генераторів інформаційного образу ВНЗ.

Виклад основного матеріалу

1. Компонентне визначення генератора інформаційного образу

На практиці виникають складнощі з формальним описом та подальшим опрацюванням генераторів, які є частиною спільноти (а не цілою спільнотою):

- може бути відсутня єдина фізична адреса генератора;
- визначення показників обсягів аудиторії та авторитетності потребує значних затрат ресурсів внаслідок великих обсягів ручної роботи.

У разі відсутності єдиної фізичної адреси накопичують адреси, які дають змогу максимально точно ідентифікувати генератор. Такими адресами для визначення генератора, залежно від технічної платформи, є:

- адреси тематичних колонок чи рубрик – у онлайн-ЗМІ (“ідентифікація за рубрикою”);
- адреси розділів, що релевантні завданням ВНЗ, – у веб-форумах та групах у соціальних мережах (“ідентифікація за розділом”);
- адреси облікових записів авторів, що створюють релевантні дискусії, – у колективних блогах та великих слабкоструктурованих спільнотах (“ідентифікація за автором”);
- адреси окремих релевантних дискусій чи статей з коментарями – у великих слабкоструктурованих спільнотах (“атомарна ідентифікація”);
- сторінки термінів (“тегів”), що ідентифікують релевантні дискусії у колаборативних платформах та блогах (“ідентифікація за терміном”).

Наведені вище адреси вказують на окремі структурні елементи генератора, які надалі називатимемо **компонентами генератора**.

У кожному з перерахованих випадків з високою імовірністю для одного генератора існує більше ніж один компонент.

Компоненти, своєю чергою, містять окремі змістові дискусії, у межах яких і генерується інформаційний образ. Аналогічно до компонент генератора, дискусії, залежно від платформи, оформляються як:

- дискусії – на форумах, групах у соціальних мережах;
- безпосередні коментарі до публікації – в онлайн-ЗМІ, блогах, колективних блогах;
- коментарі до статті – у колаборативних системах, зокрема у вікі-системах (Вікіпедія та аналогічні).

Проте за суттю про різноманітність форм дискусій можна вести мову лише у площині візуального подання для користувача; сама структура дискусії як послідовності окремих дописів користувачів залишається незмінною [5,6].

Розглянемо детальніше питання співвідношення між компонентами генератора та дискусіями. У всіх випадках, крім одного, між компонентою та дискусіями існує відношення “один до багатьох”, а у випадку атомарної ідентифікації є тривіальним, зумовлює відношення “один до одного”. Формально його можна звести до загальнішого, який об’єднує решту варіантів, що показано на рис. 2.

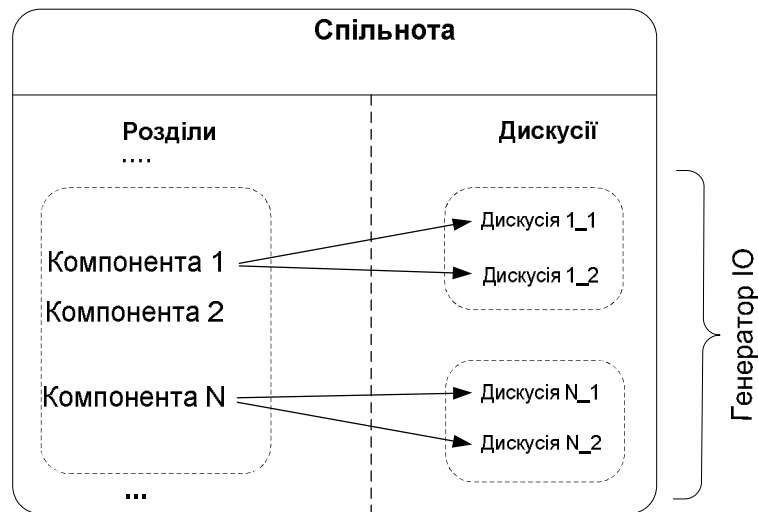


Рис. 2. Співвідношення компонент та дискусій у генераторі інформаційного образу

Лише у випадку, коли генератор визначається через набір термінів (компонента генератора визначається через термін), та сама дискусія може потрапити у різні компоненти генератора. Ця ситуація виникає, якщо дискусія маркується одночасно кількома термінами, які одночасно входять до набору релевантних термінів для визначення компонент.

Відношення компонент до дискусій “багато до багатьох”, що існує в таких випадках, значно ускладнює процес планування діяльності в генераторі та моніторингу активності спільноти у межах генератора. Як буде показано нижче, під час адресації деякі важливі кількісні характеристики генератора не можна визначити коректно. Тому до вказаного методу доцільно вдаватися лише за крайньої потреби, коли іншого способу визначення компонент генератора не існує [7].

На практиці, у разі початкової ідентифікації компонент за адресами релевантних термінів, доцільним є подальший аналіз компонент для виділення іншої ознаки для адресації, основні кроки якого наведено на рис. 3.



Рис. 3. Поетапне керування процесом формування інформаційного образу

З незначними модифікаціями запропонований алгоритм можна застосувати і для інших перетворень одного способу ідентифікації компонента на другий (рис. 4).

Цей підхід достатньо ефективний, оскільки для багатьох платформ найпростішим способом виявлення тематично релевантної частини сайту є саме ідентифікація за терміном.

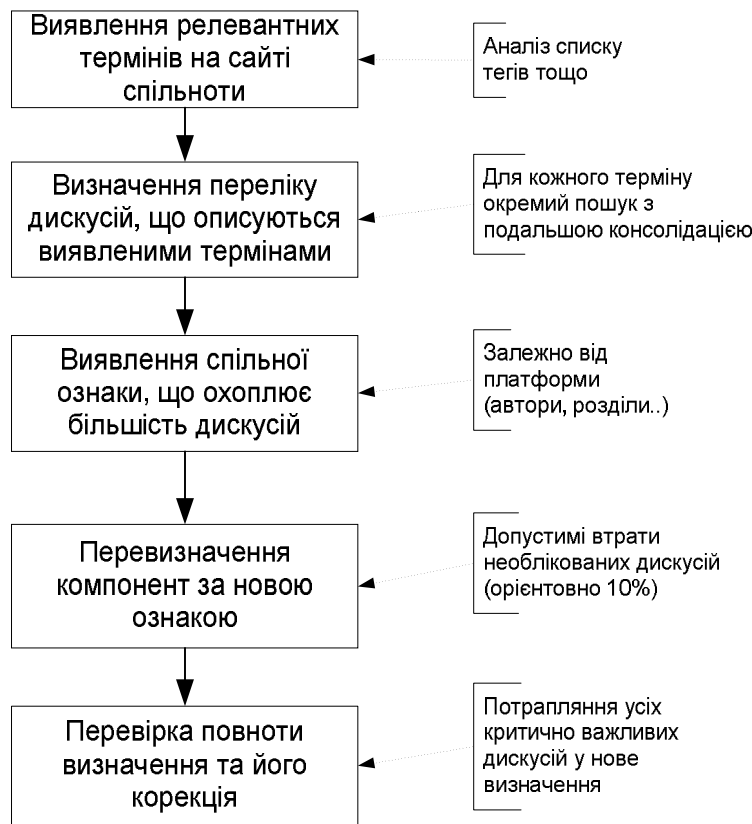


Рис. 4. Зміна способу ідентифікації компоненти

Аналогічно до розглянутої вище ситуації з відношенням між компонентами та дискусіями, яке пропонується зводити до “один до багатьох”, виникає питання щодо співвідношення між генераторами в одній спільноті та їхніми компонентами.

Під час попереднього аналізу спільноти та виділення в ній компонент, які можуть бути складовими генераторів, можлива є ситуація, коли окремі компоненти увійдуть до кількох генераторів одночасно. Наприклад, розділ “Дипломне проектування” освітнього форуму може одночасно входити до генератора “наукова діяльність ВНЗ” та до генератора “магістратура ВНЗ”. Тобто у загальному випадку між компонентами генератора та генератором може існувати відношення “багато до багатьох”.

Проте у разі, коли виділення генераторів здійснюється заради формування та закріплення окремих команд для інформаційної діяльності, виникає ряд ризиків для ВНЗ:

- можливість конфлікту між виконавцями внаслідок неузгоджень;
- розміщення суперечливої інформації;
- “перегрів” окремого компонента через надмірний потік інформації.

Як наслідок, для планування розподілу ресурсів та призначення завдань для інформаційної діяльності ВНЗ необхідно усунути такі колізії, перейшовши до відношення “один до багатьох” із зарахуванням кожної компоненти лише до одного генератора.

Водночас, для деяких завдань аналітичного характеру (пошуку оцінкових суджень, споживацького досвіду тощо) таке спрощення може не знадобитись.

Далі вважатимемо, що, якщо не вказане інше, використовується варіант “один генератор – кілька компонент” [8].

2. Формалізація структури віртуальної спільноти як середовища інформаційної діяльності ВНЗ

Узагальнюючи структуру спільноти як середовища інформаційної діяльності ВНЗ, отримуємо таку схему (рис. 5):

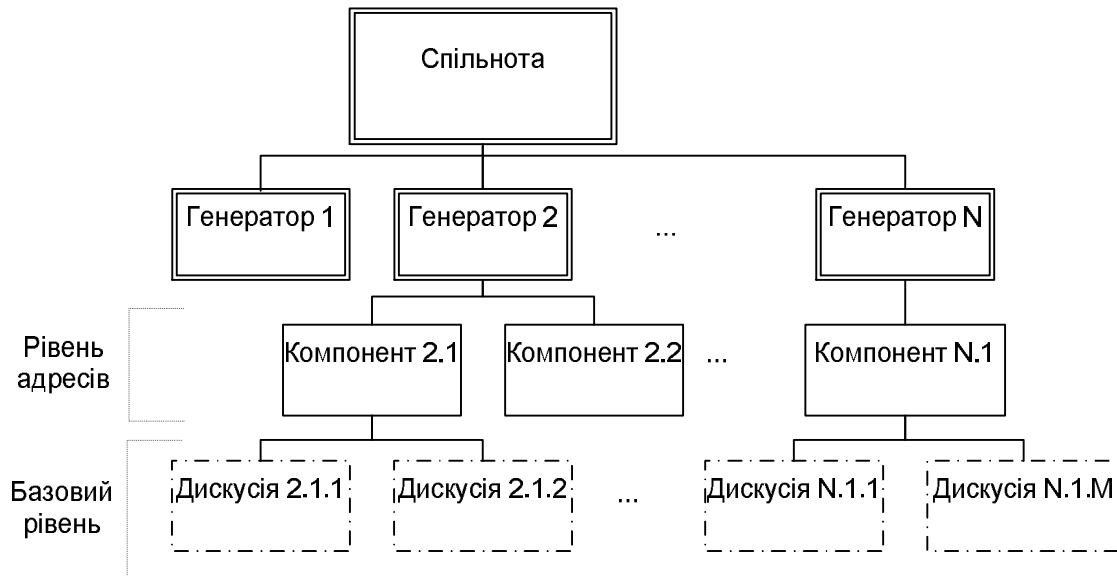


Рис. 5. Багаторівнева структура спільноти

Формалізуємо наведену вище схему віртуальної спільноти. Нехай

$$Diss(VC^*) = \{Diss_k\}_{k=1}^{N^{(Diss)}}, \quad (1)$$

де VC^* – певна віртуальна спільнота; $N^{(Diss)}$ – кількість повідомлень у спільноті VC^* ; $Diss(VC^*)$ – множина усіх дискусій спільноти.

Водночас віртуальна спільнота визначається також набором компонент:

$$Diss(VC^*) = \{Comp_j\}_{j=1}^{N^{(Comp)}}. \quad (2)$$

Компоненти також мають власну множину дискусій $Diss(Comp_j)$, причому:

$$Diss(Comp_j) \subset Diss(VC^*). \quad (3)$$

Множину компонент вважатимемо **коректно визначеною**, якщо:

$$Diss(Comp_j) \cap Diss(Comp_l) = \emptyset, \quad \forall j, l = 1 \dots N^{(Comp)} \quad (4)$$

Алгоритм отримання коректно визначеної множини компонент наведено на рис. 4.

Кожен генератор у спільноті формується з множини компонент генератора $Comp(Gen_i)$. Тоді множина дискусій генератора Gen_i :

$$Diss(Gen_i) = \bigcup_{Comp_j \in Comp(Gen_i)} Diss(Comp_j). \quad (5)$$

Множину генераторів вважатимемо **коректно визначеною**, якщо:

$$Comp(Gen_i) \cap Comp(Gen_m) = \emptyset, \quad \forall i, m = 1 \dots N^{(Gen)}, \quad (6)$$

де $N^{(Gen)}$ – кількість виділених генераторів інформаційного образу в спільноті VC^* [9].

Очевидно, що:

$$Diss(Gen_i) \subset Diss(VC^*). \quad (7)$$

3. Використання характеристик віртуальних спільнот для опису генераторів інформаційного образу

Розглянемо детальніше питання ужитковості та застосовності пропонуванних характеристик віртуальних спільнот до генераторів інформаційного образу.

Як зазначено вище, генератор часто є лише певною частиною спільноти. Це стосується як інформаційного наповнення сайту спільноти, так і її обсягів (аудиторії). Отже, технічні, комунікативні характеристики спільнот прямо переносяться на генератори, утворені на базі спільноти. Окреме, незалежне від спільнот визначення соціально-демографічних характеристик для генератора на практиці також неможливе і недоцільне через безперешкодну участь будь-якого учасника спільноти у дискусіях у межах генератора.

Проте, кількісні характеристики безпосередньо до генераторів застосувати не можна. Для великих спільнот частка генератора інформаційного образу ВНЗ та наповнення, релевантного ВНЗ, достатньо незначна (у деяких випадках навіть менша за 1 %). Як наслідок, враховувати загальні обсяги спільноти як такі, що властиві генератору, некоректно [10].

Враховуючи характер генераторів як частини спільноти, найприроднішим варіантом є визначення кількісних характеристик генератора як частки від показників усієї спільноти. Визначити частку від кожного показника можливо, проте обчислення надто трудомістке і потребуватиме постійного оновлення, зважаючи на високу динамічність процесів, що відбуваються у спільноті між окремими її частинами.

Тому пропонується для кожного генератора Gen_j встановити його узагальнену частку в спільноті $Part(Gen_j)$, що визначає рівень впливу генератора та його значущість у функціонуванні спільноти. Визначимо, що $0 \leq Part(Gen_j) \leq 1$. Якщо ж спільнота загалом є генератором, приймемо $Part(Gen_j) = 1$.

Повноцінно визначити такий коефіцієнт сьогодні можна лише експертним способом на підставі доступних даних про кількісні характеристики компонент генератора та дискусій.

Початковим наближенням для визначення частки генератора $Part(Gen_j)$ є відношення між кількістю повідомлень у спільноті взагалі, оскільки саме кількість повідомлень визначається активністю та кількістю користувачів спільноти та кількістю обговорюваних тем.

$$Part^{(Post)}(Gen_j) \approx \frac{\sum_{Diss_i \in Diss(Gen_j)} PostCount(Diss_i)}{\sum_{Diss_i \in Diss(VC^*)} PostCount(Diss_i)}, \quad (8)$$

де $PostCount(Diss_i)$ – кількість повідомлень у дискусії.

Проте отримане наближення не відображає факту, що значною мірою кількість повідомлень утворюється в темах, де існує “флейм”, розміщених відповідно в розділах, що не входять до тематичних для ВНЗ (на зразок розділів “Кнайпа”, “Курилка”, “Балачка” на популярних тематичних форумах). Для зменшення такої похибки доцільно використати частку дискусій, що належать генератору, від усіх дискусій спільноти

$$Part^{(Diss)}(Gen_j) \approx \frac{\sum_{Diss_i \in Diss(Gen_j)} Count(Diss_i)}{\sum_{Diss_i \in Diss(VC^*)} Count(Diss_i)}. \quad (9)$$

Проте і цей показник не може бути універсальним способом для визначення $Part(Gen_j)$, оскільки велика кількість “коротких” дискусій з малою кількістю повідомлень характерна для розділів, у яких низька масова активність. Дискусії та повідомлення створюють лише окремі автори або дискусії започатковуються в автоматичному режимі на основі запозичення чи

реплікації контенту з інших (переважно новинних) джерел. У такому разі показник $Part(Gen_j)$ буде відчутно завищеним.

Тобто, можемо встановити лише можливі межі для $Part(Gen_j)$. Нехай:

$$\begin{aligned} \overline{Part}(Gen_j) &= \max\left(Part^{(Diss)}(Gen_j), Part^{(Post)}(Gen_j)\right) \\ \underline{Part}(Gen_j) &= \min\left(Part^{(Diss)}(Gen_j), Part^{(Post)}(Gen_j)\right). \end{aligned} \quad (10)$$

Тоді для визначення узагальненої частки генератора одержуємо:

$$\underline{Part}(Gen_j) \leq Part(Gen_j) \leq \overline{Part}(Gen_j). \quad (11)$$

Так визначаються допустимі межі, у яких повинна міститись експертна оцінка коефіцієнта $Part(Gen_j)$.

Кожен з показників груп VC та VA для генератора Gen_j отримують, помноживши відповідний показник для спільноти на $Part(Gen_j)$. А система базових характеристик віртуальної спільноти VC_i описується кортежем:

$$VC_Ch(VC_i) = \langle VT(VC_i), VA(VC_i), VI(VC_i), VD(VC_i), VG(VC_i) \rangle, \quad (12)$$

де $VT(VC_i)$ – технологічні показники; $VA(VC_i)$ – показники обсягів; $VI(VC_i)$ – показники інтенсивності; $VD(VC_i)$ – соціально-демографічні показники; $VG(VC_i)$ – комунікативні показники [11].

Висновки

Досліджено задачу формалізації онлайн-спільнот як генераторів контенту, який є складовою інформаційного образу ВНЗ. Визначено, що у загальному випадку генератором є не уся спільнота, але лише її певна, релевантна тематиці ВНЗ, частина. Побудовано формальну теоретико-множинну модель генератора, введено поняття коректно визначеного генератора та запропоновано підходи до ідентифікації генератора як підмножини спільноти за різними ознаками. Вибір ідентифікаційної ознаки визначається технологічними характеристиками платформи, на якій функціонує спільнота.

Отримані результати є основою для проектування каталогів та баз даних генераторів інформаційного образу, математичних та імітаційних методів оптимізації інформаційної діяльності ВНЗ у онлайн-спільнотах.

1. Корж Р. О. Формалізація процесу формування інформаційного образу ВНЗ в соціальних середовищах Інтернету / Р. О. Корж, А. М. Пелецишин // Східно-Європейський журнал передових технологій. – Харків, 2013. - №5/3 (65) – С. 4-8. Бібліогр.: 10 назв. 2. Пелецишин А. М. Визначення комплексу показників віртуальної спільноти для вищих навчальних закладів / А. М. Пелецишин, Р. О. Корж, О. Р. Трач // Східно-Європейський журнал передових технологій. – Харків, 2014. – №2/2 (68) – С. 16-23. Бібліогр.: 12 назв. 3. Сугак Д. Б. Роль веб-сайта в науково-образовательной деятельности вуза / Д. Б. Сугак // Вестник СПбГУКИ. - №3(12), сентябрь. – 2012. — С.77–82. 4. Філіпова Л. Інформаційне моделювання контенту університетського веб-сайту: загальні принципи та методика / Л. Філіпова, А. Шелестова // Вісник Книжкової палати, №3, 2014. 5. Тимовчак-Максimeць О. Ю. Розроблення формалізованих запитів для виявлення веб-форумів та дискусій // О. Ю. Тимовчак-Максimeць, А. М. Пелецишин // Вісник Національного авіаційного університету : – Київ, 2012. – № 1 (50). – С. 106–111. 6. Пелецишин А. М. Аналіз комунікативної взаємодії учасників спільнот веб-форумів / А. М. Пелецишин, О. Ю. Тимовчак-Максimeць // Східно-Європейський журнал передових технологій. – Харків, 2010. – № 6/8 (48). – С. 44-49. 7. Корж Р. О. Етапи формування цілісного інформаційного образу ВНЗ та їхня характеристика / Р. О. Корж, А. М. Пелецишин // Інформація, комунікація, суспільство 2014 : матеріали 3-ої Міжнар. наук. конф. ІКС – 2014, 21-24 трав. 2014р., Україна, Львів, Славське / М-во освіти і науки України, Нац.

ун-т "Львів. політехніка". – Л., 2014. – Бібліогр.: 2 назви. 8. Lisa Harris (2010) *Community in the Electronic Classroom: Virtual Social Networks and Student Learning*. VDM Verlag Dr. Müller. P. 304. 9. Korz, R., Peleschyshyn, A., Syerov, Y., Fedushko, S. (2014). "The cataloging of virtual communities of education thematic." *Webology*, 11(1), Article 117. Available at: <http://www.webology.org/2014/v11n1/a117.pdf> 10. Кудін А.П., Кархут В.Я., Франчук В.М. Інформаційно-комунікаційні технології та управління діяльністю вищого навчального закладу: освітній портал. Збірник наукових праць Кам'янець-Подільського національного університету імені Івана Огієнка. Серія педагогічна / [редкол.: П.С.Атаманчук (голова, наук. ред.) та ін.]. – Кам'янець-Подільський: Кам'янець-Подільський національний університет імені Івана Огієнка, 2010. – Вип. 16: Формування професійних компетентностей майбутніх вчителів фізико-технологічного профілю в умовах євроінтеграції. С. 26-29. 11. Verification of the Socio-demographic Characteristics of Virtual Community Members / S. Fedushko, A. Peleshchyshyn, R. Korzh, Yu. Syerov // Сучасні проблеми радіоелектроніки, телекомунікацій, комп'ютерної інженерії: матеріали XII Міжнар. конф., присвяч. 170-річчю заснування Національного університету "Львівська політехніка", 25.02–01.03 2014 р., Львів, Славське, Україна / Нац. ун-т "Львів. політехніка". – Л. : Вид-во Львів. політехніки, 2014. – С. 632.

УДК [004.942:519.85]:630*228.3

Л. Лотиш

Національний лісотехнічний університет України

МАТЕМАТИЧНІ ТА ПРОГРАМНІ АСПЕКТИ МОДЕЛЮВАННЯ ВІКОВОЇ СТРУКТУРИ РОЗВИТКУ ДВОВИДОВОГО ЛІСУ

© Лотиш Л., 2014

Розглянуто математичне моделювання еволюції двовидового середньоширотного лісу з врахуванням вікової структури та створено програмне забезпечення для моделювання системи. Моделювання здійснюється числовим аналізом динамічної системи рівнянь методом Рунге–Кутта. Послідовно описано побудову моделі, що враховує як міжвидову конкуренцію, так і інші фактори: освітлення, заболочення, вікову структуру, кількість опадів, зовнішні впливи. Наведено результати моделювання, отримані на створеному програмному комплексі. Показано можливості використання створеної моделі локального рівня для інформаційного забезпечення розроблення та підтримки прийняття рішень у лісовому господарстві.

Ключові слова: лісове господарство, прогнозування, інформаційні системи, моделювання.

Simulation is carried out by numerical analysis of the dynamic system equations by the Runge-Kutta method. Consistently described the construction of a model that takes into account both interspecies competition and ranks of other factors: light, water-logging, age structure, rainfall, external influences. The results of simulation obtained on the created software. The possibilities of using the created model local level to ensure the development of an information and decision support in forest management.

Key words: forestry, forecasting, information systems, modeling.

Постановка проблеми

Ліс як об'єкт господарської діяльності людини одночасно є і об'єктом наукових досліджень, оскільки лісові насадження – це складне природне утворення. Ліс – складова географічного ландшафту і із сукупності різних видів дерев, чагарників, трав, мохів, тварин і мікроорганізмів, що

є взаємозалежними і впливають один на одного і на навколишнє середовище. Однак ліс слід розглядати не тільки в просторі, але і в часі, враховуючи його власний розвиток. Ці динамічні процеси є результатом складних взаємовідносин між організмами (боротьба за існування і природний відбір, постійне оновлення та розвиток, зміна вікового складу, процеси обміну речовин і енергії). Вивчати лісові біогеоценози необхідно комплексно, у всьому різноманітті зв'язків між його частинами і процесами, що відбуваються всередині нього. Це необхідно робити за системним підходом, тобто принципом системності в пізнанні біологічних явищ природи – як в деталях, так і загалом. Лісові біогеоценози характеризуються високою мінливістю в просторі, оскільки умови мікросередовища в їх межах завжди дещо відрізняються один від одного. Велика варіативність ознак зумовлена і господарською діяльністю людини – все це істотно ускладнює процес досліджень[9].

Моделювання лісових екологічних систем дає змогу передбачити наслідки того чи іншого сценарію, вибравши менш небезпечний і корисний для досліджуваної системи. За допомогою моделей можна доволі швидко прогнозувати наслідки спрямованих дій на багато років й навіть століть. Крім того, безперечним плюсом моделювання та супутнього йому прогнозування є брак негативних або навіть катастрофічних наслідків експериментування на навколишньому середовищі.

Вікова структура є одним з важливих показників лісового фонду під час оцінювання лісових ресурсів та їх сировинного потенціалу. Вона дає уявлення про площі лісів у межах вікових груп та про запаси деревини, які можна використати як в цьому, так і в наступних десятиліттях.

Стійкість лісу до засухи, заболочення, затінення та забруднення, як і до ряду антропогенних факторів значною мірою залежить від віку самих рослин. Крім господарського значення, враховувати віковий розподіл у математичних моделях важливо і для екології місцевості, в якій знаходиться лісовий масив.

Аналіз останніх публікацій та досліджень

У країнах ЄС дослідження та супутні їм програмні розробки на тему прогнозування розвитку лісових насаджень, підтримки прийняття рішень впроваджуються та використовуються як на регіональному, так і на національному рівнях. Проте комплексні системи лісогосподарського планування в своїй більшості з різних причин застосовують ряд спрощень у розрахунках, зокрема не враховують вікового розподілу (наприклад, розраховують зміну діаметра стовбура, а на його основі з використанням емпіричних відношень розраховують інші характеристики, серед яких вік (Busing, 1991; Chave, 1999; Bartelink, 2000). Серед винятків модель SILVA2, де акцент робиться на прирості вистоти (Pretzsch, 1992) та модель Moravie M.A. et al. (1997), де насамперед розраховується збільшення радіуса крони, а згодом діаметр і висота з характеристик крони [5].

HARVEST – система підтримки прийняття рішень (СППР), що була розроблена у США як стратегічне дослідження та інструмент планування – дає змогу оцінити просторову структуру наслідків довготривалих стратегій управління деревиною. Модель добре підходить для оцінювання альтернативних стратегій, прогнозування того, які альтернативні ефекти впливають на вік, просторовий розподіл лісів і структуру отриманого лісового ландшафту (Gustafson, Crow, 1999). Проте лісозаготівельну діяльність модель симулює, надаючи цілим масивам нульового значення, припускаючи, що вікова структура лісу не змінюється в процесі відновлення.

Питання розрахунку таксаційних показників з врахування вікової структури висвітлено в роботі [7] естонських науковців К. Kiviste, А. Kiviste (2009). За допомогою баз даних лісогосподарських підприємств Естонії було виведено алгебраїчні рівняння знаходження висоти стовбура, його діаметра, об'єму залежно від типу лісової ділянки, домінуючих видів та їхнього віку. Модель входить до естонської державної системи інформації про ліси та в декілька програмних продуктів для обробки даних лісовпорядкування.

Серед останніх досліджень за цюю тематикою варто звернути увагу на систему PlanWise [8] (розробники – учені Шведського університету сільськогосподарських наук А. Korosuo, Р. Wikström, 2011), де застосовується підхід багатокритеріального аналізу рішень (Multi-criteria decision analysis

або ж (MCDA). За цим підходом під час планування і моделювання різновікового лісу можна враховувати також негрошові цінності (за принципами сталого розвитку).

Мета дослідження

Метою публікації є показ можливостей застосування імітаційної комп'ютерної моделі для прогнозування динаміки лісових екологічних систем за різними сценаріями антропогенного впливу. Удосконалити математичну модель розвитку двовидового лісу, описану та розглянуту у роботі [4], враховуючи вікову структуру лісу, та представити її програмну реалізацію.

Виклад основного матеріалу

Вік насаджень та вікові фази його розвитку є важливими ознаками, що зображають динаміку лісу. Крім того, вони значною мірою характеризують стійкість рослин до забруднення, надмірного зволоження, посухи, хвороб. За розглянутою в роботі [4] системою рівнянь (1), в якій розглядалася еволюція саджанців одного віку, внесемо зміни для зображення вікового розподілу у лісовому масиві:

$$\begin{cases} \dot{R} = \sum_i (A_i - V_i)u_i + B(P^0 - P) - W, \\ \dot{u}_i = (k_i C_i V_i - D_i)u_i \end{cases} \quad (1)$$

де \dot{R} – узагальнений показник родючості – щільність ресурсу ($\text{кг}/\text{м}^2$); $u_{\text{л}}$ – щільність біомаси листяних порід ($\text{кг}/\text{м}^2$); $u_{\text{х}}$ – щільність біомаси хвойних порід ($\text{кг}/\text{м}^2$); A_i – коефіцієнт відновлення ґрунту за рахунок опадів і-ї породи; B – коефіцієнт самовідновлення ґрунту (1/рік); P^0 – асимптотичне значення родючості за відсутності лісу ($\text{кг}/\text{м}^2$); V_i – швидкість споживання ресурсу (трофічна функція) (1/рік); C_i – корегуючий множник, що описує конкуренцію; k_i – коефіцієнт росту і-ї породи; D_i – коефіцієнт природної смертності дерев (1/рік); W – вплив зовнішніх факторів, найчастіше негативний, тому з від'ємним знаком, ($\text{кг}/\text{год} \cdot \text{м}^2$); t^0 – середній час дозрівання лісу (років).

Враховуючи, що інтервали вікових груп суттєво відрізняються від прийнятого в моделі [4] кроку (одиниця) за часом, то залежно від темпів розвитку деревних порід слід додати диференціацію за віком у межах самого інтервалу на вікові групи.

Вікову структуру насаджень зобразимо, користуючись прийнятим у лісогосподарській діяльності поділом на групи віку (молодняки 1-го класу, молодняки 2-го класу, середньовікові, пристигаючі, стиглі, перестійні) та вважатимемо, що двовидову структуру лісового масиву формують бук (як листяна порода) та ялина біла (як хвойна порода) з відповідним класом віку.

При побудові вікової моделі орієнтуємося на модель [3], що являє собою в основі класичну модель Вольтера, з граничними і початковими умовами:

$$u(t,0) = \int_t^{T_m} u(u, T') u(t, T) dT; \quad (2)$$

$$u(0, T) = u(T). \quad (3)$$

Якщо вікова група включає p років (залежно від класу віку) і розглядається режим, за якого на вхід групи надходить одне і те саме значення біомаси u_0 , то на кожний рік за віком ми отримуватимемо такі щільності біомаси [3]:

1-й рік – $u_0(1 + C)$;

2-й рік – $u_0(1 + 2C), \dots$;

p -й рік – $u_0(1 + pC)$;

C – коефіцієнт, що показує, наскільки збільшується вихідна щільність біомаси за рік. Його величина зазвичай становить 0,1–0,18 та залежить від групи, щільності біомаси, родючості. Однак у

межах групи вона змінюється мало. Якщо група відповідно до класу віку складає 10 років, то прийнятий лінійний закон приросту в групі цілком виправданий.

Враховуючи, що повний об'єм біомаси в групі – це сума всіх об'ємів за роками:

$$U[j] = [p + 0,5(1+p)pC[j]] \cdot u_0, \quad (4)$$

то частка біомаси від маси групи, що за кожен рік переходить в іншу групу, можна визначити відношенням:

$$Q[j] = \frac{1+pC[j]}{p+0,5(1+p)pC[j]} \quad (4)$$

При цьому введена у (4) функція C не збігається з визначеною у [4], оскільки містить і k , і трофічну функцію, причому явно вказано лише залежність C від j .

Загалом функцію $u_i[j]$ можна описати рівняннями, схожими на систему (1), проте замість коефіцієнтів C_i^0 ; D_i^0 вводяться масиви приросту і відмирання $C_i^0[j]$ і $D_i^0[j]$, відповідно.

При $p=10$ отримаємо:

$$Q_i[j] = 0,1 + \frac{4,5C_i^0[j]}{10 + 5,5C_i^0[j]}. \quad (5)$$

Враховуючи вікову структуру, не можна не враховувати той фактор, що молодняки та стиглі дерева по-різному переходять засуху, заболоченість та інших факторів. Тому модифікуємо поправкові функції(6), що представлені в [4], за допомогою залежності від віку j .

$$C_i[j] = C_i^0 \text{Sun}_i[j] \text{Bog}_i[j] \cdot \text{Ext}[j] - \text{Rain}_i[j] \quad \text{для } j \neq 0 \quad (6)$$

Для $j=0$, при виконанні однієї з двох умов: $N^{btot} > 0$, $N > \sum_{i,j} \prod u_i[j]$ – виконується той самий вираз, проте з $j=0$ і у випадку, коли обидві умови порушуються, то $C_i[0]=0$.

Надалі будемо використовувати такі позначення: N^{btot} – повний “шар води”, що був накопичений за весь час існування лісу; N^{cr} – критична “товщина”, за якої відбувається різке угнітання лісу.

$$\text{Bog}_i[j] = \text{Bog}_i(N_i^{cr}[j], N^{btot}). \quad (7)$$

Швидкість споживання ресурсу задамо у вигляді:

$$V_i[j] = V_i^p[j] \cdot p \quad (8)$$

Функція $\text{Rain}_i(\bar{u}, N)$ дає змогу враховувати негативний вплив засухи на річний приріст біомаси. Ця функція описує загальну кількість опадів за рік $N(t)$ і враховує необхідні затрати цієї породи на випаровування та внутрішнє споживання води. Вважається, що листовим породам для нормального життя без росту необхідно більше вологи, ніж хвойним[4]. Засуху зображаємо рівнянням:

$$\text{Rain}_i[j] = \begin{cases} \text{Rain}_i \cdot \frac{V_i[j]}{\sum_{i'} V_i[i'] \cdot u_i[i']}, & j \neq 0 \\ 0, & j = 0 \end{cases} \quad (9)$$

Врахуємо, що збільшення узагальненого показника родючості відбувається завдяки переробленню щорічного опадів (A_i^0) і відмерлої біомаси, тобто дорівнює сумі[4]:

$$A_i[j] = A_i^0 + A_i^1 D_i^0[j]. \quad (10)$$

Для запису системи рівнянь використаємо такий спосіб скорочення: в добутку функцій однакової структури будемо використовувати індекси лише правого множника, тобто запис $C_i[0] \cdot V_i[0] \cdot u_i[0]$ еквівалентний $C \cdot V \cdot u[0]$. У таких позначеннях отримаємо систему (11) [3]:

$$\begin{cases} \dot{u}[0] = \sum C V_i[0] \cdot k_i(u'', j) - Q u_i[0] + C k V u_i[0] - D u_i[0] - Ext_i[0], \\ \dot{u}[j] = Q u_i[j-1] - Q u_i[j] + C k V u_i[j] - Ext_i[j], \\ \dot{P} = \sum_{ij} (A - V) u_i[j] + B(p^0 - p) - W \end{cases} \quad (11)$$

Результати моделювання

Одним з основних етапів під час розроблення математичних моделей деревостанів є аналіз, оцінювання та моделювання динамічних процесів, які в них відбуваються. Адекватне зображення зазначених процесів у вигляді математичних моделей надалі дасть змогу розробити достовірніші моделі для прогнозування їх росту і розвитку. Для деревостанів головних лісотвірних видів Карпатського регіону в статті як базові прийнято такі ознаки, як віковий склад деревостану та щільність біомаси порід.

Згідно з вищеописаною математичною моделлю було доповнено програмне забезпечення, описане у [4]. Для зображення адекватності створеної моделі було здійснено числові розрахунки в можливих природних умовах. Параметри розрахунку для моделі: кількість груп віку – 6, інтервал вікової групи – 10 років, часовий крок 1 рік, час моделювання – 150 років, кількість опадів за рік – 800 мм.

На рис.1 зображено щільності біомаси u листяних (бук) та хвойних (ялина біла) порід, що формують двовидову структуру лісового масиву відповідно до класу віку. Як видно з рис. 1, щільність ресурсу P (кг/м²) швидко зростає до досягнення межі росту породи, після чого темпи приросту знижуються.

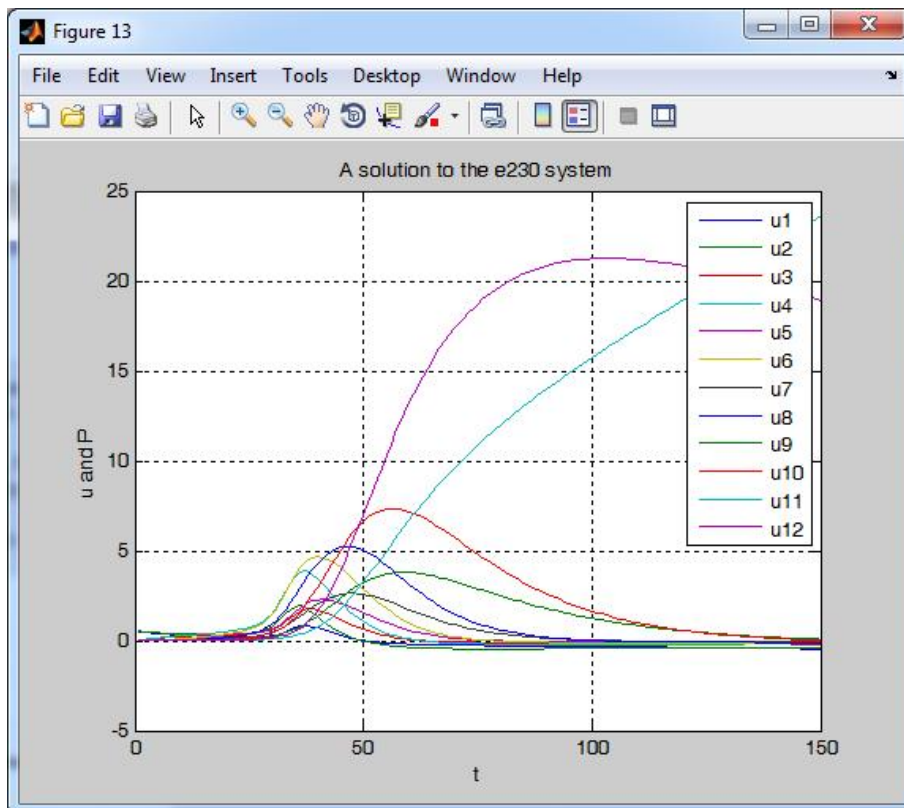


Рис. 1. Розвиток хвойних та листяних порід відповідно до групи віку

На рис. 2 показано еволюцію лісу в зоні висадки (в центральній точці масиву) залежно від часу. З рисунка видно, що спочатку на місці висадки проходить бурний ріст листяних порід та їх поширення на вільні сусідні ділянки. Можемо припустити, що з часом хвойні породи в центрі масиву активно витісняють листяні, що призводить до формування кільцевої структури: в середині старий хвойний ліс, а зовні – світлі листяні ділянки.

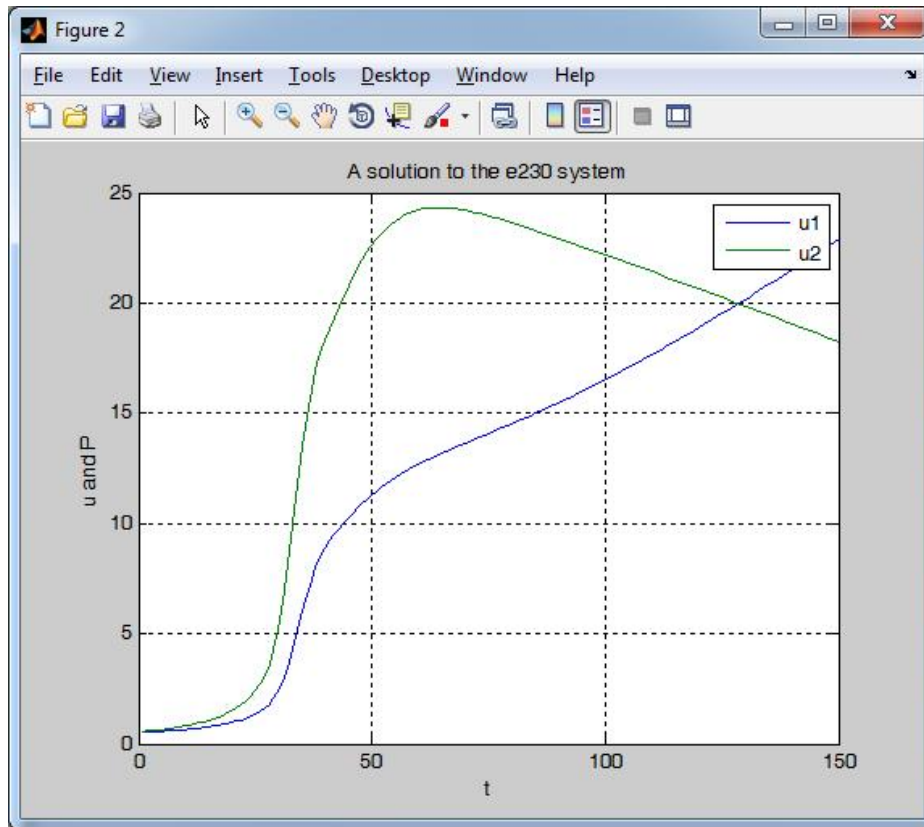


Рис. 2. Динаміка хвойних та листяних порід у центральній точці масиву

Робота математичної моделі та супутнього їй програмного комплексу тестувалася для великої кількості вхідних даних. Було змодельовано вплив недостатньої кількості опадів, перезволоження та різних антропогенних чинників. Порівнюючи отримані результати моделювання за відповідними таблицями ходу росту [1], можна зробити висновки про адекватність отриманих даних.

Суть застосованого підходу в побудові математичної моделі є такою:

- моделювання скеровано на практичне застосування (для аналізу та прогнозу);
- можливість застосування моделі як основи для складнішої просторово-розподіленої моделі, враховуючи локально-екологічну спрямованість;
- завдання, оскільки стосується певної території, ставиться в загальному випадку у вигляді інтегро-диференціальних рівнянь з відповідними граничними умовами. При цьому необхідна можливість постійної корекції початкових умов відповідно до реальних змін, що відбуваються в системі;
- особливого значення надають побудові сценаріїв розвитку екосистеми за того чи іншого зовнішнього впливу (антропогенного чи природного);
- моделювання скеровано на короткостроковий і середньостроковий прогноз. Тут потрібно зазначити, що система з цим набором параметрів існує недовго і параметри, що визначають поведінку моделі, безперервно змінюються.

Висновки

Розроблено практичну модель двовидового середньоширотного лісу та відповідний програмний комплекс. Математична модель адекватно враховує як міжвидову конкуренцію, так й інші фактори: освітлення, заболочення, вікову структуру, кількість опадів, зовнішні впливи та може бути використано як основу для побудови моделей вищого рівня деталізації та територіального охоплення.

Сьогодні ведеться робота над створенням програми для користувача, що враховуватиме характеристики територій і початкові умови, отримані експериментально для певної території. Використовуючи програмний комплекс, наведений у статті, можна апробувати модель на подіях минулих років та отримати можливі сценарії розвитку лісового господарства. Одним із застосувань моделі може стати її використання для паркової реконструкції.

Практичне застосування програмного комплексу, побудованого на реалізації такого роду моделей, у лісовому господарстві України сприятиме сталому розвитку та адаптації механізмів ведення роботи лісгосподарських та бізнес-процесів до зарубіжних стандартів управління.

1. Василюшин Р. Д. *Продуктивність та еколого-енергетичний потенціал лісів Українських Карпат: автореф. дис. на здобуття наук. ступеня д-ра. сільськогосп. наук : 06.03.02 "Лісовпорядкування і лісова таксація"* / Василюшин Роман Дмитрович – Київ, 2014. – 46 с.
2. Заугольнова Л.Б. *Ценопопуляції растений (очерки популяционной биологии)*. – Москва: Наука, 1988 – 183 с.
3. Кузнецов В.И. *Математическое моделирование эволюции леса для целей управления лесным хозяйством* / В.И. Кузнецов, Н.И. Козлов, П.М. Хомяков. – М.: ЛЕНАНД, 2005. – 232 с. – ISBN 5-9710-0027-6.
4. Лотиш І.Л. *Математичні та програмні аспекти моделювання розвитку двовидового лісу* / І.Лотиш // *Науковий вісник національного лісотехнічного університету України – Львів : РВВ НЛТУ України.* – 2014. – Вип. 24.3. – С.358-365. – ISBN 5-7763-2435-1.
5. Чумаченко С.И. *Имитационное моделирование многовидовых разно-возрастных лесных насаждений* : дис. на здобуття наук. ступеня д-ра біолог. Наук – Митищи –2006 - С. 52–53.
6. [Електронний ресурс]. – Режим доступу: http://www.nature.air.ru/biodiversity/book2_2.html.
7. *International Journal of Mathematical and Computational Forestry & Natural-Resource Sciences* [Електронний ресурс]: *Algebraic difference equations for stand height, diameter, and volume depending on stand age and site factors for Estonian state forests* / A.Kiviste1, K. Kiviste . – Режим доступу: http://mcfns.com/index.php/Journal/article/view/MCFNS.1-67/MCFNS-1%3A67_PDF.
8. *International Journal of Mathematical and Computational Forestry & Natural-Resource Sciences* [Електронний ресурс]: *An integrated MCDA software application for forest planning: a case study in southwestern Sweden* / A. Korosuo, L. O. Eriksson. – Режим доступу: <http://mcfns.com/index.php/Journal/article/view/MCFNS.3-75/MCFNS-3%3A75>.
9. [Електронний ресурс]: *Математическое моделирование лесных экосистем : учебное пособие* – Вологда-Молочное – 2012. – Режим доступу: <http://www.booksite.ru/fulltext/rusles/matmodel/text.pdf>

МЕТОДИ І АЛГОРИТМИ СУЧАСНИХ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

УДК 004.05

Є. Гнатчук

Хмельницький національний університет,
кафедра системного програмування

ДОСЛІДЖЕННЯ МЕТОДІВ ФОРМУВАННЯ ВИМОГ ДО ЕКСПЕРТНИХ СИСТЕМ ТЕХНІЧНОГО ДІАГНОСТУВАННЯ

© Гнатчук Є., 2014

Розглянуто особливості експертних систем технічного діагностування як програмних систем. Досліджено методи формування вимог до програмного забезпечення та можливість їх застосування під час формування вимог до експертних систем технічного діагностування.

Ключові слова: вимоги, програмне забезпечення, експертні системи технічного діагностування.

The singularity of expert systems of technical diagnosing as program systems are considered in article. Methods of formation of software requirements and possibility of their use when forming requirements to expert systems of technical diagnosing were probed.

Key words: requirements, software, expert systems of technical diagnosing.

Вступ

Будь-яку програмну систему створюють для розв'язання однієї або декількох задач майбутніх користувачів цієї програмної системи. У зв'язку зі стрімким розвитком комп'ютерної галузі від розробників програмного забезпечення (ПЗ) вимагається скорочення термінів створення програмного продукту, але разом з цим підвищення його якості. Тому розробники ПЗ прикладають багато зусиль для збирання і документування вимог до програмної системи. Недостатній обсяг інформації, який надходить від користувачів, не повністю сформовані вимоги, а також їхні суттєві зміни у процесі розроблення ПЗ – це основні причини, за якими розробники не можуть надати вчасно, а також у рамках бюджету всю заплановану функційність програмної системи.

Проблеми можуть виникати також через неформальне збирання інформації, припущень щодо функційності, помилкових або неузгоджених припущень, недостатньо визначених вимог, безсистемності процесу.

Постановка задачі

Одним з перспективних напрямків у галузі інформаційних технологій є використання експертних систем (ЕС).

Програмні засоби, які ґрунтуються на технології експертних систем, поширені в світі [1]. Мета використання ЕС полягає у розробленні програм, які під час розв'язування задач, важких для експерта-людини, дають результати, що не поступаються за якістю і ефективністю рішенням, одержаним експертом.

Сьогодні технологію експертних систем використовують для розв'язування різних типів задач у найрізноманітніших проблемних галузях, зокрема таких, як діагностування, енергетика,

нафтова і газова промисловість, транспорт, металургія, космос, хімія, телекомунікації і зв'язок, фінанси, освіта, медицина та ін. [2].

Оскільки від ЕС у результаті вимагається достатньо висока якість і ефективність рішень, це значно впливає на процес формування вимог до таких систем. Отже, необхідно дослідити методи формування вимог до ПЗ та оцінити можливість їх застосування у процесі формування вимог до ЕС.

Особливості експертних систем технічного діагностування

Взагалі до ЕС належать системи, які ґрунтуються на знаннях, це системи, обчислювальні можливості яких є, по-перше, наслідком нарощування їх бази знань і, по-друге, визначаються використовуваними методами [1, 2]. Методи інженерії знань, або методи ЕС, значною мірою інваріантні тому, в яких областях вони можуть застосовуватись. Сьогодні ЕС використовуються для розв'язування таких задач: прийняття рішень в умовах невизначеності або неповноти знань, інтерпретація символів та сигналів, діагностика, прогнозування, проектування, планування, керування, контроль тощо.

Задача діагностування є однією з найважливіших для медицини (визначення хвороби людини), техніки (пошук місця несправності технічного об'єкта, а також в апаратному та програмному забезпеченні комп'ютерів), сільського господарства (визначення хвороб рослин і тварин).

Експертні системи технічного діагностування (ЕСТД) мають певні особливості, що відрізняє їх від інших програмних систем і це, своєю чергою, ускладнює процес формування вимог.

У базах знань ЕСТД містяться так звані експертні знання – знання, якими володіє лише вузьке коло спеціалістів.

Однією з рис, що притаманна тільки ЕСТД, є реалізація концепції явного представлення знань, що дає можливість користувачам одержувати пояснення одержуваних рішень на якісному рівні.

Деякі експертні системи здатні вже в процесі їх практичного застосування поповнювати свої знання, якщо встановлено зворотний зв'язок з об'єктом, на який діють прийняті ЕСТД рішення.

Об'єднання технології ЕСТД з технологією традиційного програмування додає нових якостей до програмних продуктів завдяки динамічній модифікації додатків користувачем, а не програмістом, більшої “прозорості” додатка, наприклад, зберігання знань обмеженою природною мовою, що не потребує спеціальних коментарів до знань і спрощує навчання та супровід, кращої графіки, інтерфейсу і взаємодії.

Для того, щоб правильно визначити вимоги, розробники системи повинні розуміти контекст, в якому працює система.

Методи формування вимог до ЕСТД

За більшістю визначень, вимога – це характеристика або умова, яку повинна задовольняти система [3, 4].

Вимоги до програмного забезпечення поділяють за різними рівнями, але загальноприйнятими є розподіл вимог за характером на: функційні та нефункційні [3, 4]. Функційний характер мають вимоги, що визначають поведінку системи, нефункційний – вимоги, що визначають характер поведінки системи. Нефункційні вимоги не пов'язані безпосередньо з функціями, які виконує система. Вони пов'язані з такими інтеграційними властивостями, як надійність, формування коректної відповіді на запит користувача, розмір системи та інше. Для формування вимог до ЕСТД нефункційні вимоги є значиміші та критичніші, ніж окремі функційні вимоги. Це пов'язано з тим, що важливою складовою ЕСТД є база знань, що висуває ряд вимог, пов'язаних з її створенням та наповненням. Також нефункційні вимоги відображають бюджетні та організаційні обмеження компанії-замовника.

Існують серйозні труднощі, які значно уповільнюють та ускладнюють процес формування вимог до ЕСТД.

Перші труднощі виникають у зв'язку з постановкою задачі. Більшість замовників, плануючи розроблення ЕСТД, внаслідок недостатньої компетентності у питаннях застосування методів штучного інтелекту схильна значно перебільшувати очікувані можливості системи. Замовник

відводить ЕСТД роль самостійно мислячого експерта у досліджуваній галузі, який здатен розв'язувати широке коло задач. Звідси і типові попередні постановки задач зі створення ЕСТД: “Розробити ЕСТД з обробки зображень”; “Створити ЕСТД для діагностування приладів навігації” тощо. Для успішного розроблення ЕСТД необхідні не тільки чітка і конкретна постановка задач, але й розроблення докладного (хоча б мовного) опису розрахункового методу її вирішення.

Друга і основна проблема – це набуття (засвоєння) знань. Ця проблема виникає під час “передання” знань, якими володіють експерти-люди, експертній системі. Для того, щоб “навчити” знанням комп'ютерну систему, потрібно сформулювати, систематизувати і формалізувати ці знання. Більшість експертів, успішно використовуючи в повсякденній діяльності свої знання, зазнають значних труднощів при спробі сформулювати і представити у системному вигляді хоча б основну частину цих знань: ієрархію понять, які використовуються; евристики; алгоритми та зв'язки між ними. Виявляється, що для такої формалізації знань необхідні знання в області математичної логіки та методів представлення знань і знання про можливості ЕОМ.

Отже, під час розроблення ЕСТД необхідна участь спеціалістів, які мають вказану сукупність знань та виконують функцію посередників між експертами у предметній області і комп'ютерними (експертними) системами. Вони отримали назву “інженери зі знань” (knowledge engineers). Функції експерта та інженера зі знань рідко поєднуються в одній особі. Часто функції інженера зі знань виконує розробник ЕСТД. Як показав досвід багатьох розробок, для початкового набуття знань, в яких беруть участь експерти, інженери зі знань і розробники ЕС, потрібна активна робота всіх трьох категорій спеціалістів.

Враховуючи існуючі особливості ЕСТД, розглянемо методи формування вимог до ПЗ, а також можливість використання цих методів для формування вимог до ЕСТД. Існує низка методів формування вимог до ПЗ – опорні погляди, сценарії, етнографічний метод, методи структурного та об'єктно-орієнтованого аналізу і методи прототипування [5, 6].

Метод опорних погляду визнає різні погляди учасників формування вимог до системи та використовує їх як основи побудови процесу формування вимог та безпосередньо самі вимоги. На рівні системних вимог погляди можуть відображати призначення самої системи, а також її частин або підсистем, що впливає на її функціонування. Погляди утворюють ієрархію, яка використовується для визначення границь системи і допомагає аналізувати вимоги. Перевагами цього методу є те, що він враховує множину різних поглядів користувачів системи та забезпечує основу для виявлення протиріч у вимогах.

Свою чергою, похідними від цього методу, залежно від трактування поняття “погляд”, є методи SADT, CORE та VORD [5–8]. При використанні методів SADT (Structured Analysis and Design Technique) і CORE (Context Object Request Event) відбираються такі погляди, що визначають дані, які будуть створені та використані при роботі системи, і способи опрацювання цих даних. Методи опорних поглядів можна використовувати як джерело інформації про системні вимоги, а також для виявлення протиріч у вимогах. Але використовувати їх для формування нефункційних вимог до ЕСТД недоцільно, оскільки ці методи не дають змоги дослідити зв'язки між поглядами та типами учасників, що формують вимоги.

Метод VORD (Viewpoint-Oriented Requirements Definition) розглядає погляди, що є зовнішніми до системи. Цей метод використовується для формування нефункційних вимог. Перевагою його є те, що при використанні для формування вимог до ЕСТД є можливість створення ієрархії поглядів, а також їх групувати, що робить зрозумілою структуру системи та визначає зв'язки між підсистемами та із зовнішнім середовищем.

Метод сценаріїв при формуванні вимог є необхідним для деталізації вже сформованих вимог, оскільки описує послідовність роботи користувача з системою [5–7]. Переважно сценарій містить описи стану системи, опис нормального проходження подій, опис виключних ситуацій та способів їх обробки, інформацію стосовно інших дій, які можна здійснювати під час виконання сценарію. Варіанти використання – це метод формування вимог на основі використання сценаріїв. Перевагою цього методу є можливість визначення користувачів системи, а також види взаємодій у системі.

Використовуючи цей метод під час формування вимог до ЕСТД, можна сформувати системні вимоги та вимоги користувача.

Етнографічний метод використовується для розуміння і формування соціальних та організаційних аспектів експлуатації системи [6, 7]. Важливість етнографічного методу полягає у тому, що він допомагає визначити неявні вимоги до системи, які відображають реальні особливості її експлуатації. Недоліком цього методу є труднощі при формуванні всіх вимог до предметної галузі та вимог організаційного характеру. Для розроблення ЕСТД цей момент є дуже важливим, оскільки саме особливості предметної галузі визначають контекст системи, а, отже, список вимог до неї.

Метод прототипування можна використовувати для аналізу ризиків і на початковому етапі розроблення планів керування програмним проектом [7–9]. Доведено, що прототипування зменшує кількість проблем, пов'язаних з розробленням та формуванням вимог. Крім того, прототипування зменшує загальну вартість розроблення системи. Іншими перевагами цього методу є можливість виявлення неповних або неузгоджених вимог. Слід зазначити, що метод прототипування є ефективним для формування системних вимог та вимог користувача. При формуванні вимог до ЕСТД важливо також формувати нефункційні вимоги, отже, окремо метод прототипування використовувати не можна.

Методи структурного та об'єктно-орієнтованого аналізу забезпечують основу для детального формування системних вимог та аналізу вимог [7, 8]. Але ці методи мають такі недоліки: не забезпечують ефективного механізму формування нефункційних системних вимог, у результаті застосування цих методів формується об'ємна документація, що призводить до формування надто деталізованих та неістотних вимог.

Можливості застосування розглянутих методів для формування вимог до ЕСТД наведено у таблиці.

Можливості застосування розглянутих методів для формування вимог до ЕСТД

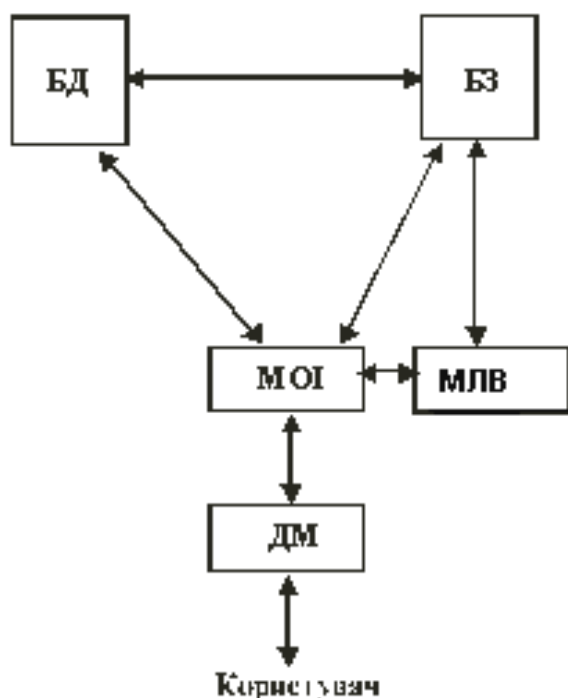
№ з/п	Методи формування вимог до ПЗ	Можливості застосування для формування вимог до ЕСТД	
		Функційні вимоги	Нефункційні вимоги
1.	Методи опорних поглядів: – метод SADT	Використовується для формування вимог	Не дозволяють дослідити зв'язки між поглядами та типами учасників, що формують вимоги
	– метод CORE	Використовується для формування вимог	
	– метод VORD	Використовується для формування вимог	
2.	Метод сценаріїв	Використовується для формування системних вимог та вимог користувача	Використовується для деталізації вже сформованих вимог
3.	Етнографічний метод	Використовується для формування вимог до системи, які відображають реальні особливості її експлуатації	Не забезпечує ефективного механізму формування вимог
4.	Метод прототипування	Використовується для формування системних вимог	Окремо метод прототипування використовувати не можна
5.	Метод структурного аналізу	Використовується для формування системних вимог та вимог користувача	Не забезпечує ефективного механізму формування вимог
6.	Метод об'єктно-орієнтованого аналізу	Використовується для формування системних вимог та вимог користувача	Не забезпечує ефективного механізму формування вимог

Аналізуючи таблицю, можна зробити висновок, що більшість методів формування вимог до ПЗ можна використовувати як окремо, так і як їх комбінацію для формування функційних вимог до

ЕСТД. Що стосується нефункційних вимог, то жоден метод окремо використовувати не можна, враховуючи вимоги до предметної галузі ЕСТД.

Для успішного функціонування ЕСТД необхідно: забезпечити збирання, організацію та зберігання знань про проблемну область; багатокористувальницький спосіб реалізації спілкування користувача з ЕСТД; механізми реалізації машини логічного виведення та побудови логічного ланцюжка, що реалізує виведення; передавання параметрів; виконання функціональних модулів; операційні можливості – інтерфейс із системою програмування, міжмодульний інтерфейс, системні послуги, можливості щодо адаптації до різних ОС і апаратних баз.

Наприклад, розглянемо структуру типової ЕСТД, зображену на рис. 1. Складові системи визначають системні вимоги та вимоги користувача. Формування нефункційних вимог залежить від предметної галузі.



– ДМ – діалоговий модуль, що забезпечує взаємодію користувача з ЕСТД, реалізує опитування користувача у процесі роботи системи, занесення зібраної інформації до тимчасової бази даних та видання результатів діагностування;

– МОІ – модуль опрацювання інформації, що забезпечує організацію та опрацювання інформації, яка з'являється у процесі діагностування;

– МЛВ – модуль логічного виведення, що забезпечує опрацювання діагностичної інформації;

– БЗ – база знань, у якій зберігаються необхідні для роботи ЕСТД знання про предметну галузь;

– БД – база даних, яка є тимчасовою протягом сеансу роботи користувача, і у ній зберігається одержана від користувача та у процесі роботи системи інформація.

Рис. 1. Структура типової ЕСТД

Отже, процес формування вимог до експертних систем технічного діагностування ускладнюється через особливу структуру цих систем, зокрема наявність баз знань та особливості предметної галузі. Ці особливості потрібно враховувати, вибираючи метод формування вимог. Подальших досліджень потребують методи виявлення суперечливості та дублювання вимог, а також оцінювання повноти сформованих вимог.

Висновки

Дослідження методів формування вимог до програмного забезпечення дійшли висновку, що не існує універсального методу формування вимог. Враховуючи особливості експертних систем технічного діагностування, як методи формування функційних вимог доцільно використовувати будь-який з розглянутих методів, а також комбінувати їх. Для формування нефункційних вимог доцільно використовувати метод сценаріїв та метод опорних поглядів.

1. Попов Э. В. Экспертные системы 90-х гг. Классификация, состояние, проблемы, тенденции // *Новости искусственного интеллекта*. – 1991. – № 2. – С. 84–101. 2. Рыбина Г. В. Особенности и принципы построения интегрированных экспертных систем для диагностики сложных технических систем // *Приборы и системы управления*. – 1998. – № 9. – С. 12–16. 3. ГОСТ

28195-89. *Оценка качества программных средств. Общие положения. 4. IEEE Standard Glossary of Software Engineering Terminology / IEEE Std 610.12-1990.* 5. Майерс Г. *Надежность программного обеспечения : пер. с англ. – М. : Мир, 1980. – 360 с.* 6. Липаев В. В. *Программная инженерия. Методологические основы. – М. : ТЕИС, 2006. – 608 с.* 7. Соммервилл И. *Инженерия программного обеспечения : 6-е издание; под ред. А. А. Минько. – М. : Издательский дом “Вильямс”, 2002. – 618 с.* 8. Орлов С.А. *Технологии разработки программного обеспечения. Разработка сложных программных систем : учебник для ВУЗов. – СПб. : Питер, 2004. – 527 с.* 9. Брауде Э. *Технология разработки программного обеспечения. – СПб. : Питер, 2004. – 655 с.*

УДК 536.532 (088.8)

Г. Юрчик

Національний університет “Львівська політехніка”,
кафедра електронних засобів інформаційно-комп’ютерних технологій

СТРУКТУРНО-АЛГОРИТМІЧНИЙ МЕТОД АВТОКАЛІБРУВАННЯ ТЕРМОЕЛЕКТРИЧНИХ ВИМІРЮВАЛЬНИХ КАНАЛІВ ТЕМПЕРАТУРИ В УМОВАХ ЇХ ЕКСПЛУАТАЦІЇ

© Юрчик Г., 2014

Розглянуто метод автокалібрування (комплектної перевірки) термоелектричних вимірювальних каналів (ТВК) температури в умовах їх експлуатації. За методом послідовно формуються калібровані адитивний та мультиплікативний електротеплові впливи безпосередньо на робочий кінець первинного перетворювача (ПП) ТВК. Зареєстровані додаткові результати спостережень вихідного сигналу ТВК дають можливість визначити за формулою контрольовану температуру об’єкта, максимально наближену до реального її значення завдяки автокорегуванню прогресуючих похибок як ПП, так і вторинних перетворювальних ланок ТВК та знайти поправку, яка автоматично вводиться в результати поточних вимірювань температури.

Ключові слова: термоелектричний перетворювач температури, дрейф градувальної характеристики, структурно-часова надлишковість вимірювальних перетворень, автокорекція адитивної і мультиплікативної похибок, точність методу автокалібрування.

The method of auto-calibration (complete calibration) of measuring channels thermoelectric (TEC) temperature in the conditions of use. According method formed successively calibrated additive and multiplicative electrothermal effects directly on the working end of the primary converter (PP) TEC. Logged additional observations Output enable DEC to determine the formula temperature controlled facility as close to the actual value due to its auto- progressive errors of both PP and secondary conversion units TEC and find an amendment that is automatically entered in the results of current measurements of temperature.

Key words: thermoelectric converter temperature drift calibration performance, structural and temporal redundancy measurement conversions, auto-additive and multiplicative errors, the accuracy of the method of auto calibration.

Вступ

У багатьох промислових галузях України, таких як теплова і атомна енергетика, машинобудування, приладобудування, медицина, харчова промисловість, сільське господарство тощо для контролю температурних параметрів різних теплових технологічних процесів (ТП) широко використовуються як аналогові, так і цифрові термоелектричні вимірювальні канали

(ТВК), що входять до багатофункціональних автоматизованих систем управління цими процесами. При цьому висока ефективність та енергоощадність сучасних технологій, пов'язаних із забезпеченням оптимальних режимів функціонування потужного теплосилового обладнання, тепловою обробкою різних речовин та матеріалів і др. повною мірою і безпосередньо залежать від метрологічної надійності засобів вимірювання температури і зокрема ТВК, під якою розуміють здатність засобу вимірювання зберігати нормовану точність не тільки в межах міжповірного інтервалу, але й протягом тривалішого часу їх експлуатації. Як свідчать статистичні дані [1–4] похибка ТВК під впливом різних дестабілізуючих факторів (агресивність контрольованого середовища, високі температури тощо) протягом міжповірного інтервалу може значно перевищувати допустимі значення. Основними джерелами похибки вимірювань температури є відносно швидкий дрейф градувальної характеристики (ГХ) первинного термоелектричного перетворювача (ТЕП), зміна коефіцієнтів перетворення і дрейфу нуля вторинних ланок ТВК, а також дія на них різних дестабілізуючих впливових величин. Водночас чергова періодична повірка ТВК є доволі складною і малоефективною процедурою, яка передбачає обов'язковий демонтаж ТЕП з місця встановлення на технологічному обладнанні з метою його заміни або переградування в лабораторних умовах. А необхідність метрологічних досліджень вторинних перетворювачів ТВК вимагає багато часу і, як правило, зумовлює в багатьох випадках небажану зупинку неперервних теплових ТП. Недоліком цього поелементного методу повірки ТВК є також значна невизначеність у математичній моделі результуючої похибки вимірювання, зумовлена недостатністю експериментальних даних про систематичні похибки, які виникають внаслідок дії на ТВК різних зовнішніх дестабілізуючих факторів.

У зв'язку з вищевикладеними складностями процедури поелементної повірки ТВК, а також невизначеністю оцінки його результуючої похибки у статті розглядається альтернативний метод автокалібрування або комплектної повірки ТВК безпосередньо в умовах їх експлуатації. Висока оперативність методу, можливість автокорекції прогресуючої похибки ТВК в реальному часі підвищують метрологічну надійність ТВК і водночас ефективність сучасних теплових ТП та якість виготовленої продукції.

Метод автокалібрування термоелектричних вимірювальних каналів температури

Метод автокалібрування використовують, вбудовуючи до структури ТВК засоби формування каліброваних адитивних і мультиплікативних електротеплових впливів на робочий кінець ТЕП безпосередньо в умовах експлуатації. Послідовне здійснення цих впливів, протягом часу меншого за теплову інерційність контрольованого об'єкта, відповідно до запропонованого алгоритму функціонування ТВК, дає можливість на основі одержаних додаткових результатів спостережень вихідних сигналів ТВК визначити за формулою значення температури контрольованого об'єкта, максимально наближеного до її реального значення. Цього досягають завдяки алгоритмічній автокорекції не тільки прогресуючої похибки ТЕП, зумовленої відносно швидким дрейфом і нелінійністю його ГХ, дрейфом нуля та зміною коефіцієнтів перетворення вторинних вимірювальних ланок ТВК, але й додаткової похибки внаслідок дії на ТВК в умовах експлуатації різних дестабілізуючих впливових величин.

Структурну схему мікропроцесорного ТВК з автокалібруванням наведено на рис. 1. Метод автокалібрування ТВК полягає в такому. Робочий кінець трьохелектродного ТЕП ТВК знаходиться в контрольованому середовищі, температуру якого в межах ($T_{x\min}$ – $T_{x\max}$) необхідно вимірювати. Три вільні кінці ТЕП розміщено в коробці 4 із сталюю температурою T_0 . Перемикачем 5 перекомутують вільні кінці, утворюючи два варіанти ТЕП з одним спільним електродом, тобто 1.1–1.2 та 1.1–1.3. У процесі пусконаладжувальних робіт на технологічному об'єкті здійснюють первинне калібрування ТВК за відомої температури T_K , яку обирають за співвідношенням:

$$T_K = \sqrt{T_{x\max} \cdot T_{x\min}} \quad (1)$$

Перетворюють за допомогою АЦП7 термоелектрорушійну силу (ТЕРС) на виході ТЕП (електроди 1.1–1.2) на цифровий код

$$N_1' = S(e_{n1}T_K - e_{01}T_0), \quad (2)$$

де S – крутизна характеристики АЦП7; e_{n1} – початкове значення термоелектричного коефіцієнта робочого кінця основного ТЕП (1.1-1.2) при температурі калібрування T_K ; e_{01} – термоелектричний коефіцієнт вільних кінців ТЕП при температурі T_0 , який є фактично незмінним протягом тривалого часу експлуатації ТВК і визначається з ГХ ТЕП.

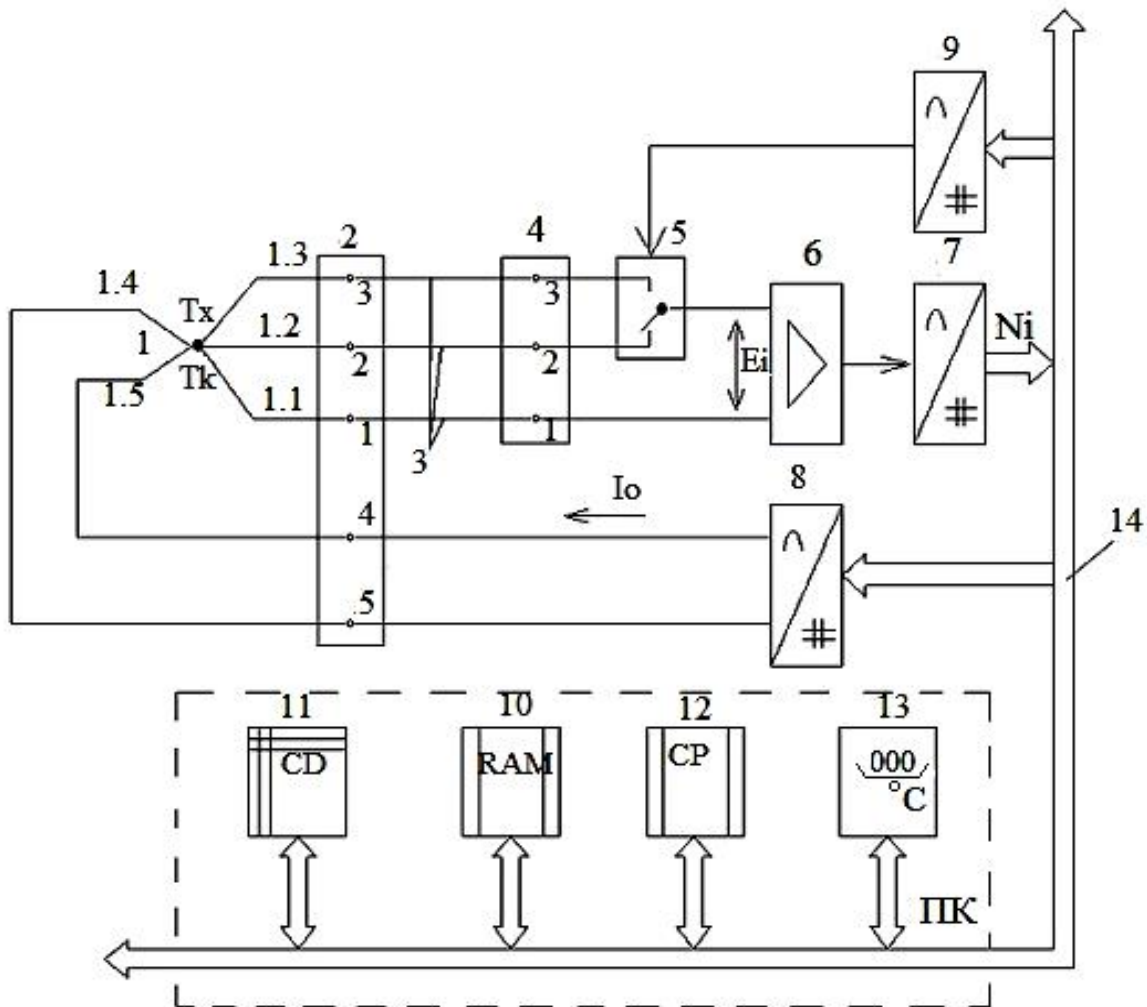


Рис. 1. Структурна схема мікропроцесорного термоелектричного вимірювального каналу температури з автокалібруванням: T_x – контрольована температура об'єкта;

T_K – відоме значення температури робочого кінця ТЕП при первинному калібруванні ТВК;

1 – подвійний ТЕП з вимірювальними електродами 1.1, 1.2 і додатковим електродом 1.3; 1.4

і 1.5 – електроди, що використовуються для адитивного електротеплового впливу на робочий кінець ТЕП;

2 – клемна колодка ТЕП; 3 – подовжувальні компенсаційні дроти; 4 – коробка стабілізації або компенсації змін температури вільних кінців ТЕП; 5 – керований перемикач; 6 – нормуючий підсилювач сигналу ТЕП;

7 – аналого-цифровий перетворювач (АЦП); 8, 9 – цифро-аналогові перетворювачі постійного струму (ЦАП);

10 – блок оперативної пам'яті; 11 – пристрій управління; 12 – обчислювальний пристрій; 13 – цифровий індикатор контрольованої температури T_x ; 14 – загальна шина персонального комп'ютера ПК

Цифровий код ТЕРС вільних кінців ТЕП (1.1–1.2) визначається за співвідношенням

$$N_{01} = S e_{01} T_0 = S e_{n1} T_K - N_1' . \quad (3)$$

За рівнянням (3) одержуємо вираз для початкового значення термоелектричного коефіцієнта робочого кінця ТЕП (1.1-1.2)

$$e_{n1} = \frac{N_{01} + N_1'}{ST_K} \quad (4)$$

Цифрові коди N_1' та N_{01} реєструють і запам'ятовують. Далі змінюють чутливість ТЕП шляхом підключення до його робочого кінця, замість основного електрода 1.2, однотипного з ним електрода 1.3, який виконано так, що термоелектричний коефіцієнт його “гарячого” кінця в $K_1 = 0,9-0,95$ разу менший ніж термоелектричний коефіцієнт робочого кінця електрода 1.2. Такого виконання електрода 1.3 досягають, наприклад, штучним старінням його “гарячого” кінця. Зміна чутливості ТЕП призводить до зменшення ТЕРС на його виході, цифровий код якої матиме значення:

$$N_2' = S(e_{n2}T_K - e_{01}T_0) = S\left[e_{n1}\left(1 - \frac{\Delta e'}{e_{n1}}\right)T_K - e_{01}T_0\right] = S(e_{n1}K_1T_K - e_{01}T_0), \quad (5)$$

де ϵ_{n2} – коефіцієнт чутливості робочого кінця утвореного ТЕП (електроди 1.1-1.3); $\Delta e' = (e_{n1} - e_{n2})$

Одержаний код N_2' реєструється і запам'ятовується. Потім до вимірювальної схеми під'єднується, замість додаткового електрода (1.3), основний електрод (1.2). На цьому процедура первинного калібрування ТВК перед початком його тривалої експлуатації завершується.

При здійсненні процедури автокалібрування ТВК в процесі його експлуатації спочатку одержують цифровий код ТЕП, що відповідає контрольованій температурі T_x :

$$N_1'' = S[(e_{n1} + \Delta e_{x1})T_x + \Delta e_1 - \Delta e_{01}T_0], \quad (6)$$

де Δe_{x1} – неконтрольовані зміни термоелектричного коефіцієнта робочого кінця ТЕП внаслідок його поступової фізичної деградації, а також через відмінність температури T_x від температури первинного калібрування ТВК; Δe_1 – адитивна складова інформативного сигналу, зумовлена, по-перше, появою паразитної ТЕРС внаслідок розвитку структурної неоднорідності електродів ТЕП за їх довжиною і накладання на них градієнтного температурного поля, а, по-друге, – дрейфом нуля підсилювача б.

Цифровий код N_1'' реєструють і запам'ятовують у ПК. З урахуванням виразу (3) цифровий код (6) можна подати у вигляді :

$$N_1'' = Se_{n1}T_x + S\Delta e_{x1}T_x + S\Delta e_1 - Se_{01}T_0 = Se_{n1}T_x + \Delta N_1 + \Delta N_2 - N_{01}, \quad (7)$$

де $\Delta N_1 = S\Delta e_{x1}T_x$ – мультиплікативна похибка вимірювання в цифровому коді від зміни чутливості робочого кінця ТЕП; $\Delta N_2 = S\Delta e_1$ – адитивна похибка вимірювання в цифровому коді, зумовлена розвитком неоднорідності електродів ТЕП, а також дрейфом нуля підсилювача б.

З виразу (7) випливає, що дійсне значення температури, яка контролюється, можна визначити за формулою :

$$T_x = \frac{N_1'' + N_{01} - \Delta N_1 - \Delta N_2}{Se_{n1}} \quad (8)$$

Похибки вимірювання ΔN_1 та ΔN_2 невідомі, і їх вплив на результат визначення температури можна зменшити, якщо ввести компенсуючі поправки, що мають зворотні знаки до знаків похибки. Для знаходження поправок вимірюють ТЕРС ТЕП (1.1–1.3), утвореного за допомогою додаткового третього електрода (1.3) шляхом перекомутації вільних кінців ТЕП. Цифровий код на вільних кінцях ТЕП (1.1–1.3) матиме значення :

$$N_2'' = S[(e_{n1} + \Delta e_{x2})T_x + \Delta e_2 - e_{01}T_0] = S[(e_{n1} + \Delta e_{x1})T_x K_2 + \Delta e_2 - e_{01}T_0], \quad (9)$$

де $K_2 = \left(1 - \frac{\Delta e''}{e_{n1} + \Delta e_{x1}}\right)$; $\Delta e'' = (e_{n1} + \Delta e_{x1}) - (e_{n2} + \Delta e_{x2})$, Δe_{x2} зміни термоелектричного коефіцієнта робочого кінця ТЕП (1.1-1.3).

Враховуючи, що умови фізичної деградації робочих кінців ТЕП(1.1-1.3) і (1.1-1.2) практично однакові, можна стверджувати, що $\Delta e_{x1} = \Delta e_{x2}$. Цифровий код N_2'' реєструють і запам'ятовують в ПК. Потім відповідно до алгоритму калібрування ТВК додатково нагрівають робочий кінець ТЕП, використовуючи термоелектричний ефект Пельтьє[5]. Для цього через робочий кінець ТЕП (1.1-1.3) за допомогою другої пари електродів 1.4 і 1.5 подвійного ТЕП пропускають протягом визначеного часу t оптимальний постійний струм I_0 певного напрямлення. Температура робочого кінця ТЕП зростає на величину $\Delta T = 4 - 5^\circ C$. Цифровий код ТЕРС на вільних кінцях ТЕП (1.1-1.3) набуде значення:

$$N_3'' = S[(e_{n1} + \Delta e_{x1})(T_x + \Delta T)K_2 + \Delta e_2 - e_{01}T_0], \quad (10)$$

де $\Delta T = \frac{PI_0}{cm}nt$ – перегрів робочого кінця ТЕП відносно температури контрольованого об'єкта; P – коефіцієнт Пельтьє робочого кінця ТЕП; c і m – усереднені питома теплоємність і маса робочого кінця ТЕП; $t = nt$ – час нагрівання робочого кінця ТЕП; t – теплова стала робочого кінця ТЕП; $n=(0,1 \neq 0,2)$. Цифровий код N_3'' реєструють та запам'ятовують у ПК.

Після реєстрації і запам'ятовування цифрового коду N_3'' миттєво під'єднують до вимірювальної схеми замість додаткового електрода 1.3 основний електрод 1.2, відновлюючи при цьому початкове значення чутливості ТЕП ($e_{n1} + \Delta e_{x1}$) і одержують відповідне значення цифрового коду ТЕРС на вільних кінцях ТЕП

$$N_4'' = S[(e_{n1} + \Delta e_{x1})(T_x + \Delta T) + \Delta e_1 - e_{01}T_0]. \quad (11)$$

Цифровий код N_4'' також реєструють і запам'ятовують. Враховуючи, що адитивна похибка вимірювання для двох ТЕП варіантів (1.1)-(1.2) і (1.1)-(1.3) однакова, можна вважати справедливою рівність $\Delta e_1 = \Delta e_2 = \Delta e$, що дає можливість за допомогою одержаних додаткових результатів спостережень визначити поправки на адитивну і мультиплікативну похибки за формулами:

$$\Delta N_1 = \frac{(N_1'' - N_2'')(N_1' + N_{01})}{(N_1' - N_2')} - \frac{T_x}{T_K}(N_1' + N_{01}); \quad (12)$$

$$\Delta N_2 = N_1'' + N_{01} - \frac{(N_1'' - N_2'')(N_4'' - N_1'')}{(N_4'' - N_3'') - (N_1' - N_2'')}. \quad (13)$$

З врахуванням одержаних значень поправок ΔN_1 і ΔN_2 визначається реальне значення контрольованої температури відповідно до виразу (8) за формулою:

$$T_x = \left[\frac{N_1'' + N_{01}}{N_1' + N_{01}} + \frac{1}{2} \left(\frac{N_1'' - N_2''}{N_1' - N_2'} - M \right) \right] T_K, \quad (14)$$

$$\text{де } M = \frac{(N_1'' - N_2'')(N_4'' - N_2'')}{[(N_4'' - N_3'') - (N_1' - N_2'')](N_1' + N_{01})}$$

– алгоритмічний коефіцієнт.

Блок-схему алгоритму функціонування термоелектричного вимірювального каналу температури з автокалібруванням наведено на рис. 2.

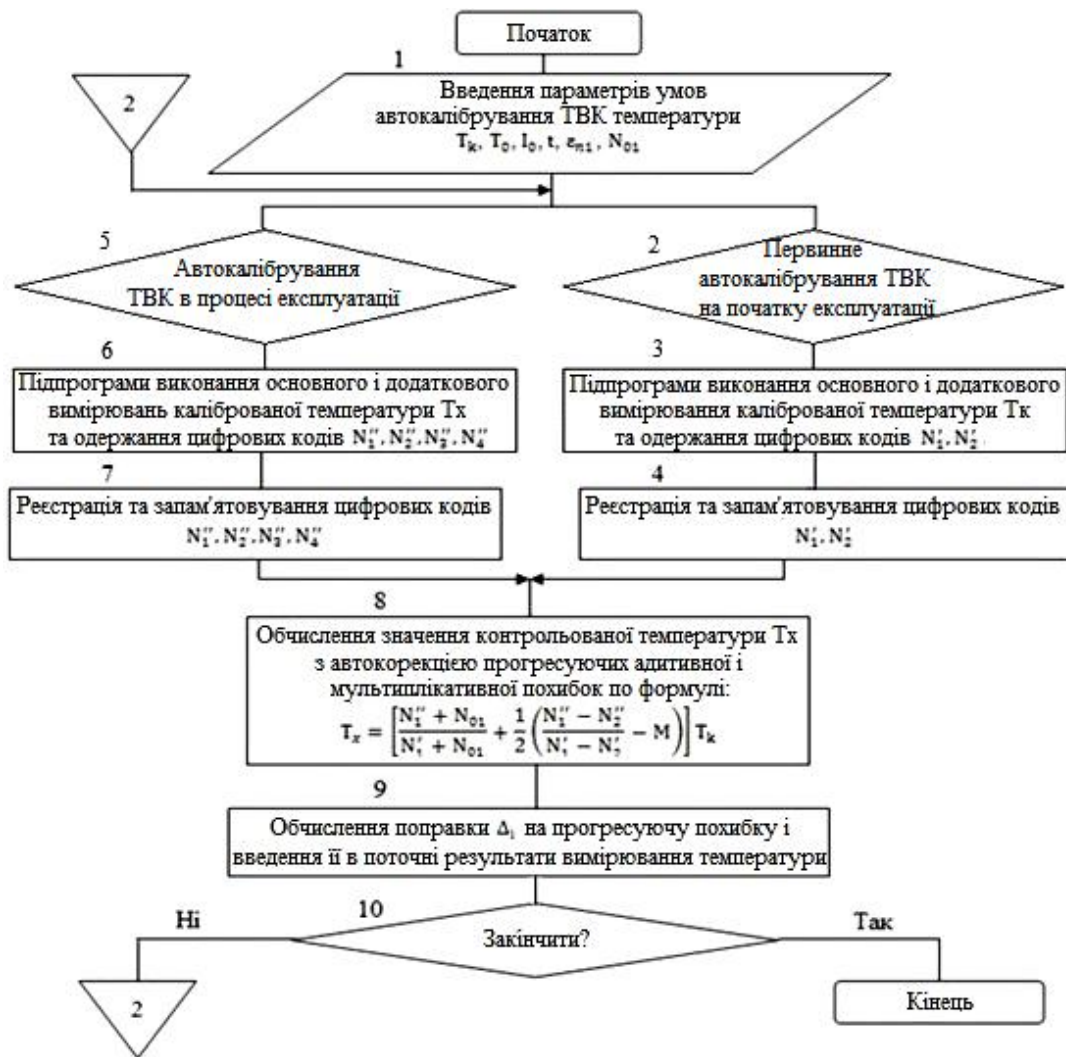


Рис.2. Блок-схема алгоритму функціонування термоелектричного вимірювального каналу температури з автокалібруванням

Висновки

Отже, при автокалібруванні ТВК одержують значення контрольованої температури об'єкта, максимально наближене до реального значення, яке не залежить від нелінійності і прогресуючого дрейфу ГХ ТЕП, дрейфу нуля і зміни коефіцієнта перетворення підсилювача інформативного сигналу, дії на ТВК різних дестабілізуючих факторів, і визначається тільки цифровими кодами результатів рівноточних вимірювальних перетворень. Похибка методу автокалібрування ТВК переважно визначається похибкою формування каліброваних електротеплових впливів на ТЕП і знаходиться на рівні $\pm 1,0\%$ у заданному робочому діапазоні вимірюваних температур. Значення поправки до поточних результатів вимірювання температури обчислюється за виразом :

$$\Delta i = T_x - T_{xi}, \quad (15)$$

де T_x – значення температури, одержане за виразом (14) під час чергового калібрування ТВК; T_{xi} – поточний результат прямого вимірювання температури до моменту калібрування ТВК.

1. Саченко А.А. Совершенствование методов измерения температуры / А.А. Саченко, Е.Я. Твердый. – К.: Техника, 1983. – 101 с. 2. Ярышев Н.А. Теоретические основы измерения нестационарной температуры / Н.А. Ярышев. – Л.: Энергоатомиздат; Ленинградское отделение, 1990 – 255с. 3. Термоелектричні термометри та їх метрологічне забезпечення : навч. посібник / М.П. Березненко, Ю.О. Скрипник, Г.І. Хімічова, Л.О. Глазков, Г.В. Юрчик. – К.: ІСДО, 1994 – 204 с.

4. Геращенко О.А. Температурные измерения / О.А. Геращенко, А.Н. Гордов, В.И. Лах и др. – К.: Наукова думка, 1989. – 494с. 5. Термоэлементы и термоэлектрические устройства: Справочник / Л.Н. Анатычук. – К.: Наукова думка, 1979. – 768с.

UDK 004.932.2

S. Bogucharskiy, V. Mashtalir

Kharkiv National University of Radio Electronics,
Informatics department

ON MATRIX MODIFICATION OF CLARANS CLUSTERING METHOD IN LARGE VIDEO SURVEILLANCE DATABASES

© Bogucharskiy S, Mashtalir V., 2014

Розглянуто алгоритми кластеризації для дуже великих баз даних (VLDB) при використанні в обробці зображень і відео. Виняткова галузь застосування потребує представлення вихідних даних у вигляді багатовимірних векторів. Саме тому запропоновано матричні модифікації традиційних методів k -медоїдів, розбиття навколо медоїдів (PAM), кластеризації великих додатків (CLARA), а також CLARA на основі довільного пошуку (CLARANS). Проаналізовано переваги та недоліки цих методів.

Ключові слова: обробка відео, VLDB, k -медоїди, PAM, CLARA, CLARANS.

Clustering algorithms for Very Large Data Bases (VLDB) are observed in application with image and video processing. Such a specific case requires initial data presentation as multidimensional vectors. That is why matrix modifications of traditional k -medoids, Partitioning Around Medoids, Clustering LARGE Applications and CLARA based on RANDOMized Search methods are proposed. Benefits and drawbacks of them all are examined.

Key words: video processing, VLDB, k -medoids, PAM, CLARA, CLARANS.

Introduction

Multidimensional data clustering aims at finding so called groups (classes, clusters, segments) of surveyed objects in information arrays being analyzed, which is an important part of data mining [1-5]. Its results are widely used for many applications. At the same time, there are quite a lot of practical tasks connected with processing of varied media, logical data, text and images, but the most popular and well-studied methods of cluster analyses become ineffective because of huge amounts of data needed to be processed.

Clustering algorithms for very large databases (VLDB) [1, 5], which form a separate branch in cluster analyses, turn out to be ahead of others for solving such kind of problems. One of the prime algorithms of such a type is CLARANS (Clustering Large Applications based on RANDOMized Search) [6] based on well-known k -medoids method, PAM (Partitioning Around Medoids) [1] and CLARA (Clustering LARGE Applications) algorithms, which are quite effective for processing not very large data arrays.

For the most part of existing clustering algorithms it should be noted that data to be processed are defined as multidimensional vectors forming traditional data table or multidimensional vector sequences (in case data are sequentially obtained for processing). At the same time, initial video processing information is usually presented in a matrix form, in which case this form may contain the whole image or its fragment. It is clear that in order to use some existing clustering methods for image processing, initial image should be vectorized somehow and devectorized to initial form after its processing.

The article presents CLARANS method modification that copes with initial data in a form of $(m \times n)$ -matrix $x(k) = \{x_{i_1 i_2}(k)\}$, $i_1 = 1, 2, \dots, m$; $i_2 = 1, 2, \dots, n$; $k = 1, 2, \dots, N$ is a number of observation in the analyzed data array. By doing so, it is considered that the whole of this array should be divided into p clusters, each of which may be described based on its centroid $c(l)$, $l = 1, 2, \dots, p$ defined also in a form of $(m \times n)$ -matrix.

Matrix modification of k -medoids method

As the classical k -means clustering method [5] is closely related to least squares criterion and Euclidean metric, k -medoids method is based on a matrix metric and least modules criterion which ensures its robustness to outliers. Suppose $x(k)$ and $w(l)$, then there exist $(n \times 1)$ -vectors with the following distance between them:

$$D(x(k), c(l)) = \|x(k) - c(l)\|_1 = \sum_{i=1}^n |x_i(k) - c_i(l)|, \quad (1)$$

and the medoid is selected among available observations, for which the sum of distances within clusters (1) is minimal. It is not essentially hard to calculate the medoid-center, which components are the medians of corresponding components of each cluster elements.

At the same case, when $x(k)$ and $c(l)$ are $(m \times n)$ -matrices, it is suggested to use the following distance instead of (1).

$$D(x(k), c(l)) = \sum_{i_1=1}^m \sum_{i_2=1}^n |x_{i_1 i_2}(k) - c_{i_1 i_2}(l)| = I_m^T |x(k) - w(l)| I_n \quad (2)$$

where I_m, I_n are $(m \times 1)$ - and $(n \times 1)$ -vectors, formed by unities $|x(k) - w(l)| = \{|x_{i_1 i_2}(k) - w_{i_1 i_2}(l)|\}$.

The process of finding medoids-centers can be implemented as follows:

i). Randomly select p observations $x(k)$ from data array and set them as initial medoids-centers $c_{(0)}(l), l = 1, 2, \dots, p$.

ii). Assign the rest $N - p$ observations to the cluster with centroid that is nearest to each observation according to (2).

iii). In each cluster adjust all components of observations $x_{i_1 i_2}(k)$ in increasing order and find the median for each component $med x_{i_1 i_2}(k)$.

iv). Form a new centroid for each cluster $c_1(l) = \{med x_{i_1 i_2}(k)\}$.

v). Repeat steps iii), iv) until

$$D(c_t(l), c_{t+1}(l)) \leq e \quad \forall l. \quad (3)$$

vi). After fulfillment of (3), set $c_{t+1}(l) = c(l)$ and calculate radius of each cluster as follows

$$R(l) = \arg \max_k D(c(l), x(k) \in c(l)). \quad (4)$$

The above method is efficient for not large N , but for very large data volumes enormous calculations for each observation in connection to the medoids reduce the method usage.

Matrix modification of partitioning around medoids

PAM method is more effective from computational point of view, as it does not calculate centroids, but sets existing observations to medoids of clusters.

To find p medoids, PAM starts from arbitrary selection of p image-matrices, much like it is done in k -medoids method. Further, at each iteration of the method, replacement of selected observation $x(k)$ is made to not yet selected image $x(q)$ from available sample set, until clusters will be eventually formed.

To estimate the replacement of $x(k)$ to $x(q)$, cost C_{rkq} is calculated at each iteration for all the rest of the data $x(r)$. In such a case, the cost is a distance function between observations

$$D(x(k), x(q)) = I_m^T |x(k) - x(q)| I_n. \quad (5)$$

During the replacement four different situations may occur.

i). Suppose $x(r)$ belongs to the cluster with medoid $x(k)$ at some iteration of the algorithm. Let $x(r)$ be closer by distance (5) to $x(t)$, than $x(q)$, i.e.

$$D(x(r), x(q)) \geq D(x(r), x(t)) \quad (6)$$

where $x(t)$ is the second medoid close to $x(r)$. Thus, if $x(k)$ is replaced to $x(q)$ as a medoid, $x(r)$ moves to cluster presented by $x(t)$. In such a case, the cost of replacement can be calculated as follows

$$C_{rkq} = D(x(r), x(t)) - D(x(r), x(k)), \quad (7)$$

and it is non-negative.

ii). Suppose $x(r)$ belongs to cluster presented by $x(k)$, with this $x(r)$ is 'less common' to $x(t)$ rather than to $x(q)$, i.e.

$$D(x(r), x(q)) < D(x(r), x(t)). \quad (8)$$

Then, if $x(k)$ is replaced to $x(q)$, $x(r)$ will be related to cluster presented by $x(q)$. Thus, the cost of replacement equals to

$$C_{rkq} = D(x(r), x(q)) - D(x(r), x(k)). \quad (9)$$

Unlike (7), it may be positive and negative as well, depending on image similarity to $x(r)$: $x(k)$ or $x(q)$.

iii). Suppose $x(r)$ does not belong to cluster presented by $x(k)$, and $x(t)$ is a representative of this cluster. Further, let $x(r)$ be closer to $x(t)$ than to $x(q)$. Then, even if $x(k)$ is replaced to $x(q)$, $x(r)$ remains in cluster presented by $x(t)$ anyway. In this circumstances, the cost is

$$C_{rkq} = 0. \quad (10)$$

iv). Suppose $x(r)$ belongs to cluster presented by $x(t)$, but $x(r)$ is less closer to $x(t)$ than to $x(q)$. Then, the replacement of $x(k)$ to $x(q)$ leads to transfer of $x(r)$ to cluster $x(q)$ from $x(t)$. In this circumstances, the cost equals

$$C_{rkq} = D(x(r), x(q)) - D(x(r), x(t)) \quad (11)$$

and it is always negative.

Combining all the four cases, it is clear that the total cost of the replacement of $x(k)$ to $x(q)$ is

$$TC_{kq} = \sum_r C_{rkq}. \quad (12)$$

This algorithm can be presented by the schema presented on Fig. 1.

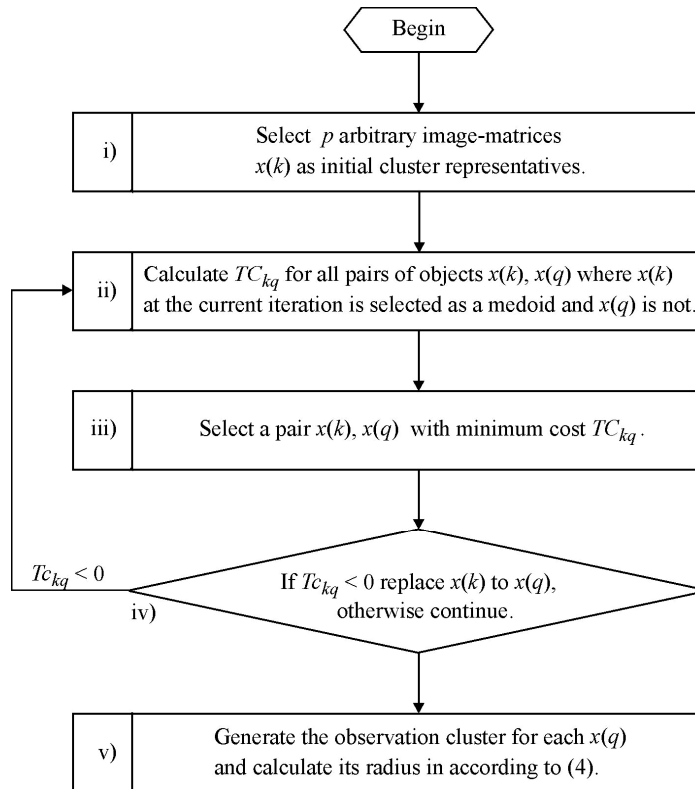


Fig.1. Method of partitioning around medoids

Although PAM method does not define medoids-centers, its implementation requires computations of distances, which makes it ineffective for quite large databases.

CLARA method

CLARA is an extension of PAM designed for large sample collections. Instead of searching for medoids of clusters in the whole sample set of N observations, this method finds medoids in each subsample that is much less by volume. For this purpose, several randomly selected subsamples are processed. The authors of this method have experimentally shown [7] that for the most part of tasks solved by them it was sufficient to form five subsamples containing $40+2p$ observations in each. This algorithm can be presented by the following schema (see Fig. 2).

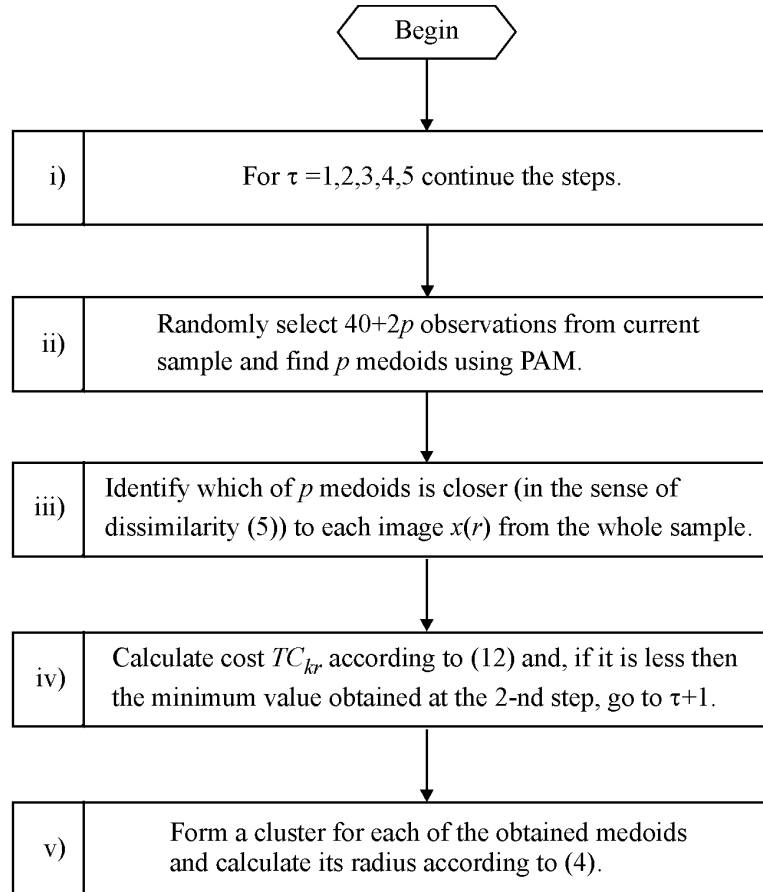


Fig. 2. CLARA method

The authors of the method have shown its effectiveness for sample collections with nearly 1000 observations (that is not that much for video samples). Though there is a risk of losing some information contained in analyzed observations.

CLARANS modification

CLARANS method is based on PAM and CLARA. Clustering process is treated as searching on the graph $G_{N,p}$, each of its nodes corresponds to one of p medoids. Cost is assigned for each of the nodes, which is defined by the distance between each of $N-p$ analysed images and current medoid values and calculated using equation (12).

It should be mentioned that earlier described PAM method also can be treated as searching $G_{N,p}$ graph minimum. With this, all the neighbours of current node are analyzed at each step (equations (5)-(12)), and then the current node is replaced by an observation that provides maximum decrease of the cost.

Such process continues until global minimum is reached. Thus, at each iteration it is necessary to analyze $p(N - p)$ distances, which is too difficult for large N .

Apparently, CLARA analysis takes less images, and it is restricted to a subgraph of less dimensionality $G_{(40+2p)}$, $p < G_{N, p}$, though, as it has been mentioned above, there is a risk of losing data with important information.

CLARANS does not analyze each image at each node. Also it is not restricted to a subgraph only. For each node it randomly forms an area of local neighbourhood.

Clustering process based on modified CLARANS method can be implemented in the following sequence of steps.

- i). Specify method parameters 'numlocal' and 'maxneighbor', assign $i = 1$ and define 'minconst'.
- ii). Select arbitrary graph node $G_{N, p}$ and specify it as a 'current' one.
- iii). Set $j = 1$.

iv). Randomly select an image-neighbour S for the 'current' and according to equation (12) count costs for the 'current' and S .

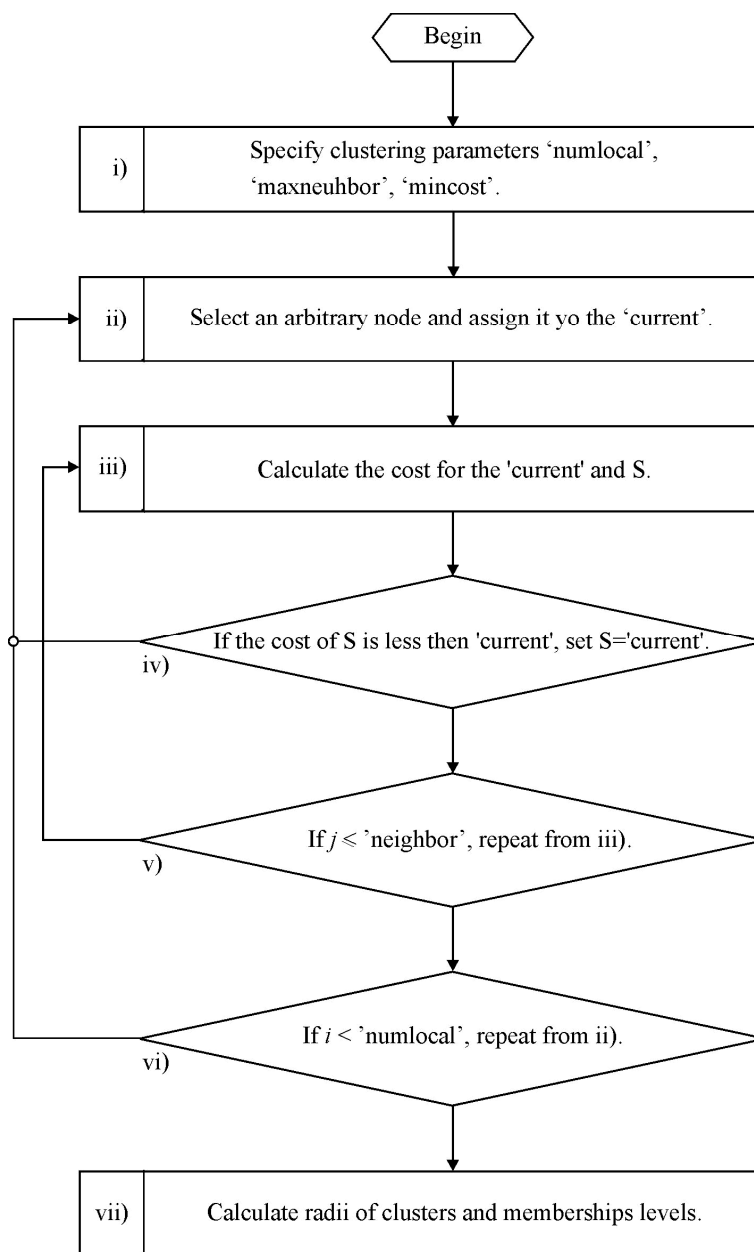


Fig.3. Modified CLARANS method

- v). If the cost of S is lower, set S as ‘current’ and repeat from step iii).
- vi). Otherwise increase j to $(j+1)$ and if $j \leq \text{‘neighbor’}$, repeat from step iv).
- vii). Otherwise if $j > \text{‘neighbor’}$, compare the cost of ‘current’ with ‘minconst’. If the cost of ‘current’ < ‘minconst’ assign ‘minconst’=‘current’.
- viii). Increase i by 1 ($i = i + 1$), and if $i < \text{‘numlocal’}$ finish, otherwise repeat from step ii).

After $p = \text{‘numlocal’}$ clusters are formed, radius of each cluster can be calculated in addition according to equation (4). Moreover, as clusters almost always overlap in real-world applications (especially in image processing), levels of membership image to each cluster can be calculated. With this, if the distance between observation $x(k)$ and medoid $c(l)$ equals $D(x(k), c(l))$, then the level of belonging $x(k)$ to $c(l)$ can be defined using the equation

$$m(x(k), c(l)) = \frac{D^{-1}(x(k), c(l))}{\sum_{q=1}^p D^{-1}(x(k), c(q))}. \quad (13)$$

Data processing algorithm based on modified CLARANS may be presented by the scheme shown on Fig. 3.

It is clear that clustering effectiveness depends greatly on the correct selection of parameters ‘maxneighbor’ and ‘minconst’, that requires users’ qualification. It is also clear that the more ‘maxneighbor’ value is (under the limit $N - p = N - \text{‘numlocal’}$), the more the process is alike to PAM and the more complicated it is. Thus, subjectivity level here is quite big, which is actually always present in tasks connected with self-learning paradigm, though it does not prevent finding acceptable solutions.

Conclusion

Modification of CLARANS method has been proposed, which is designed for clustering in very large databases (VLDB). The proposed modification copes with processing in a matrix form, and fuzzy clustering is applied for situations with overlapping clusters. The method is simple in implementation and does not need any complicated computations.

1. Han J., Kamber M. *Data Mining: Concepts and Techniques*. – 2-nd ed. – San Francisco: Morgan Kaufmann, 2006. – 800 p. 2. Gan G., Ma C., Wu J. *Data Clustering: Theory, Algorithms, and Applications*. – Philadelphia: SIAM, 2007. – 466 p. 3. Abonyi J., Feil B. *Cluster Analysis for Data Mining and System Identification*. – Basel: Birkhäuser, 2007. – 303 p. 4. Olson D.L., Dursun D. *Advanced Data Mining Techniques*. – Berlin: Springer, 2008. – 180 p. 5. Xu R., Wunsch D.C. *Clustering*. – Hoboken: John Wiley&Sons, 2008. – 358 p. 6. Kohonen T. *Self-Organizing Maps*. – 1-st ed. – Berlin: Springer, 1995. – 501 p. 7. Ng R.T., Han J. *Efficient and clustering methods for spatial data mining // 20-th Int. Conf. on Very Large Data Bases*. – Santiago de Chile, 1994. P.144-155. 8. Kaufman L., Rousseeuw P.J. *Finding Groups in Data: An Introduction to Cluster Analysis*. – N.Y.: John Wiley&Sons, 1990. – 342 p.

ПОБУДОВА ТА ФУНКЦІОНУВАННЯ СИСТЕМ ЕЛЕКТРОННОЇ КОНТЕНТ-КОМЕРЦІЇ

© Висоцька В., Чирун Л., 2014

Проаналізовано основні проблеми електронної контент-комерції та функціональних сервісів опрацювання контенту. Запропонована модель дає можливість створити методи та засоби опрацювання інформаційних ресурсів у системах електронної контент-комерції та реалізувати процеси формування, управління та супроводу контенту.

Ключові слова: інформаційний ресурс, комерційний контент, контент-аналіз, контент-моніторинг, контентний пошук, система електронної контент-комерції.

General problems of electronic content commerce and functional services of content processing are analyzed in the article. Proposed model gives an opportunity to create a method and instrument of information resources processing in electronic commerce systems and to implement the processes of content formation, management and support.

Key words: information resources, commercial content, content analysis, content monitoring, content search, electronic content commerce systems.

Вступ. Загальна постановка проблеми

Розроблення і впровадження систем електронної контент-комерції є одним із важливих напрямів розвитку електронного бізнесу. Характерною рисою таких систем є можливість автоматичного опрацювання інформаційних ресурсів для збільшення обсягів продажу контенту постійному користувачу, активного залучення потенційних користувачів та розширення меж цільової аудиторії [1, 2]. Зокрема, принципи і технології електронної контент-комерції активно застосовують при створенні систем on-line/off-line продажу та аналізу/обміну/збереження контенту, Інтернет-магазину, cloud storage/computing [2]. Проблематика проектування, створення, впровадження та супроводу систем електронної контент-комерції є актуальною з огляду на такі фактори, як активний розвиток досліджень в галузі е-бізнесу та відсутність теоретичного обґрунтування стандартизованих методів і потреба в уніфікації програмних засобів опрацювання інформаційних ресурсів у таких системах [1–3]. Постійно з'являються нові підходи та способи вирішення цієї проблеми, але важливим питанням залишається невідповідність між відомими методами і засобами опрацювання інформаційних ресурсів та принципами побудови систем електронної контент-комерції. Відсутні загальні підходи до створення систем електронної контент-комерції та уніфіковані методи опрацювання інформаційних ресурсів у цих системах [2, 3]. Фахівці із проектування, реалізації та впровадження систем електронної контент-комерції мають справу із різними процесами опрацювання інформаційних ресурсів, які сприяють досягненню мети щодо збільшення обсягів продажу контенту постійному користувачу, активному залучення потенційних користувачів та розширення меж цільової аудиторії [1–5].

Формулювання мети

Метою роботи є побудова процесів опрацювання інформаційних ресурсів у системах електронної контент-комерції для визначення функціональних вимог до підсистем формування, управління та супроводу комерційного контенту. Відповідно це передбачає моделювання кожного з видів дій, які реалізують дані підсистеми з подальшим об'єднанням та

узагальненням їх в єдиній узгодженій моделі, яка описує вимоги до процесів функціонування систем електронної контент-комерції загалом.

Аналіз отриманих наукових результатів

Основними процесами опрацювання інформаційних ресурсів у системах електронної контент-комерції є формування, управління та супровід контенту (рис. 1), схема зв'язків яких є такою: *формування контенту* ® *управління контентом* ® *супровід контенту*.

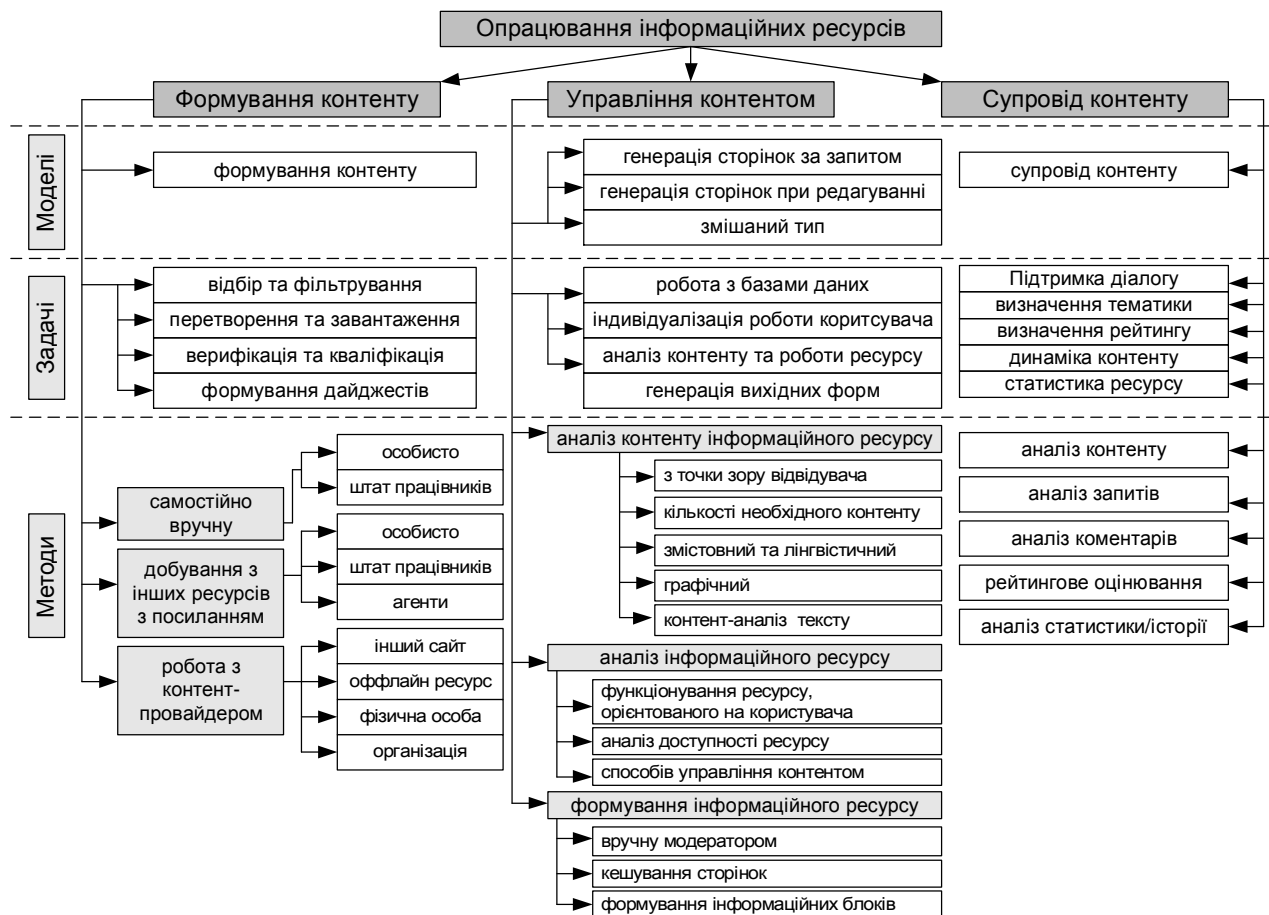


Рис. 1. Особливості процесу опрацювання інформаційних ресурсів у системах електронної контент-комерції

Функціонування системи електронної контент-комерції описується такими схемами зв'язків основних компонентів цієї системи:

1) для процесу формування інформаційного ресурсу системи: *контент* ® *формування комерційного контенту* ® *база даних комерційного контенту* ® *управління комерційним контентом* ® *інформаційний ресурс системи*;

2) для процесу формування відповіді на запит користувача: *запит користувача* ® *управління комерційним контентом* ® *інформаційний ресурс системи* ® *супровід комерційного контенту* ® *база даних користувачів системи*;

3) для процесу формування звіту роботи системи для модератора: *запит модератора* ® *супровід комерційного контенту* ® *база даних користувачів системи* ® *управління комерційним контентом* ® *звіт для модератора*;

4) для процесу модерації внутрішніх параметрів системи: *запит модератора* ® *формування комерційного контенту* ® *база правил* ® *супровід комерційного контенту* ® *база правил* ® *управління комерційним контентом* ® *результат*.

Модель системи електронної контент-комерції S описується як

$$S = \langle X, Q, Formation, H, C, V, Management, Support, Z, T, Y \rangle, \quad (1)$$

де величина $X = \{x_1, x_2 \mathbf{K}, x_{n_X}\}$ – множина вхідних даних $x_i \in X$ з різних джерел інформації при $i = \overline{1, n_X}$; величина $Q = \{q_1, q_2 \mathbf{K}, q_{n_Q}\}$ – множина запитів $q_d \in Q$ користувачів при $d = \overline{1, n_Q}$; величина $Formation$ – оператор формування комерційного контенту; величина $H = \{h_1, h_2 \mathbf{K}, h_{n_H}\}$ – множина внутрішніх параметрів $h_k \in H$ системи S при $k = \overline{1, n_H}$; величина $C = \{c_1, c_2 \mathbf{K}, c_{n_C}\}$ – множина комерційного контенту $c_r \in C$ при $r = \overline{1, n_C}$; величина $V = \{v_1, v_2 \mathbf{K}, v_{n_V}\}$ – множина параметрів впливу $v_l \in V$ зовнішнього середовища на систему S при $l = \overline{1, n_V}$; величина $Management$ – оператор управління комерційними контентом; величина $Support$ – оператор супроводу комерційного контенту; величина $Z = \{z_1, z_2 \mathbf{K}, z_{n_Z}\}$ – множина сторінок $z_w \in Z$ інформаційного ресурсу в системі S при $w = \overline{1, n_Z}$; величина $T = \{t_1, t_2 \mathbf{K}, t_{n_T}\}$ – час $t_p \in T$ транзакції опрацювання інформаційного ресурсу в системі S при $p = \overline{1, n_T}$; величина $Y = \{y_1, y_2 \mathbf{K}, y_{n_Y}\}$ – множина статистичних даних $y_j \in Y$ роботи системи S при $j = \overline{1, n_Y}$.

Величина $y_j = \{a_1, a_2, \dots, a_8\}$, де a_1 – кількість відвідувань за період часу Δt , a_2 – середній час перебування на інформаційному ресурсі (хв:с) за період часу Δt , a_3 – показник відмовлень (%) за період часу Δt , a_4 – досягнута мета пошуку комерційного контенту, a_5 – динаміка життєвого циклу комерційного контенту (%), a_6 – загальна кількість переглядів сторінок за період часу Δt , a_7 – кількість переглядів сторінок за одне відвідування, a_8 – нові відвідування (%), a_9 – абсолютно унікальні відвідувачі, a_{10} – джерело трафіка (прямі переходи, переходи з пошукових систем, переходи з інших сайтів тощо) у % тощо [5].

Ріст обсягу контенту і швидкості його поширення породив поняття контентних потоків, аналіз яких вимагає використання нового інструментарію опрацювання інформаційних ресурсів у системах електронної контент-комерції для формування, управління та супроводу контенту (рис. 2).

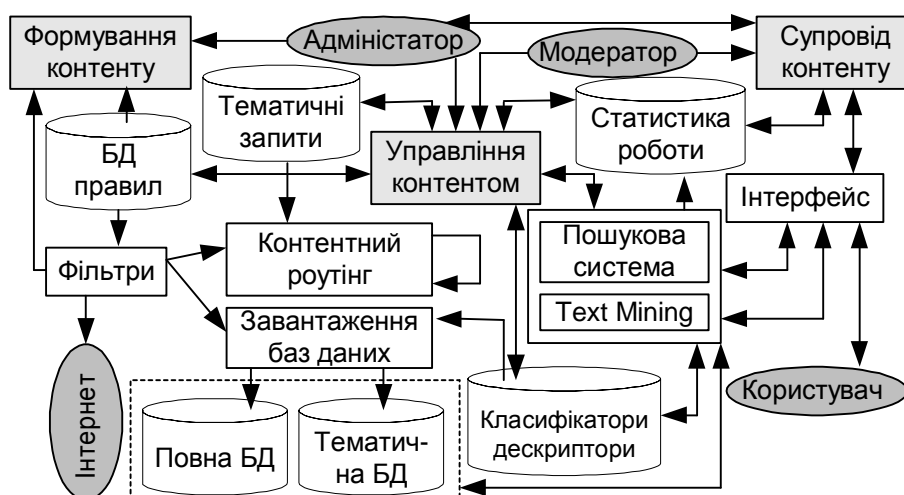
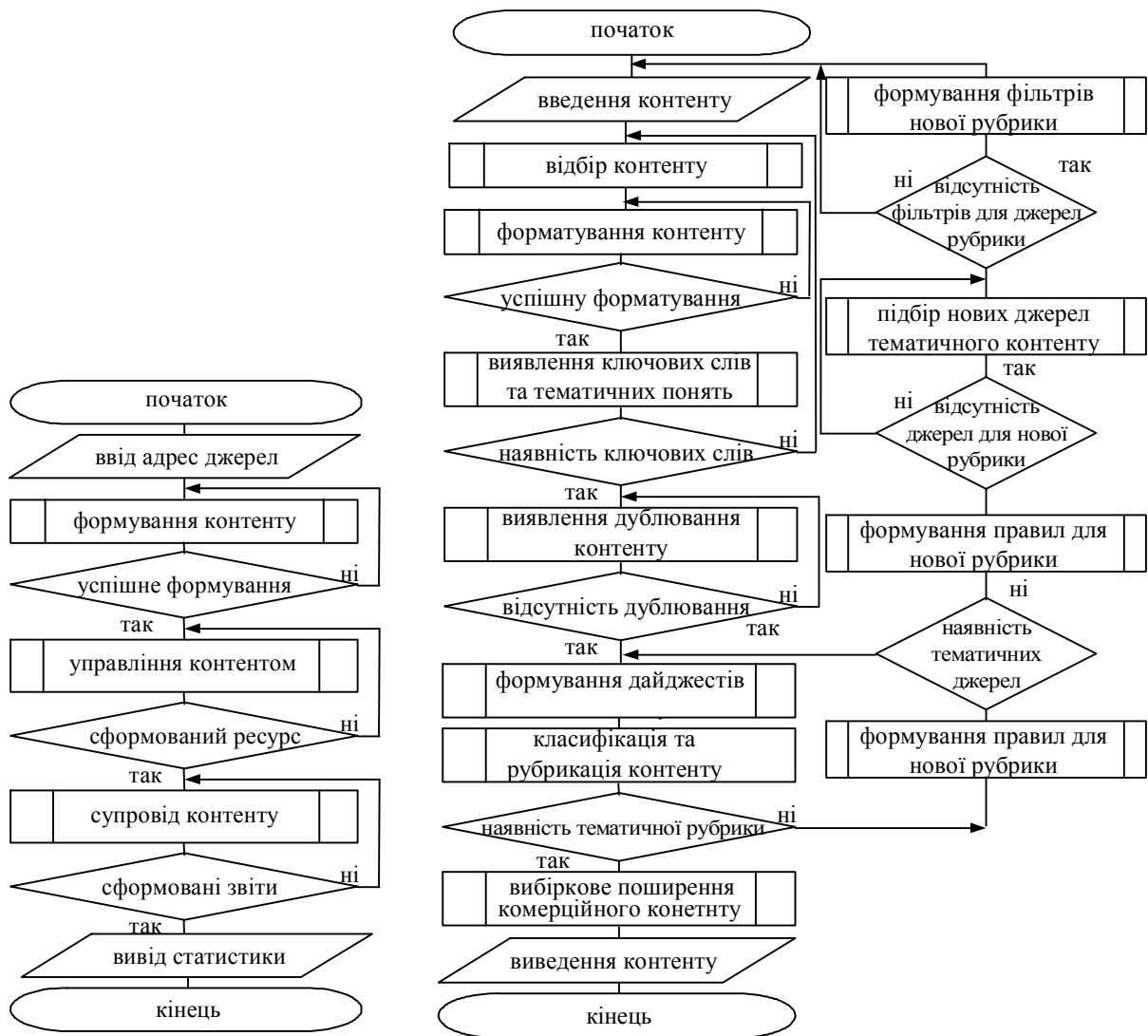


Рис. 2. Функціональна схема процесу опрацювання інформаційних ресурсів в системах електронної контент-комерції

Класичний математичний апарат та поширені інструментальні засоби не здатні адекватно відображати аналіз контентного масиву фіксованого розміру та навігацію в ньому [3, 4, 6–8]. Архітектура СЕКК має рівні ієрархії, які забезпечують незалежність збережених даних від програм, що їх використовують, та можливість розвитку системи без руйнування існуючих застосувань. На рис. 3, а подано схему процесу опрацювання інформаційних ресурсів в СЕКК, а на рис. 3, б – етапи формування комерційного контенту в системах електронної контент-комерції. СЕКК працюють на основі такої схеми: *контент-менеджер* → *підсистема редагування* → *база даних* → *ядро* → *користувач*.



а б
Рис. 3. Процес опрацювання (а) та формування інформаційних ресурсів (б)

Система електронної контент-комерції містить ядро адміністрування, підсистему авторизації/аутентифікації, менеджер шаблонів і менеджер контенту для розв'язування задач з позиції користувача (рис. 4).

При створенні ядра системи електронної контент-комерції використовують об'єктно-орієнтовану модель та абстрактні об'єкти зі властивостями/методами (рис. 5). Взаємодія з основними об'єктами ядра системи реалізують через інкапсуляцію. Для цього в класах реалізують інтерфейсні методи, призначені для маніпуляцій всередині об'єкта з даними/властивостями.

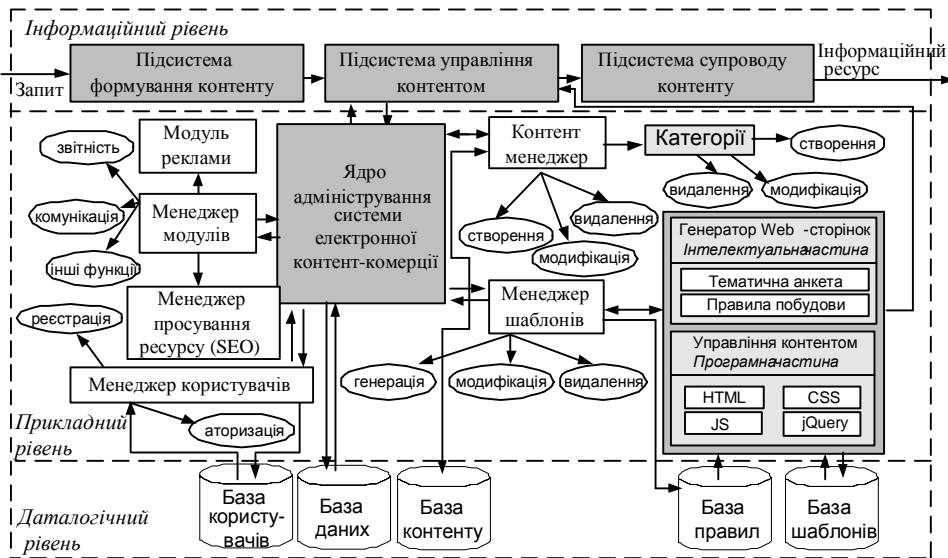


Рис. 4. Структура системи електронної контент-комерції

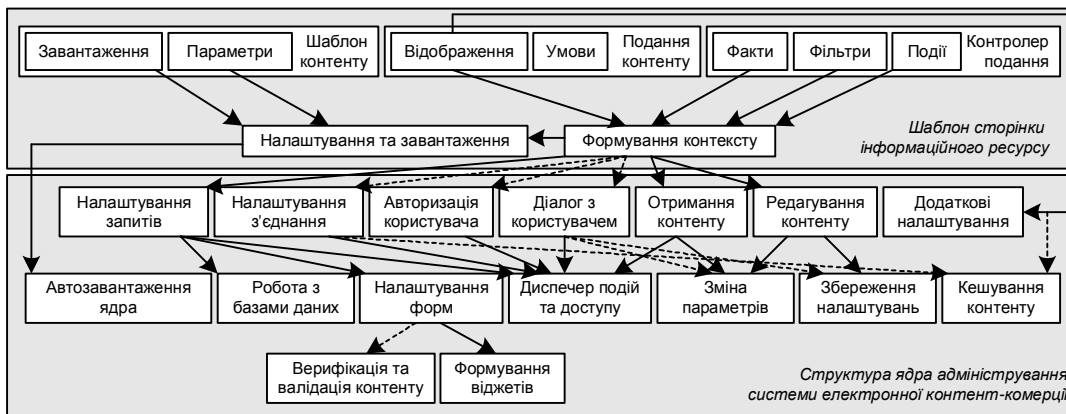


Рис. 5. Діаграма структури системи електронної контент-комерції

Процесами опрацювання інформаційних ресурсів у системах електронної контент-комерції є формування, управління та супровід контенту (рис. 6).

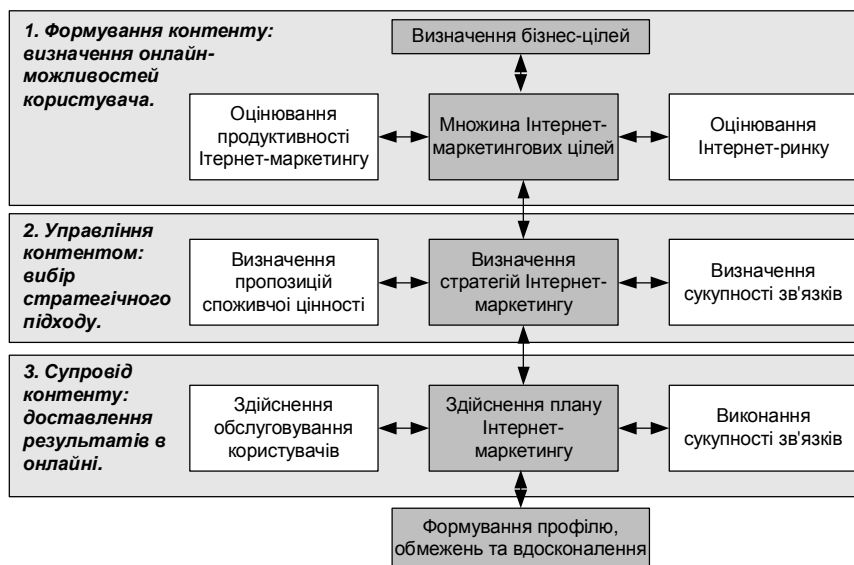


Рис. 6. Основні процеси опрацювання інформаційних ресурсів у системах електронної контент-комерції

Отримані дані враховують при створенні або оновленні інформаційного ресурсу та удосконаленні архітектури СЕКК. Раніше модератор самостійно шукав та опрацював необхідний контент: збір контенту з різних джерел даних, аналіз та фільтрація контенту, формування комерційного контенту як кінцевого продукту згідно із індивідуалізованими даними користувача системи. Реалізацією процесу опрацювання інформаційних ресурсів в СЕКК полегшують роботу модератора, автоматизуючи збирання контенту з джерел, аналіз та фільтрацію контенту.

Процес формування контенту складається із декількох етапів: *модератор* ® *створення контенту* ® *база даних* ® *систематизація контенту* ® *база даних* ® *поширення контенту* ® *редактор, або Інформаційний ресурс (джерело)* ® *збирання контенту* ® *база даних* ® *систематизація контенту* ® *база даних* ® *поширення контенту* ® *модератор* і реалізується у вигляді контент-моніторингу контенту та створення бази даних відповідно до інформаційних потреб споживачів (рис. 7). Після збирання і первинного опрацювання контент приводять до єдиного формату, класифікують відповідно до визначеного рубрикатора та йому приписують дескриптори, включаючи ключові слова. При застосуванні інтернет-маркетингу етапи систематизації контенту забезпечують постійне поповнення бази даних оперативними даними, ефективний одночасний доступ багатьох користувачів до бази даних, зручні засоби пошуку необхідного контенту.

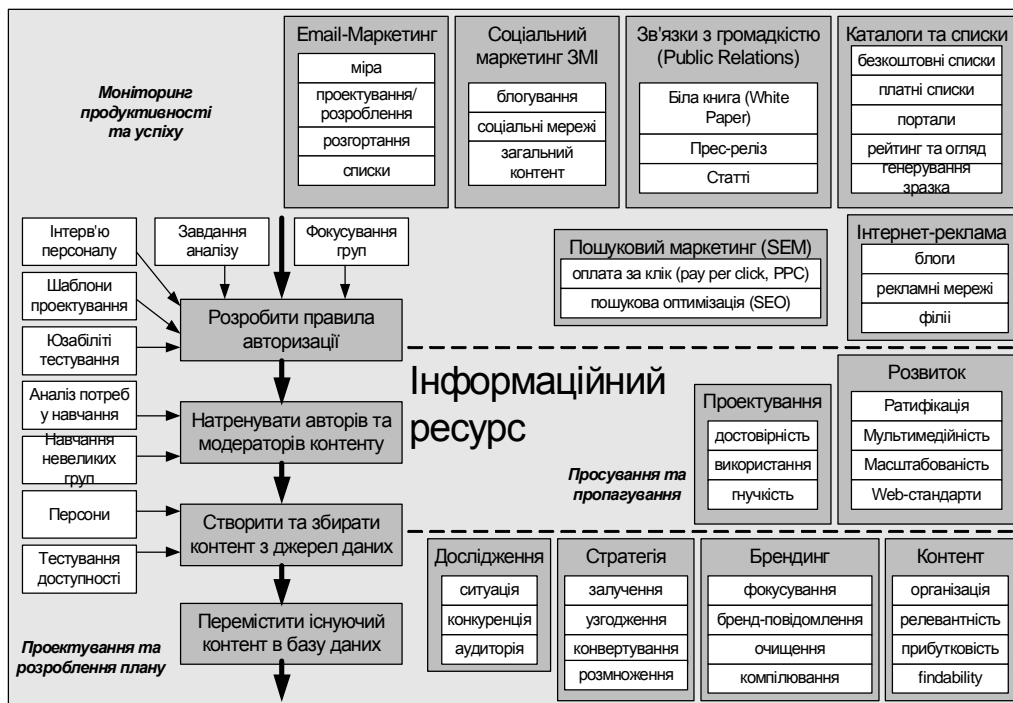


Рис. 7. Методи збирання/створення контенту та маркетингу

Поділ на етапи формування комерційного контенту в СЕКК підвищує ефективність при адмініструванні системи; забезпечує економію ресурсів, інтернет-трафіку та анонімність користувачів; автоматичне сканування джерел даних.

Управління контентом інформаційного ресурсу та відвідувачів (рис. 8–9), їх моделювання є процесом кількісного дослідження динаміки окремих тематичних напрямів та технічний аналіз інформаційного ресурсу.

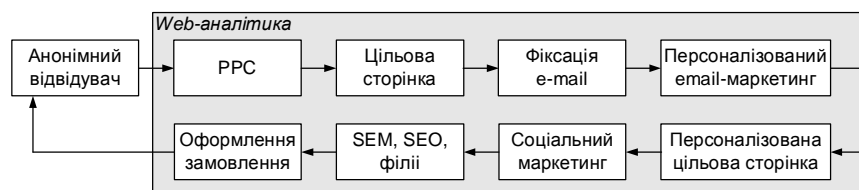


Рис. 8. Схема процесу аналізу відвідування інформаційного ресурсу

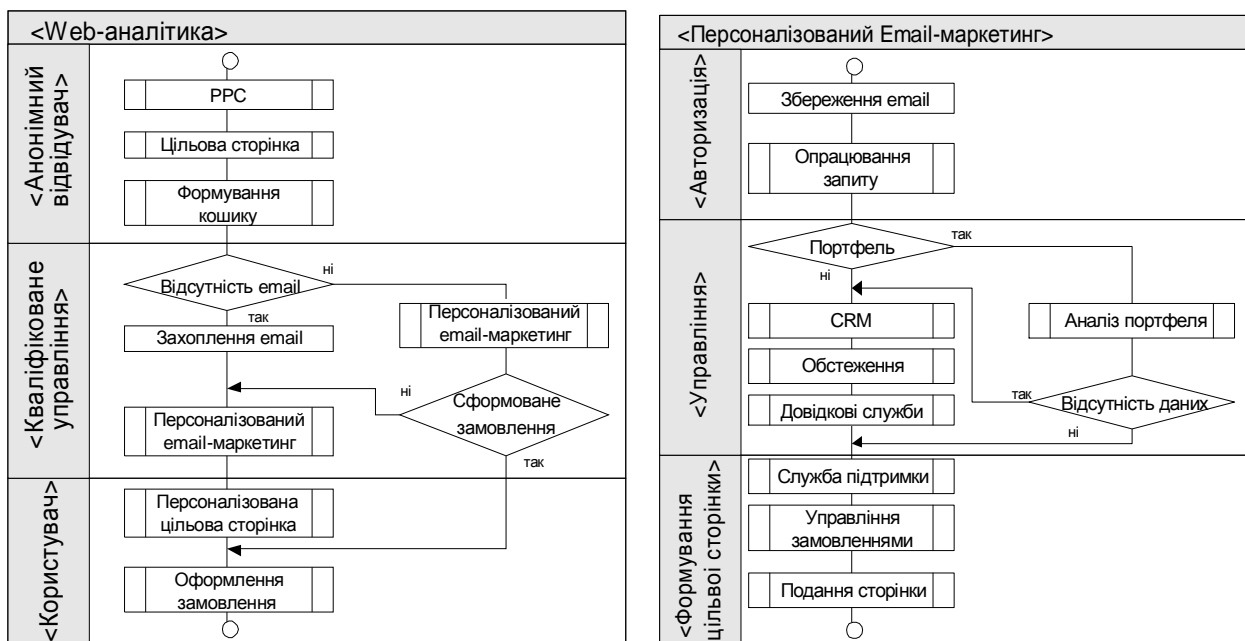


Рис. 9. Схема аналізу контенту відвідувачів (а) та інформаційного ресурсу (б)

Процес управління контентом складається із декількох етапів (рис. 9): користувач ® опрацювання контенту ® база даних ® аналіз контенту ® база даних ® подання контенту ® користувач.

Аналіз результатів управління контентом впливає на швидкість розвитку тематичних напрямів та контентного простору. Стійкі статичні зв'язки між контентом свідчать про кореляцію тематик, ефективність посилань на публікації джерел, більш ранні цитування, републікації тощо. Механізми, основні на узагальнених методах кластерного аналізу, виявляють контент у потоках, що формує навколо себе нові тематичні напрями. Кластерний аналіз, теорія фракталів і автотельних процесів при їхньому коректному застосуванні кількісно оцінюють ступінь зв'язку в тематичних контентних потоках. Оперативний аналіз контенту відвідувачів сприяє реалізації процесу управління контентом за допомогою генерації сторінок через інформаційні блоки, який поділяється на типи: тематична; по останніх зверненнях; комбінована.

Процес супроводу контенту – це оперативні етапи узагальнення, модерації та структурування комерційного контенту (рис. 10), тобто Користувач ® структурування контенту ® база даних ® модерація контенту ® база даних ® узагальнення контенту ® модератор. Із підсистемою супроводу контенту СЕКК має такі можливості: формування рейтингу комерційного контенту; формування інформаційного портрету постійного користувача СЕКК; аналіз характеристик (коментарі, відгуки, побажання тощо) на комерційний контент з боку користувача СЕКК; збирання, накопичення та опрацювання інформації про потреби кінцевого/потенційного користувача СЕКК та споживача контенту; формування інформаційного портрету контентних потоків.

Система електронної контент-комерції значно полегшує роботу модератора з формування, управління та супроводу комерційного контенту (рис. 11). Вона передбачає такі основні етапи опрацювання контенту інформаційних ресурсів:

1. Збирання контенту з різних джерел даних системою.
2. Фільтрація створеного і/або зібраного контенту системою.
3. Аналіз (визначення ключових слів, рубрикація, формування дайджестів) контенту системою.
4. Аналіз отриманих даних для концентрування уваги модератора на необхідній тематиці.

Підготовка даних для оперативної роботи з ними.

5. Аналіз потреб потенційних та постійних користувачів системи.

6. Формування кінцевого продукту згідно із проаналізованими даними та потребами потенційних/постійних користувачів.
7. Поповнення списку джерел даних згідно із результатами аналізу запитів користувачів.
8. Поповнення правил та словників фільтрів згідно із результатами аналізу контенту.
9. Виявлення нових тематичних рубрик потоку контенту.

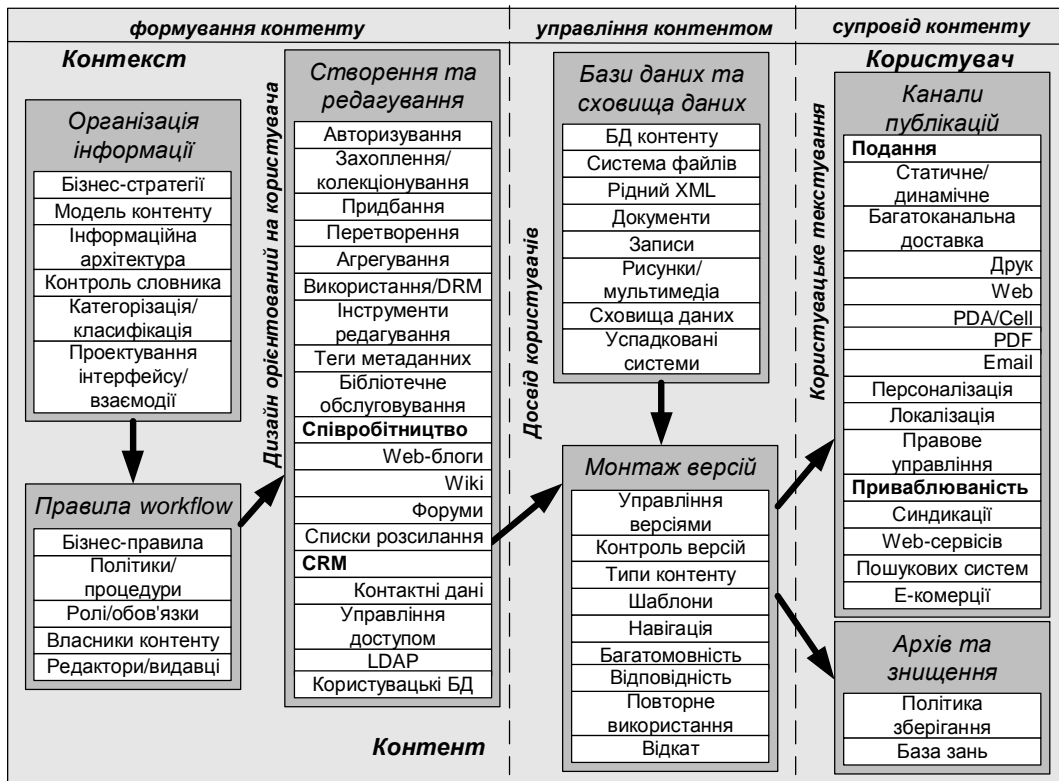


Рис. 10. Підсистеми формування, управління та супроводу контенту в системах електронної контент-комерції

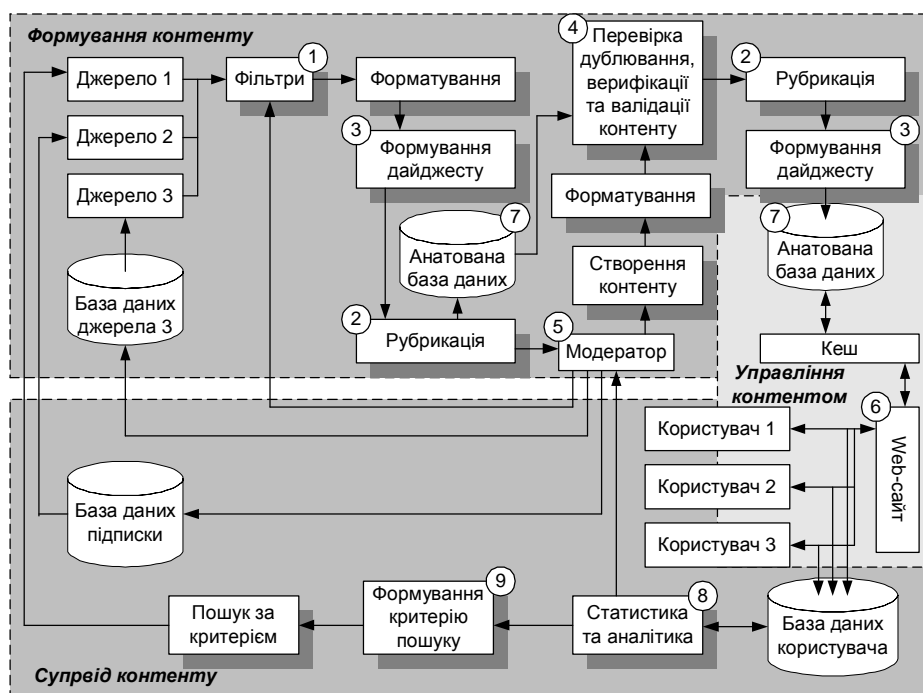


Рис. 11. Схема взаємодії підсистем опрацювання інформаційних ресурсів

На рис. 12 подано класифікацію розроблених методів опрацювання інформаційних ресурсів у СЕКК з детальним переліком реалізованих етапів формування, управління та супроводу комерційного контенту в цих системах.



Рис. 12. Методи опрацювання інформаційних ресурсів у системах електронної контент-комерції

Розроблені методи опрацювання інформаційних ресурсів в системах електронної контент-комерції дають можливість сформулювати вимоги до підпрограм опрацювання інформаційних ресурсів.

Контент-аналіз інформаційних ресурсів

Особливістю сучасності є постійний ріст темпів виробництва контенту. Цей процес є об'єктивним і позитивним, але виникла проблема: прогрес у галузі виробництва контенту призводить до пониження загального рівня інформованості потенційного користувача. Крім збільшення обсягів контенту до масштабів, яке призводить до неможливості його безпосереднього опрацювання та швидкості його поширення, виникає низка специфічних проблем (таблиця).

Основні негативні чинники формування комерційного контенту

Назва	Основна причина	Рішення
Інформаційний шум	Структурованість масивів контенту.	Фільтри, контент-моніторинг, аналіз сайту, контент-аналіз.
Паразитичний контент	Поява як додатків.	Фільтри, контент-моніторинг, контент-аналіз.
Нерелевантність контенту	Невідповідність потребам користувачів.	Створення анотованої бази даних, пошукових образів первинного контенту та їх кластеризація, контент-аналіз.
Дублювання контенту	Дублювання в джерелах.	Контент-аналіз, сканери і фільтри на базі статистики та критеріїв.
Навігація в потоці контенту	Швидкий ріст обсягу і поширення контенту.	Аналіз сайту, фільтри, контент-моніторинг, контент-аналіз.
Надмірність пошуку	Дублювання і нерелевантність.	Анотований пошук, контент-аналіз та реферування.

Негативні чинники у формуванні комерційного контенту ускладнюють процес пошуку необхідних даних при скануванні різних джерел інформації. Збільшення фізичного обсягу та змінна актуальності/динаміки контентних потоків (постійне систематичне та нерегулярне оновлення) призводить до дублювання, інформаційного шуму та надмірності результатів пошуку контенту. Охоплення та узагальнення великих динамічних потоків контенту, які неперервно генерують в Інтернет-джерелах, вимагає якісно нових методів/підходів пошуку – таких, як контент-моніторинг (рис. 13). Вхідною інформацією для контент-моніторингу є текст природною мовою як послідовність символів, вихідна інформація – це таблиці розділів, речень і лексем аналізованого тексту. Контент-моніторинг є програмним засобом автоматизації знаходження найважливіших складових у потоках контенту. Це змістовний аналіз потоків контенту з метою постійного отримання необхідних якісних/кількісних зрізів протягом наперед не визначеного проміжку часу.

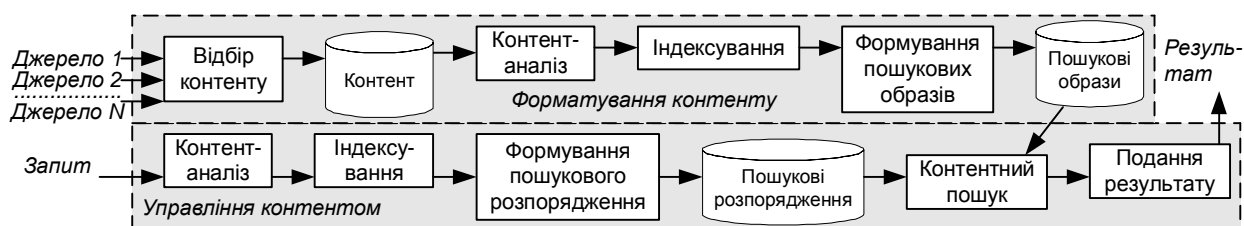


Рис. 13. Структурна схема процесу контент-моніторингу під час формування комерційного контенту

Складовою контент-моніторингу є контентний пошук та контент-аналіз тексту. Контент-аналіз призначений для пошуку контенту в масиві даних за змістовими лінгвістичними одиницями (алг. 1). Одиниця рахунку є кількісною мірою одиниці аналізу, що дає змогу реєструвати частоту (регулярність) появи ознаки категорії аналізу в тексті (кількість певних слів або їх поєднань, рядків, друкованих знаків, сторінок, абзаців, авторських аркушів, площа тексту тощо).

Алгоритм 1. Контент-аналіз текстового комерційного контенту.

Етап. 1. Визначення набору критеріїв для текстового комерційного контенту.

Крок 1. Формування набору критеріїв як тип джерела (форум, електронна пошта, Інтернет-газета, чат, Інтернет-журнал); тип контенту (стаття, електронний лист, баннер, коментарій); учасники комунікації (відправник, одержувач, реципієнт).

Крок 2. Визначення розміру (мінімальний обсяг або довжина), частоти появи, способу/місця розповсюдження та час появи контенту.

Крок 3. Фільтрування згідно сформованим набором критеріїв контентного потоку та зберігання ідентифікованого релевантного контенту.

Етап. 2. Контент-аналітичний відбір. Формування вибіркової сукупності контенту за критеріями обмеженої вибірки з більшого масиву.

Етап. 3. Виявлення змістовних одиниць аналізу текстового комерційного контенту (словосполучення, речення, тема, ідея, автор, персонаж, соціальна ситуація, частина тексту, кластеризованна за змістом категорії аналізу). Вимоги до вибору лінгвістичної одиниці аналізу: достатньо велика для інтерпретації значення; достатньо мала, щоб не інтерпретувати багато значень; легко ідентифікується; кількість одиниць достатньо велика для проведення вибірки.

Етап. 4. Виділення одиниць рахунку аналізу текстового контенту.

Крок 1. Якщо одиниці рахунку збігаються з одиницями аналізу, то знаходять частоти появи виділеної змістовної одиниці, інакше перейти до кроку 2.

Крок 2. Модератор на основі аналізованого контенту висуває одиниці рахунку, наприклад, протяжність текстів; площа тексту, заповнена змістовними одиницями; кількість рядків (абзаців, знаків, колонок тексту); розмір/вид файлу; кількість рисунків з певним змістом/сюжетом тощо.

Етап. 5. Порівняння змістовних одиниць аналізу з одиницями рахунку.

Крок 1. Класифікація за угрупованнями із оціненням ваги змістовних категорій в загальному обсязі тексту. Класифікатором є загальна таблиця, в яку зведені всі категорії аналізу і одиниці аналізу. Фіксують одиниці виразу категорій.

Крок 2. Статистичні розрахунки зрозумілості та атрактивності контенту.

Етап. 6. Розроблення інструменту контент-аналізу.

Крок 1. Створення закодованого протоколу контенту для компактності подання даних та швидкого порівняння результатів аналізу різного контенту.

Крок 2. Заповнення протоколу контенту властивостями (автор, час видання, обсяг тощо).

Крок 3. Заповнення протоколу комерційного контенту підсумками його аналізу (кількість вживання в ньому певних одиниць аналізу і висновки щодо категорій аналізу). Протокол кожного контенту заповнюється на основі підрахунку даних всіх його реєстраційних карток.

Етап. 7. Розроблення таблиці контент-аналізу. Тип таблиці визначають у вигляді системи скоординованих і субординованих категорій аналізу: кожна категорія (питання) передбачає ряд ознак (відповідей), за якими квантифікується зміст тексту.

Етап. 8. Розроблення кодувальної матриці контент-аналізу.

Крок 1. Якщо обсяг вибірки ≥ 100 одиниць, то аналізується набір матричних листів, інакше виконати крок 2.

Крок 2. Якщо вибірка < 100 одиниць, то проводиться двовимірний аналіз. В цьому випадку для кожного контенту формується кодувальна матриця.

Етап. 9. Проведення аналізу тексту згідно із створеними кодувальними матрицями.

Етап. 10. Інтерпретація результатів. Виявляють і оцінюють характеристики контенту на основі статистичного набору підрахованих коефіцієнтів за певний період часу на визначену категорію. Охоплює всі здобуті фрагменти тексту, висновки спираються не на частину результатів, а враховуються всі без винятку.

Застосування контент-аналізу при моніторингу інтернет-джерел даних дає змогу автоматизувати процес знаходження найважливіших складових в потоках контенту при відборі даних з цих джерел. Це усуває дублювання контенту, інформаційний шум, паразитичний контент, надмірність результатів пошуку тощо. Цей метод застосовують в подальших етапах формування комерційного контенту для отримання більш точного релевантного результату – створення унікального комерційного контенту, який користується попитом серед користувачів СЕKK.

Висновки і перспективи подальших наукових розвідок

З позиції системного підходу застосовано принципи опрацювання інформаційних ресурсів у системах електронної контент-комерції для реалізації життєвого циклу контенту, що дало змогу розробити методи формування, управління та супроводу комерційного контенту. Розроблено комплексний метод формування комерційного контенту для скорочення часу та зменшення ресурсів виробництва контенту, що дає можливість створити засоби опрацювання інформаційних ресурсів та реалізувати підсистему автоматичного формування контенту. Створено оперативний метод управління комерційним контентом для скорочення часу та зменшення ресурсів продажу контенту, що дає можливість реалізувати підсистему управління комерційним контентом. Реалізовано комплексний метод супроводу контенту для скорочення часу та зменшення ресурсів аналізу цільової аудиторії системи електронної контент-комерції, що дає можливість розробити підсистему супроводу контенту.

1. Береза А. Електронна комерція / А. Береза, І. Козак, Ф. Левченко. – К: КНЕУ, 2002. – 326 с.
2. Берко А. Системи електронної контент-комерції / А. Берко, В. Висоцька, В. Пасічник. – Л.: НУЛП, 2009. – 612 с.
3. Большакова Е. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика / Е. Большакова, Д. Ландэ, А. Носков, Э. Клышинский, О. Пескова, Е. Ягунова. – М: МИЭМ, 2011. – 272 с.
4. Клифтон Б. Google Analytics: профессиональный анализ посещаемости веб-сайтов / Б. Клифтон. – М: Вильямс, 2009. – 400 с.
5. Корнеев В. Базы данных. Интеллектуальная обработка информации / В. Корнеев, А. Гареев, С. Васютин, В. Райх. – М: Нолидж, 2000. – 352 с.

⁴Ю.Морозов, ¹Р. Назаревич, ²А. Назаревич, ³Є.Струк, ⁴В. Мархивка
¹Фірма “SoftServe”,

²Карпатське відділення Інституту геофізики ім. С.І. Субботіна НАН України,
 Національний університет “Львівська політехніка”,
³кафедра автоматизованих систем управління,
⁴кафедра електронних обчислювальних машин

ІНТЕРНЕТ-ПОРТАЛ ГЕОФІЗИЧНОГО МОНІТОРИНГУ

© Морозов Ю., Назаревич Р., Назаревич А., Струк Є., Мархивка В., 2014

Представлено прототип розробленого Інтернет-порталу геофізичного моніторингу як підсистеми програмного комплексу збирання та обробки даних геодинамічного моніторингу. Прототип реалізовано на основі системи керування базами даних *MySQL*, який може розміщуватися як на власному сервері організації-замовника розробки, так і на публічних інтернет-серверах. Описано взаємодію бази даних та клієнтських модулів у рамках даного програмного комплексу.

Ключові слова: геодинамічний моніторинг, система збирання даних, база даних, обробка даних.

The prototype of developed geophysical Internet portal as subsystem of software of geodynamic monitoring data acquisition and processing is presented. The prototype is based on database management system *MySQL* and it can be placed both on their own servers of an awarding authority and on public Internet servers. Interaction of database and client modules within the bounds of this software is described.

Key words: geodynamic monitoring, data acquisition system, database, data processing.

Вступ

Сучасний моніторинг природних та техногенних геодинамічних процесів є дуже актуальною проблемою людства. Достатньо згадати катастрофічні наслідки кількох останніх найсильніших світових землетрусів – Суматранських (2004–2005 рр.), Китайського (2008 р.), Гаїтянського та Чілійського (2010 р.), Японського (2011 р.). Зараз такий моніторинг неможливо уявити без використання комп’ютерних та інтернет-технологій. Вони використовуються в інтелектуальних геофізичних вимірювальних системах, в локальних системах збирання і передавання даних, у регіональних та світових центрах збирання та обробки моніторингової інформації [1–5].

Показано особливості застосування комп’ютерних та інтернет-технологій у створенні інтернет-порталу опрацювання даних геофізичного моніторингу.

Вихідні дані

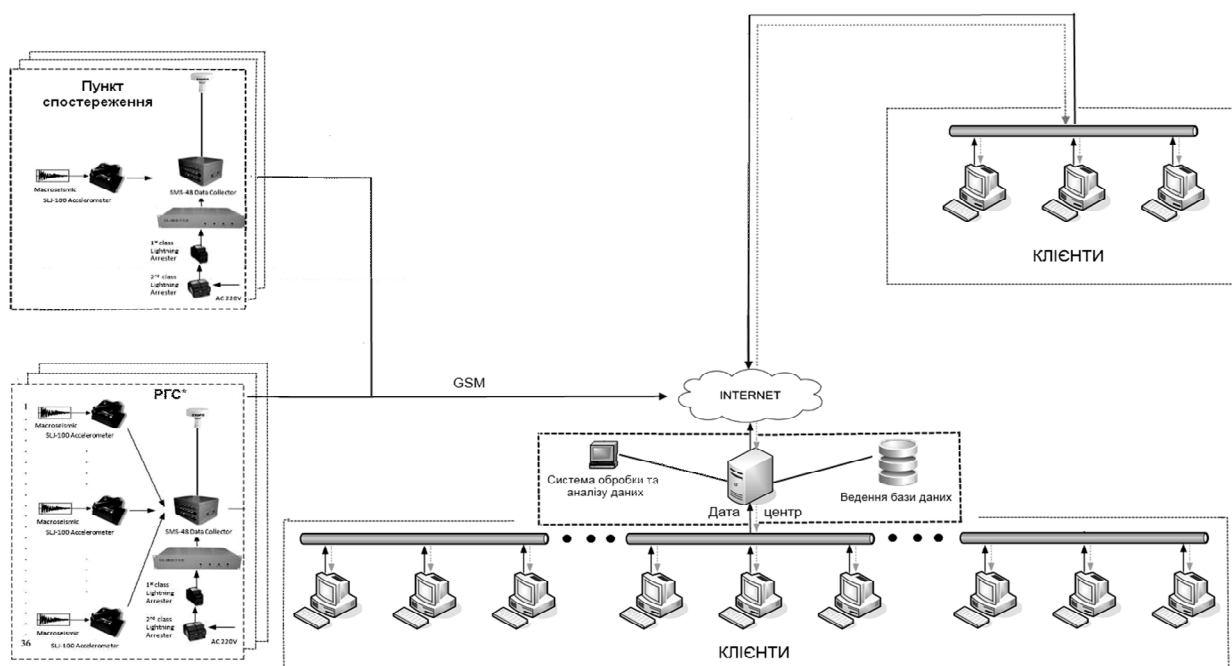
Сейсмопрогностичні дослідження у світі ведуться вже понад 50 років [6, 7]. Співробітники Карпатського відділення Інституту геофізики ім. С.І. Субботіна НАН України вже понад 30 років здійснюють моніторингові сейсмопрогностичні геофізичні дослідження у Закарпатті [8–18]. Ці дослідження проводять з використанням різних геофізичних методів на мережі режимних геофізичних станцій (РГС) та пунктів спостережень. Сьогодні на РГС та пунктах спостережень працює цифрова геофізична апаратура. Дані реєструються на місці та періодично доставляються у Львів для обробки й аналізу.

У зв'язку з розширенням в останній час доступності мережі Інтернет постало завдання використати її для оперативного передавання даних з пунктів спостережень у створюваний регіональний геофізичний інформаційний центр. При цьому доцільно комплексно розв'язати задачі передавання даних з пунктів спостережень, їх архівування в базу даних, обробки та аналізу. Саме таке завдання поставлено при створенні інтернет-порталу опрацювання даних геофізичного моніторингу.

Загальна структура та алгоритм роботи інтернет-порталу

Загальна структурна схема інтернет-порталу опрацювання геофізичних даних (рис. 1) містить мережу РГС (які самі є локальними системами збирання та обробки даних) і пунктів геофізичних спостережень, пов'язаних з регіональним геоінформаційним центром (дата-центром) через мережу Інтернет. У дата-центрі розміщено сервер, базу даних, модуль адміністратора і локальну мережу користувачів (клієнтів). Передбачено також віддалений клієнтський доступ до сервера і бази даних через Інтернет.

Для створення програмного забезпечення інтернет-порталу опрацювання даних геодинамічного моніторингу використано мову програмування Java. Програмно інтернет-портал створено у вигляді інтерактивної веб-сторінки, яка може розміщуватися як на власному сервері організації-замовника розробки (Карпатського відділення інституту геофізики ім. С.І. Субботіна НАН України (як на рис. 1)), так і на публічних інтернет-серверах. Функціонально програмний комплекс складається з таких модулів: прийому геофізичної інформації; зв'язку з базою даних; клієнтських модулів; модуля адміністратора системи.



*РГС – режимна геофізична станція

Рис. 1. Структурна схема інтернет-порталу опрацювання геофізичних даних

Модуль прийому реалізує отримання геофізичної інформації інтернет-каналами з пунктів спостережень. Модуль зв'язку з базою даних забезпечує ведення бази геофізичних даних з прив'язкою до часу, пункту спостережень і геофізичного методу (геофізичної апаратної системи – джерела даних). Клієнтський модуль забезпечує пошук та перегляд накопичених даних у графічному вигляді та завантаження їх на комп'ютер користувача (диференційовано залежно від рівня доступу для конкретного користувача). Модуль адміністратора системи забезпечує адміністрування та доступ до режимів модифікації програмного забезпечення.

Аналіз інформаційних потоків та вимог до порталу

Завданням програмного забезпечення інтернет-порталу опрацювання геофізичних даних на стадії створення прототипу є введення даних з вхідного буфера (куди вони записуються модулем прийому даних) в базу даних і видача їх за запитом клієнтів (з врахуванням їхнього рівня доступу) для візуального перегляду, корегування та обробки. Для прикладу на рис. 2 наведено формат вхідного файлу з геотермічними даними.

№	Дата	Температура
1	2011.02.23	0000
2	00:00:00	50028257
3	00:00:10	50028328
4	00:00:20	50028385
5	00:00:30	50028435
6	00:00:40	50028485
7	00:00:50	50028536
8	00:01:00	50028581
9	00:01:10	50028625

Рис. 2. Формат файлу з вхідними геотермічними даними

Завдання створення і ведення бази даних геофізичного моніторингу (показане тут на прикладі геотермічних даних) є одним з ключових при розробці системи. На основі структури бази даних надалі будується клієнтська та серверна частини інтернет-порталу.

База даних повинна забезпечувати зберігання даних з прив'язкою до:

- пункту спостережень;
- типу апаратури;
- номера каналу;
- часу і дати.

У випадку найбільшого потоку інформації (за винятком сейсмічних каналів і спеціальних режимів для геоакустичної, інфразвукової, нахиломірної та деякої іншої апаратури), коли дані з геофізичної апаратури надходять щосекунди, маємо такі обсяги даних (враховуючи, що на 1 відлік передбачено 10 знаків (5 байтів)):

Об'єми геофізичних даних за різних частот вимірювань

Частота вимірювань	1 секунда		10 секунд		1 хвилина	
	Обсяг даних		Обсяг даних		Обсяг даних	
Часовий інтервал	значень	байтів	значень	байтів	значень	байтів
година	3600	18000	360	1800	60	300
доба	86400	432000	8640	43200	1440	7200
місяць	2592000	12,1 млн.	259200	1,21 млн.	43200	216000
рік	31 млн	155 млн.	3,1 млн	15,5 млн.	518400	2,59 млн.

Як систему керування базами даних передбачено використати відомий програмний продукт *MySQL*. Оскільки у вибраній базі даних максимальна кількість рядків однієї інформаційної таблиці становить 50 мільйонів, то при максимальному потоці даних (див. таблицю) отримані результати доцільно об'єднувати у річні таблиці. Для оптимізації подальшої роботи з даними передбачено можливість помісячного партиціювання таблиць.

Зважаючи на максимальний можливий обсяг даних, які можна зберігати в базі даних, а для *MySQL 3.23+* це становить 8 мільйонів терабайт (2^{63}), зрозуміло, що цього з запасом достатньо для зберігання даних.

Для геотермічних даних, на прикладі опрацювання яких ми реалізуємо інтернет-портал, інтервал вимірювання становить 10 секунд а максимальний річний обсяг даних – 3,1 млн. значень.

Передбачено зберігання даних у 2-х форматах – прямому (10 знаків) і з плаваючою комою (6 знаків – мантиса, 4 знаки – порядок).

Також база даних повинна підтримувати стандарт SQL запитів 2003 року для того, щоб можна було здійснювати вибірку даних за необхідними параметрами.

Шляхи оптимізації структури бази даних та форматів подання даних у ній на цьому етапі не розглядаються – ці питання стануть розроблятися надалі, враховуючи реальні характеристики даних з наявної на пунктах спостережень різної геофізичної апаратури.

Вимоги до апаратного та системного програмного забезпечення

Вимоги до апаратного та програмного забезпечення, на якому повинен розміщуватись інтернет-портал, не жорсткі, оскільки вибране програмне забезпечення портоване на велику кількість платформ (AIX, BSDi, FreeBSD, HP-UX, Linux, Mac OS X, NetBSD, OpenBSD, OS/2 Warp, SGI IRIX, Solaris, SunOS, SCO OpenServer, SCO UnixWare, Tru64, Windows 95, Windows 98, Windows NT, Windows 2000, Windows XP, Windows Server 2003, WinCE, Windows Vista і Windows 7).

Мінімальні вимоги до апаратної частини сервера інтернет-порталу такі:

- Процесор: Pentium 3 і вище;
- Оперативна пам'ять: 512 Мбайт і більше;
- Об'єм вінчестера: 80 Гбайт і більше;
- Периферія: LAN 10 Мбіт, CD/DVD-ROM

Для самої бази даних доцільно використати ще 2 накопичувачі, об'єднані в масив типу RAID 1, що забезпечить автоматичне дублювання даних, а отже, і підвищену надійність їх зберігання. Один з цих накопичувачів може бути також віддаленим мережним сховищем інформації. Об'єм цих накопичувачів доцільно визначити з розрахунку накопичення геофізичних даних за період не менше 10 років і з врахуванням планів розвитку моніторингової мережі – збільшення кількості пунктів спостережень і комплексу апаратури на цих пунктах.

Проектування інтернет-порталу

На основі наведених вище результатів аналізу було розроблено загальну схему роботи та взаємодії клієнтського та серверного модулів (рис. 3).

Засоби програмування

Як уже зазначалося вище, програмно систему створено у вигляді інтерактивної веб-сторінки, і для цього використано мову програмування Java. Її вибрано тому, що вона надає необхідної функціональності, а саме:

- Ця мова є об'єктно-орієнтованою, якою зручно писати великі та складні проекти.

• Java є платформонезалежною, це дає змогу розгорнути розроблений інтернет-портал на будь-якому апаратно-програмному комплексі, для якого є віртуальна машина Java (а вона зараз існує практично для усіх платформ).

Для мови Java існує велика кількість безкоштовних бібліотек, а також програмного забезпечення з відкритим кодом.

Віртуальна машина Java вміє автоматично керувати пам'яттю – це один з аспектів, який полегшує написання програм. Для розроблення програмного продукту використано такі технології:

Java EE (Enterprise Edition) – набір специфікацій і відповідної документації для мови Java, який описує архітектуру серверної платформи для задач середніх і великих проектів;

JSP (Java server page) – технологія, що дає змогу веб-розробникам динамічно генерувати HTML, XML та інші веб-сторінки;

Java Servlet – стандартизований набір функцій для створення динамічного контенту для веб-сервера, використовуючи платформу Java.

Розробляли програмний продукт, використовуючи таких середовища:

NetBeans 7.0 – середовище розроблення мовою Java (перевагами цього середовища є наявність великої кількості інструментів та функцій у стандартному варіанті установки, це дозволяє не встановлювати додаткові плагіни для роботи, також перевагою є наявність встановлюваного із середовищем веб-сервера сервлетів *GlassFish 3.0*).

Eclipse 3.6 середовище розробки мовою Java (перевагами цього середовища є велика гнучкість при налаштуванні, велика кількість додатків та плагінів, які надають необмежений простір для розширення функцій середовища). Також перевагою є переносимість середовища, його не потрібно встановлювати, достатньо раз розпакувати, налаштувати і встановити потрібні плагіни, і надалі він працюватиме на будь-якій системі даного типу.

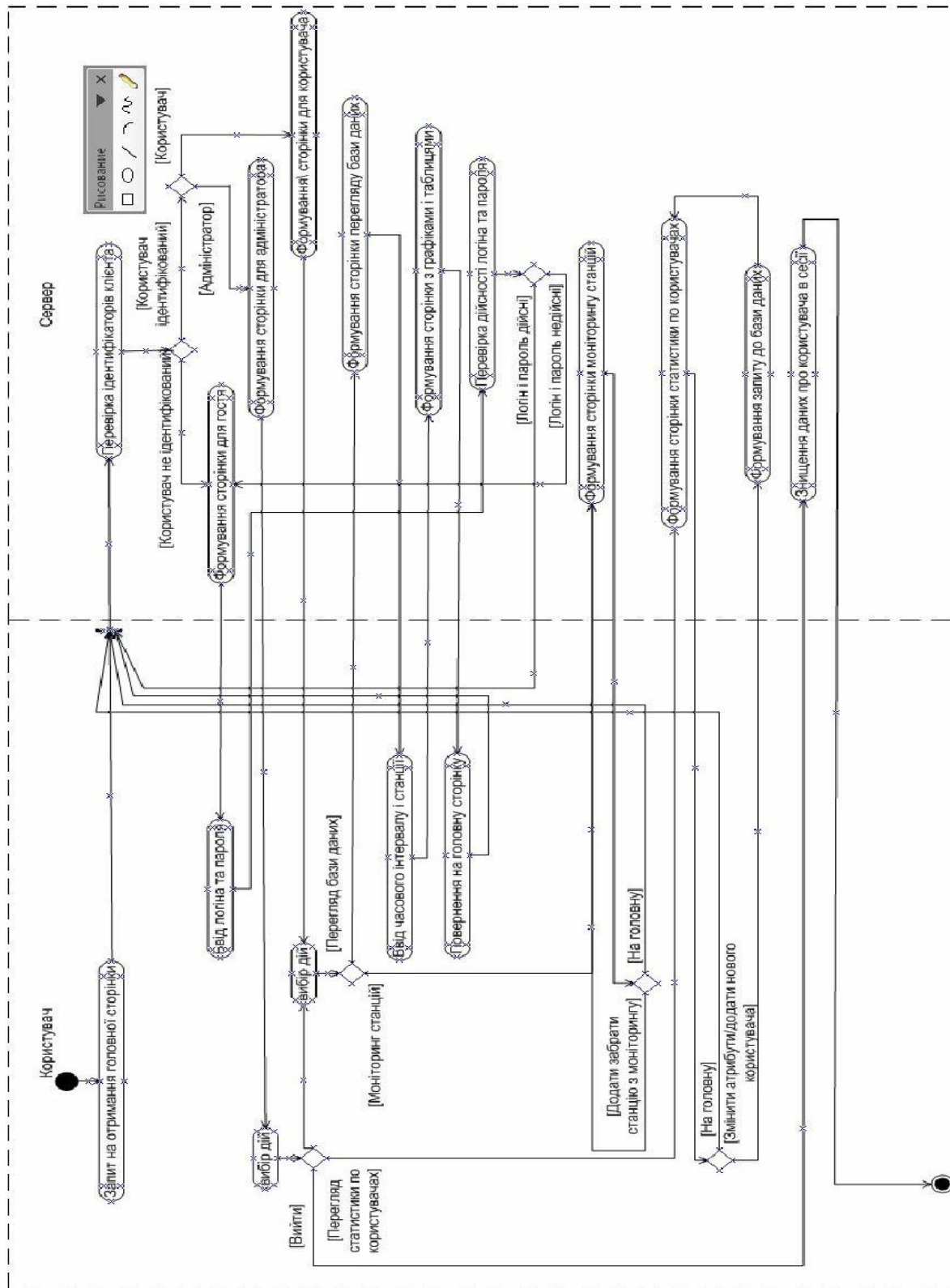


Рис. 3. Алгоритм взаємодії клієнтського та серверного модулів

Підтримувані архітектури для Eclipse 3.6:

- AIX (PPC/Motif)
- FreeBSD (x86/GTK 2)
- HP-UX (HP9000/Motif)
- Linux (x86, x86-64, PPC, IA-64/GTK 2)
- Linux (x86/Motif) Mac OS X (x86, x86-64, PPC/ Cocoa)
- OpenSolaris (x86, x64, SPARC/ GTK 2)
- Solaris 8 (SPARC/GTK 2, Motif)
- Microsoft Windows (Win32, Win64)

Як систему керування базами даних (СКБД) використано систему з відкритим кодом *MySQL* [20, 21]. Зараз *MySQL* – одна з найпоширеніших систем керування базами даних.

Вона використовується насамперед для створення динамічних веб-сторінок, оскільки має чудову підтримку з боку різноманітних мов програмування для Інтернету.

Для некомерційного використання *MySQL* є безкоштовним. Можливості сервера *MySQL* такі:

- простота встановлення та використання;
- підтримується необмежена кількість користувачів, що одночасно працюють із БД;
- кількість рядків у таблицях може сягати 50 млн.;
- висока швидкість виконання команд;
- наявність простої та ефективної системи безпеки;
- підтримка транзакцій.

Одним з основних компонентів проєктованого інтернет-порталу є контейнер сервлетів Apache Tomcat – контейнер сервлетів, який написаний мовою Java і реалізує специфікацію сервлетів та специфікацію JavaServer Pages (JSP), які є стандартами для розроблення веб-додатків мовою Java. У складі використаного контейнера сервлетів Apache Tomcat є такі компоненти: Jasper (перепроєктований механізм JSP), Catalina (перепроєктований контейнер сервлетів) і Coyote (стек HTTP).

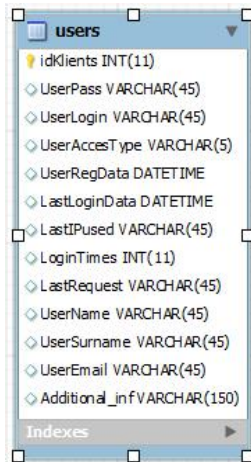
Реалізація компонентів інтернет-порталу

Модуль зв'язку з базою даних

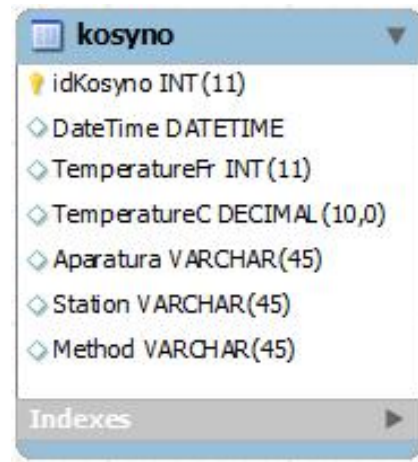
На основі аналізу наведених вище вимог було розроблено структуру бази даних, у якій дані поділяються на дві основні підмножини.

Перша підмножина – таблиця користувачів (у ній зберігаються персональні дані кожного користувача (рис. 4, а)).

- Ідентифікатор користувача (UserID);
- Дата реєстрації користувача (UserRegData);
- Логін користувача (UserLogin);
- Пароль користувача (UserPass);
- Права доступу користувача до ресурсів системи (UserAccessType);
- Дата останнього входження до системи користувача (LastLoginDate);
- IP-адреса останнього входження до системи користувача (LastIPDate);
- Кількість заходів до системи (LoginTimes);
- Останній запит, посланий користувачем (LastRequest);
- Ім'я користувача (UserName);
- Прізвище користувача (UserSurname);
- Е-мейл користувача (UserEmail);
- Додаткова інформація про користувача (Additional_inf).



а



б

Рис. 4. Таблиця з інформацією про користувачів (а) та таблиця з даними (на прикладі геотермічного пункту спостережень “Косино”) (б)

Друга підмножина – таблиці даних, у яких зберігаються дані з пунктів спостережень. Ці таблиці мають приблизно таку структуру (рис. 4, б):

- Ключ;
- Назва станції;
- Назва геофізичного методу;
- Назва типу апаратури;
- Назва типу і номер каналу;
- Дата та час вимірювання;
- Геофізичні дані

Клієнтський модуль

Модуль містить:

- систему аутентифікації користувачів;
- систему вибору та візуалізації геофізичних даних і формування їх у клієнт-файл.

Завдання останньої зокрема візуалізувати вибрані дані у формі графіків та діаграм. Для цього використано безкоштовну бібліотеку jfreecharts, яка надає функції для побудови графіків та діаграм і подальше їхнє збереження у вигляді графічних об’єктів. Графіки та діаграми будуються на основі колекції об’єктів DefaultCategoryDataset.

Клієнтський модуль повинен надавати користувачу інтуїтивно зрозумілий та зручний інтерфейс для взаємодії з базою даних і опрацювання геофізичної інформації.

Основні функції клієнтського модуля:

- Візуалізація даних (відображення графіків, числових значень та ін.);
- Коректування даних у базі даних (відповідно до рівня доступу);
- Створення файлів з вибраними даними для завантаження на комп’ютер користувача.

Для роботи з інтернет-порталом створено інтерфейс, через який формуються запити на вибір даних для візуалізації.

Під час розроблення клієнтського модуля використано бібліотеку GWT (Google Web Toolkit). Ця бібліотека дає змогу писати програмний код мовою Java, який потім перетворюється на JavaScript та XML і працює за технологією AJAX.

Під час розроблення інтерфейсу клієнтського модуля реалізовано такі функції:

- Створення таблиці, в якій зображаються дані зі станцій;
- Створення керівних елементів;
- Створення списку переглядання (станції, методи, апаратура);
- Створення службових часових міток (служба протоколу активності).

До цього модуля також підключено розроблені нами раніше утиліти, призначені для конвертування та препроцесингу геофізичних даних [16–18], зокрема утиліту “Converter-D”, ця утиліта працює у середовищі Windows.

Алгоритм роботи клієнтського модуля наведено на рис. 5, а приклад його роботи – на рис. 6.

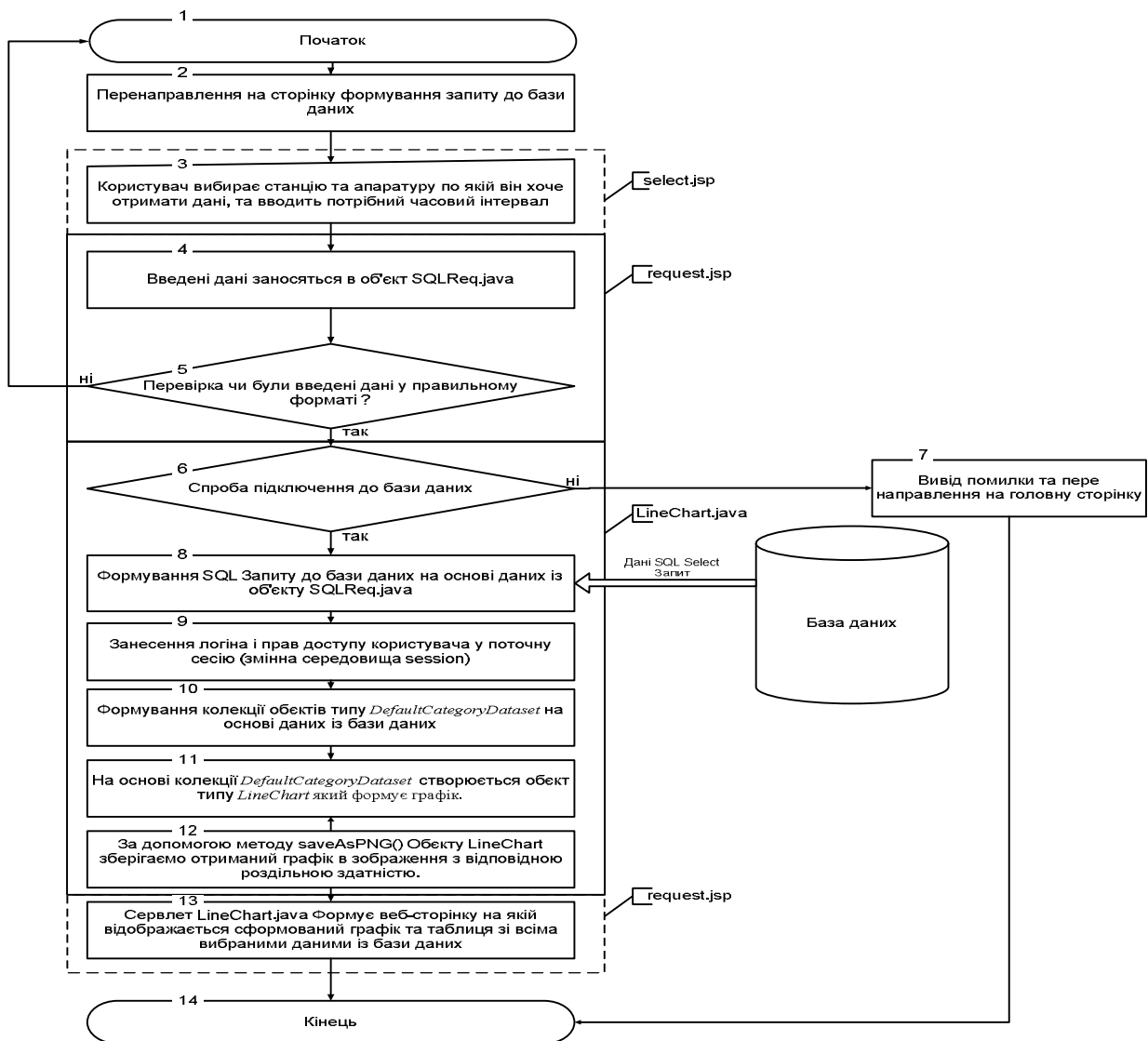


Рис. 5. Схема алгоритму роботи системи візуалізації

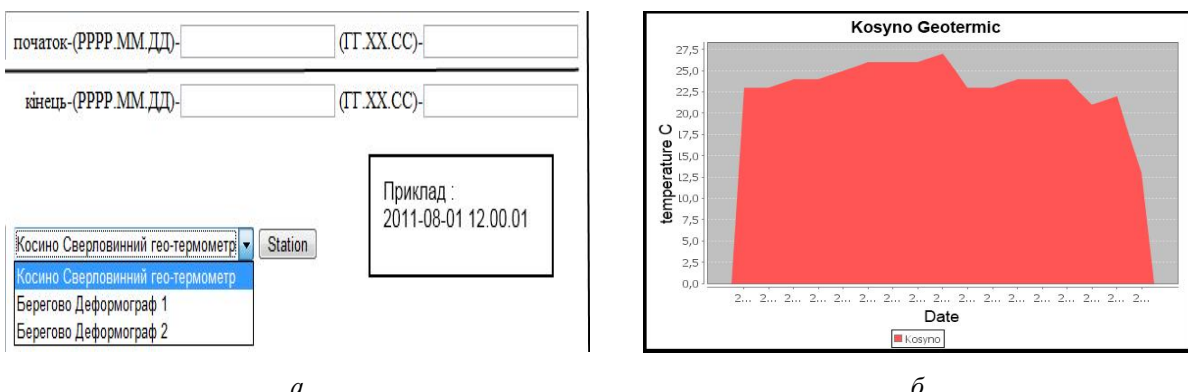


Рис. 6. Інтерактивні інтерфейсні сторінки клієнтського модуля:

а – сторінка формування запиту до бази даних; б – графічна візуалізації температурних даних з пункту “Косино” (на прикладі зовнішньої температури повітря)

Клієнтський модуль також забезпечує доступ до інформації в базі даних на двох рівнях з диференціацією по підрівнях відповідно до результатів аутентифікації користувача (алгоритм такої аутентифікації наведено на рис. 7).

Рівні доступу (клієнтські)

Перший рівень (**R1**) (гість) – доступ до перегляду даних (без можливості їх редагування).

Другий рівень (**R2**) (користувач) – доступ до перегляду та редагування і обробки даних.

Підрівні доступу по:

- Станції, на якій проводились виміри;
- Геофізичному методу;
- Апаратурі.

Рівень доступу **R0** – це рівень доступу адміністратора сервера, що надає право доступу до контролю та модифікації програмного забезпечення.

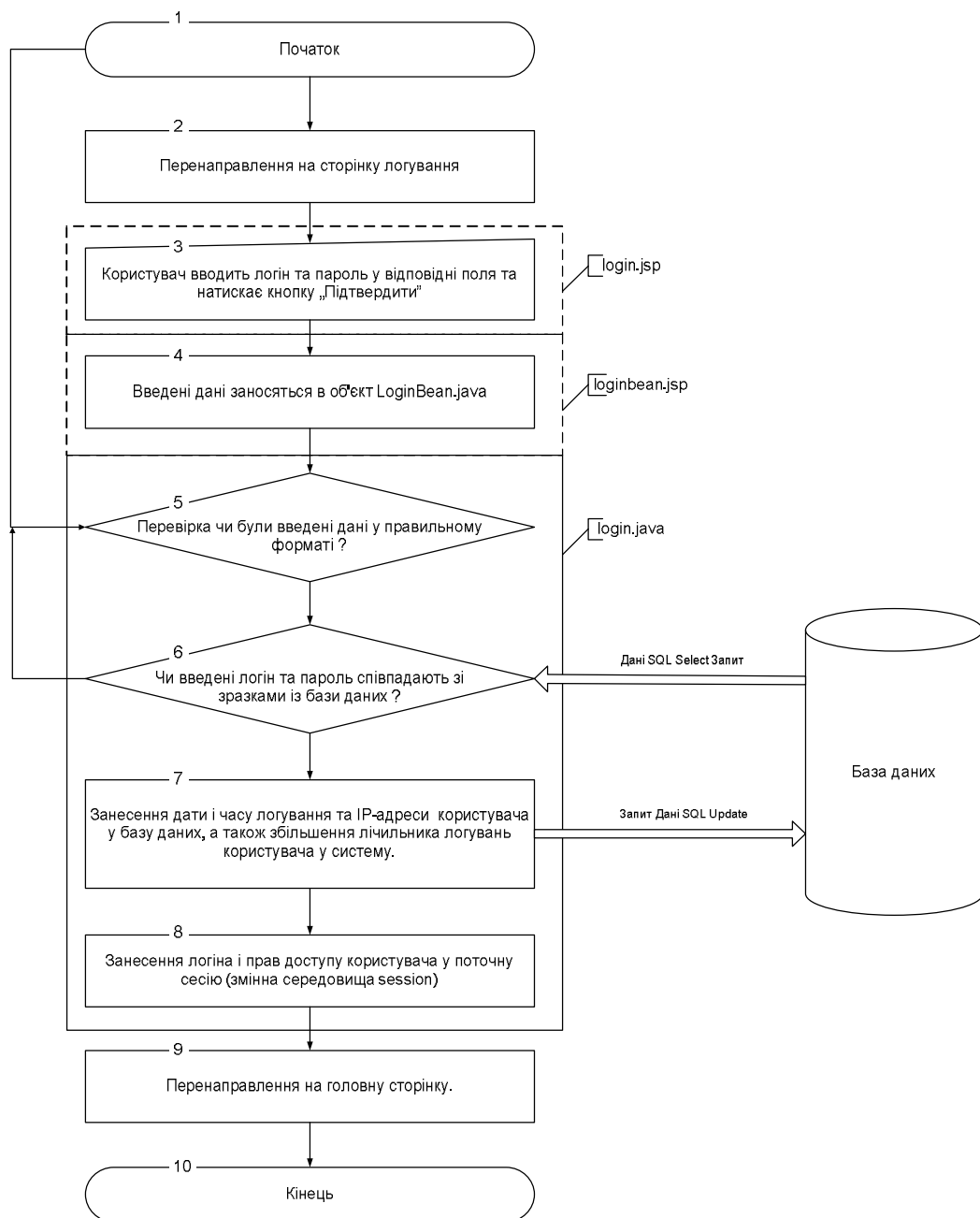


Рис. 7. Схема алгоритму роботи системи аутентифікації користувача

Що стосується подальшої обробки геофізичних даних та вимог до формування файлів, які мають видаватися системою на запит користувача, то відповідні приклади наведено зокрема в роботах [16–18], де показано способи попередньої обробки (препроцесингу) деформографічних даних з використанням розробленої нами утиліти “Converter-D”.

Тестування програмного продукту

Функціональне тестування створеного прототипу інтернет-порталу опрацювання даних геодинамічного моніторингу проведено під спеціально створеним тестовим обліковим записом на прикладі геотермічних даних з свердловинного пункту “Косино”, який є одним з перших у планах підключення до Інтернету. Приклад тестового запиту та графічного виводу даних наведено на рис. 7, а приклад тестування режиму аутентифікації користувача – на рис. 8.

Аутентифікація користувача

Логін	<input type="text" value="test"/>
Пароль	<input type="password" value="...."/>
<input type="button" value="Підтвердити"/>	

Ви ввели неіснуючий логін та пароль

[На Головну сторінку](#)

Рис. 8. Тест сторінки аутентифікації користувача

Висновки

Отже, створене програмне забезпечення інтернет-порталу геофізичного моніторингу вирішує поставлені задачі та реально експлуатується. У міру розвитку обладнаної Інтернетом мережі моніторингових геофізичних спостережень воно може нарощуватись і вдосконалюватись відповідно до потреб, оскільки володіє достатньою гнучкістю та масштабованістю.

1. Геофизическая служба Российской академии наук [електронний ресурс]: <http://www.gsras.ru/> – 2014.
2. Global Network for the Forecasting of Earthquakes [електронний ресурс] <http://seismonet.org/> – 2014.
3. European-Mediterranean Seismological Centre [електронний ресурс] <http://www.emsc-csem.org/> – 2014.
4. Gravity time dependency research (Європейська мережа маятникових нахиломірних станцій) [електронний ресурс] <http://www.dynamicgravity.org/> – 2014.
5. Шевчук Б.М., Задірака В.К., Гнатів Л.О., Фрасер С. Технологія багатофункціональної обробки і передачі інформації в моніторингових мережах. – К.: Наук. думка, 2010. – 368 с.
6. Рикитакє Т. Предсказание землетрясений. – М.: Мир, 1979. – 388 с.
7. Аранович З.И., Маламуд А.Я., Негребецкий С.А., Трапезников Н.Л. Метрология, принципы построения и функциональные особенности сейсмометрических регистрирующих систем / Автоматизация сбора и обработки сейсмологической информации. – М.: Радио и связь. – 1983. – С. 5–19.
8. Максимчук В.Ю., Кузнецова В.Г., Вербицкий Т.З. та ін. Дослідження сучасної геодинаміки Українських Карпат. – К.: Наук. думка. – 2005. – 256 с.
9. Назаревич А.В. Автоматическая система передачи результатов геофизических наблюдений по линии электросвязи. // Современные геодинамические процессы и их изучение в связи с проблемой прогноза землетрясений. – К.: Наук. думка. – 1986. – С. 48–51.
10. Назаревич А.В. Автоматический цифровой геоакустический комплекс. // Сейсмопрогностические исследования на территории УССР. – К.: Наук. думка. – 1988. – С. 116–123.
11. Назаревич А.В. Оптимізація методико-апаратурного забезпечення мережі геодинамічних спостережень України // Праці НТШ. – Львів. – 1997. – С. 148–157.
12. Назаревич А., Назаревич Л. Оптиелектронний вимірювальний канал до кварцового деформографа / Геодинаміка. – 1999. – № 1(2). – С. 116–120.
13. Назаревич А.В., Назаревич Л.С., Баштейвич М.В., Назаревич О.В.,

Микита А.Ю. Геоінформаційні технології в геомоніторингових дослідженнях // Геодезія, картографія і аерофотознімання. – 2003. – № 63. – С. 266–271. 14. Лящук Д.Н., Назаревич А.В., Назаревич Л.Є. Геоелектромагнітноемісійний метод в моніторинзі локальних геодинамічних процесів // Вісник КНУ ім. Т.Шевченка. “Геологія”. – 2003. – № 26–27. – С. 92–97. 15. Кендзера О., Вербицький Т., Вербицький С., Вербицький Ю. Цифровий сейсмограф для регіональних спостережень та результати його випробувань. –Геодинаміка. – 1998. – №1. – С. 120–126. 16. Назаревич А.В., Мицик Б.Г., Баштевич М.В., Назаревич Р.А. Деформографічні дослідження сейсмотектонічних процесів в Українському Закарпатті (геоінформаційні аспекти). IX International Conference “Geoinformatics – Theory and Applied Aspekts”. 11-14 May 2010, Kyiv, Ukraine (CD). 17. Мицик Б.Г., Назаревич А.В., Баштевич М.В., Назаревич Р.А. Безконтактні смісні вимірювачі мікропереміщень у деформографічних геофізичних дослідженнях // Відбір і обробка інформації. – 2011. – Вип. 35 (111). – С. 69–76. 18. Назаревич Р., Мархивка В., Струк Є., Назаревич А. Конвертація та препроцесинг даних деформографічного моніторингу // Вісник НУ “ЛП” “Комп’ютерні науки та інформаційні технології”. – Львів. – 2011. – № 694. – С. 334–340. 19. Turner James. MySQL and JSP Web applications. – Sams Publishing, 2002. – 560 с. 20. Маркин А.В. Построение запросов и программирование на SQL. – Рязань, 2008, – 312 с. 21. Jeffrey Richter. Programming Applications for Microsoft Windows – Microsoft Press, 2004. – 723 p.

УДК 004

Р. Базилевич, В. Андрієнко

Національний університет “Львівська політехніка”,
кафедра програмного забезпечення

ДОСЛІДЖЕННЯ ТА АНАЛІЗ АЛГОРИТМІВ ОСТРІВКУВАННЯ ЕНЕРГЕТИЧНИХ МЕРЕЖ

© Базилевич Р., Андрієнко В., 2014

Досліджено алгоритми бісекціювання графів, їх застосовність до острівкування енергетичних мереж. Проаналізовано алгоритм Кернігана–Ліна, спектральний метод та підхід k-середніх, щодо обчислювальних затрат і придатності до застосування.

Ключові слова: бісекціювання графу, острівкування енергомережі.

The graphs bisectioning algorithms are investigated and their applicability to islanding of the power system. The Kernighan-Lin algorithm, spectral method and multilevel kernel k – means approach have been analyzed with respect to the computational complexity and fitness for use.

Key words: graph bisectioning, power system islanding.

Вступ

Кероване острівкування запобігає хвильовому поширенню аварії в енергетичній мережі з забезпеченням постачання енергії частині її споживачів. Розподілений генератор продовжує постачати живлення, навіть коли деякий з генераторів вилучений [1]. Існують два види острівкування – кероване (умисне) та некероване (природне, неумисне). При керованому острівкуванні генератор в аварійній частині сам від’єднується від мережі і розподілений генератор продовжує жити місцеву мережу. При некерованому острівкуванні генератор не встигає відключитись від живлення мережі після аварії в ній. Таке острівкування є небезпечним явищем, яке супроводжується хвильовим поширенням аварії, несподіваними “живими дротами” (такими, що перебувають під напругою), пошкодженням обладнання споживачів електроенергії та іншими небажаними подіями. Перенавантаження енергосистеми може призвести до нестабільної роботи, а

та, своєю чергою, до некерованого острівкування, яке веде до катастрофічних аварій. Кероване ж проводиться з метою локалізації аварії і збереження працездатності максимальної кількості споживачів енергомережі. Кероване острівкування може покращити відновлення систем, управління навантаження мережі і генерацію струму ній. Як зазначили Rajamani et al.(1999) [13] без такого підходу Індія могла б надовго залишатись без світла у 1999 році. За повідомленням Agematsu et al. (2001) [3] некероване острівкування, де не здійснюється контроль за балансуванням активної та реактивної потужності, у 1999 році надовго залишило б Токіо без світла, коли літак врізався в лінії електропостачання. Єдиний вихід з такої ситуації – кероване острівкування. Методи керованого острівкування мають діяти швидко і вчасно, повинні враховуватись витрати на їх розроблення та впровадження, бо якщо вони істотно зростатимуть із розмірами енергосистеми, то їх застосовність буде низькою.

Розглянемо методи та алгоритми керованого острівкування. Математичною моделлю енергетичної мережі є зважений граф $G(N,B)$, вершини якого відображають джерела постачання енергії, споживачів та розгалуження, а ребра – лінії передавання енергії. Ваги вершин характеризують потужність, напругу та інші характеристики джерел енергії, ваги ребер – параметри ліній її передавання. Метою керованого острівкування є відокремлення частини споживачів від джерела аварії з мінімізацією кількості ліній розриву. Тобто задача формально зводиться до розбиття графу на дві частини, одна з яких містить джерело аварії, а інша – решту мережі з мінімізацією кількості розривів мережі. З математичного погляду ця задача належить до комбінаторного типу класу NP , оскільки має експоненційну, а саме факторіальну обчислювальну складність. Для задач великих розмірностей методи отримання точного розв'язку, наприклад, гілок та границь, чи повний перебір всіх варіантів, є непридатними, оскільки вимагають дуже великих обчислювальних затрат. Тому більшість досліджень в області розбиття графів переважно мають на меті пошук ефективних евристичних алгоритмів[2]. Більшість запропонованих методів острівкування присвячені розбиванню енергомережі на дві частини, тобто її бісекціюванню.

Аналіз алгоритмів

Для керованого острівкування енергомереж та розбиття графів запропоновано ряд методів та алгоритмів [1–21]. В статті Peiravi and Pdarabadi (2009) [2] порівняно обчислювальні вимоги для двох підходів до бісекціювання енергосистеми – “спектрального методу” та “багаторівневого підходу ядра k -середніх”. Кращий розв'язок у керованому острівкуванні енергосистем можна отримати лише тоді, коли вона розривається на дві частини, оскільки такий підхід призведе до більших острівців порівняно з багатосекційним її розривом. Відновлення бісекціюваної енергосистеми є легшим, ніж розірваної на багато частин. Обчислювальну складність кожного підходу визначено для декількох великих та польської енергосистем. Karupis and Kumar (1999) [9] пропонують розв'язувати задачу за три етапи: укрупнення графу, його поділ та вдосконалення. Вони також представили швидшу модифікацію алгоритму Kernighan-Lin [11]. Метод k – середніх використовується до таких застосувань як сегментація зображення, соціальний і генний аналіз, мережевий аналіз систем [2]. Kernighan and Lin(1970) [11] запропонували евристичний ітераційний алгоритм бісекціювання, починаючи з початкового ділення навпіл, де в кожній ітерації відшукується підмножина вузлів з кожного попереднього розбиття з найменшою кількістю ребер у перетині.

Cherng et al. (1999)[4] розробили дворівневий алгоритм бісекціювання, який поєднує гібридний метод об'єднання в кластери з повторюваним процесом розбиття інтегрованих схем. Пізніше, Cherng et al. (2003) [5] представили багаторівневий алгоритм бісекціювання шляхом інтеграції техніки кластеризації і повторюваного покращення. Областю застосування цього методу є енергетичні системи, де поділ графу може бути використаний для різноманітних цілей. Задача полягає у відокремленні шини з системи на дві чи більше груп, щоб задовольнити визначену мету. Ця техніка використовується Wang and Vittal (2004)[17] на основі групування даних.

Wang (2008) [16] запропонував адаптивний алгоритм кластеризації мережі з врахуванням початкового потоку потужності і отриманої архітектури, що допоможе дослідити її перехідну стабільність. Розміри малих груп визначають з використанням ряду обмежень. Спектральний аналіз

Лапласа для розбиття графу був застосований Mohar (1991) [18], Hagen (1992)[7] та іншими. При бісекціюванні графи розбиваються на дві частини, використовуючи друге найменше власне значення матриці Лапласа. Відповідні власний та обчислений вектор отримуються на основі розбиття. Щоб сформувати матрицю Лапласа, необхідно мати матрицю суміжності графу. Матриця Лапласа ненапрямленого незваженого графу є симетричною з одним рядком і колонкою для кожного вузла. Якщо власні значення матриці Лапласа відсортовано за збільшенням, власний вектор, відповідний другому найменшому значенню матриці Лапласа (Fiedler вектор) може бути використаний у евристичних алгоритмах для різноманітних маніпуляцій з графами, включно зі спектральним розбиттям графу. Друге найменше власне значення матриці Лапласа є більшим за 0 тільки якщо граф є зв'язним. Матриця Лапласа формується так: елемент, для якого $i = j$, є сумою гілок, прилеглих до відповідної вершини. Інші значення визначаються як протилежне значення гілки, яка спрямована від вершини i до вершини j . Далі визначаються власні значення і матриця власних векторів. Другий стовпчик (або рядок) цієї матриці називається Fiedler вектор, його може бути використано для бісекціювання графу, як показано на рис. 2, де по осі ординат - власний вектор та по осі абсцис - вузли. Обчислювальні затрати для визначення власних значень і Fiedler вектор є "вужким місцем" алгоритму бісекціювання, оскільки мають обчислювальну складність $O(n^3)$. Визначення Fiedler вектор є найбільш затратною операцією в обчислювальному відношенні частиною розбиття.

Для пришвидшення отримання розв'язку в багатьох випадках Fiedler вектор апроксимують. Багаторівневі версії спектрального методу, які використовують його на різних рівнях, широко використовуються. Holzrichter and Oliveira (1999) [8] запропонували спектральний підхід, в якому визначення вектору Fiedler здійснюється з застосуванням Davidson algorithm [8].

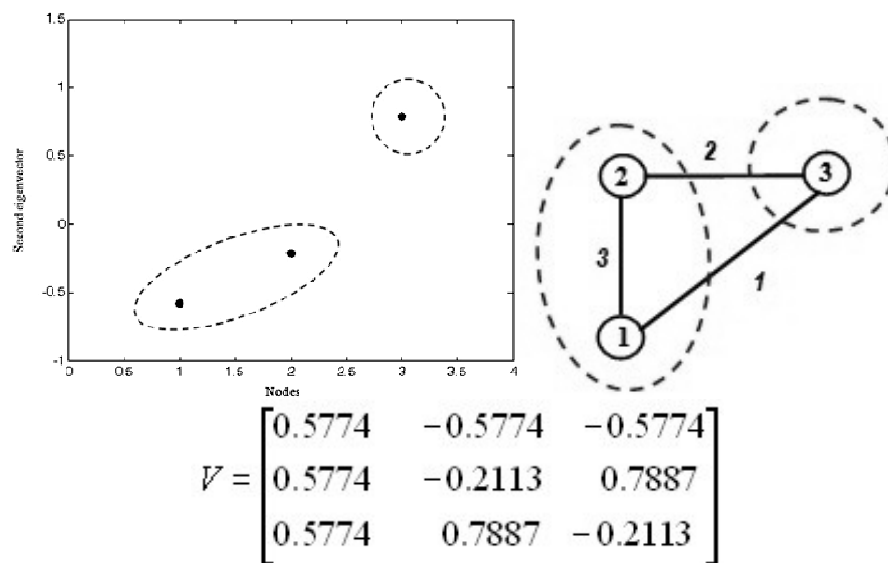


Рис. 1. Розбиття графу з використанням Fiedler вектора [2]

Хоча мінімальна нестабільність генерації/навантаження є дуже важливим фактором в острівкуванні, інші вимоги також необхідно враховувати. За пропозицією Rehtanz (2003) [14] спектральне бісекціювання енергосистеми може бути виконане так:

1. Визначити енергетичний потік мережі.
2. Моделювати енергомережу графом $G(N,B)$ з n вершинами і b ребрами:
 - 2.1. Кожний вузол енергомережі є вершиною графу G .
 - 2.2. Кожна лінія передавання енергомережі є ребром графу G .
 - 2.3. Визначити вагу кожного ребра графу G як абсолютну величину потоку потужності лінії передавання.
3. Визначити матрицю Лапласа Q графу $G(N,B)$:

3.1. Обчислити матрицю суміжності A та діагональну матрицю D графу $G(N,B)$.

3.2. Задати матриця Лапласа як $Q = D-A$.

4. Обчислити друге найменше власне значення λ_2 матриці Лапласа Q .

5. Обчислити зв'язний з λ_2 дійсний власний вектор x_2 .

6. Відобразити x_2 в евристичному розбитті вектора графу $G(N,B)$.

6.1. Впорядкувати записи x_2 , щоб отримати посортований вектор v індексів вузла.

6.2. Помістити всі вузли в розбиття U .

6.3. Для $i=1$ до $n-1$:

Перемістити вектор π_i з розбиття U до розбиття W . Обчислити скорочення розміру розбиття.

7. Знайти оптимальне розбиття (U^*, W^*) , котре матиме мінімальне скорочення встановленого розміру з поміж $(n-1)$ різноманітних розбиттів для знаходження кінцевого результату.

Спектральне розбиття і використання багаторівневого ядра k середніх схожі математично, і це допомагає розробити алгоритм розбиття, кращий щодо швидкості, зберігання пам'яті і якості, як показано в Dhillon (2005) [6]. Peiravi and Idarabadi (2009) [11] запропонували застосувати метод багаторівневого ядра k середніх для керованого острівкування. Оскільки в цьому підході немає необхідності обчислювати власні значення і власні вектори, він є набагато швидшим, ніж спектральний метод. Підхід використовує запропонований цими авторами метод Metis.

Для керованого острівкування доцільно також використати метод оптимального згортання схеми [19–23]. За цим методом вперше (1974 р.) запропоновано розбиття схеми здійснювати за декілька етапів: виділення ієрархічно вкладених кластерів шляхом згортання схеми за вибраними критеріями, вибір на дереві згортання бажаного початкового розбиття, найкраще наближення до заданих обмежень та його оптимізація. Як показали дослідження, метод має лінійно-логіфімічну обчислювальну складність, що робить його придатним для задач великих розмірностей, можливість забезпечення різноманітних обмежень, зокрема вибору довільної кількості частин розбиття, виконання багаторівневого розбиття та забезпечує високу якість отриманих розв'язків.

Обчислювальна складність алгоритмів

Обчислювальні складності алгоритмів острівкування енергосистеми залежать від їх методики. Це “вузьке місце” підходу визначає можливість застосування запропонованого алгоритму, оскільки необхідною є швидка реакція у випадку виникнення аварії. Для спектрального підходу така складність визначається кількістю затрат, необхідних на обчислення власних значень і власних векторів матриці Лапласа. Для щільної матриці Лапласа ця складність є $O(n^3)$. У реальних енергосистемах, які описуються сильно розрідженим графом, це не вимагало б значних обчислювальних затрат.

Обчислювальна складність в підході багаторівневого ядра k - середніх є $O(n^2)$ для кожного повторення. У сучасних великих енергосистемах, де має місце висока розрідженість, кожна ітерація буде мати тільки $O(nz)$ операцій, де z означає кількість відмінних від нуля записів в матриці, яка є пропорційною до кількості ліній енергосистеми. Тут немає потреби обчислювати власні вектори, що значно зменшує обчислювальну складність і робить її близькою до $O(nz)$, що є набагато менше, ніж для спектрального підходу.

Peiravi and Idarabadi (2009) [2] дослідили спектральний і підходу багаторівневого ядра k - середніх для трьох типових енергосистем: IEEE 118, IEEE 300 і польської 2746. Набагато менше обчислювального часу потрібно для підходу багаторівневого ядра k – середніх в порівнянні із спектральним підходом (рис. 2).

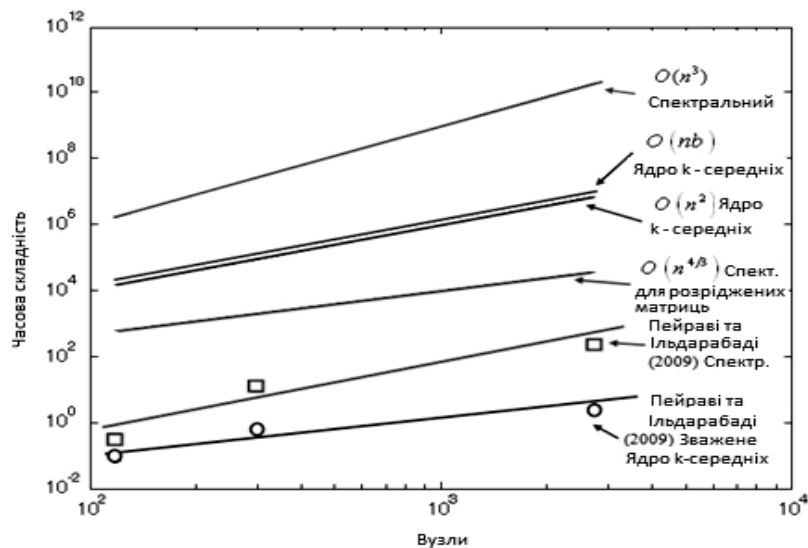


Рис. 2. Порівняння очікуваних вимог часу[2]

Висновки

Всі операції зміни структури енергосистеми для її острівкування при виникненні аварії повинні бути виконані за дуже малий час. Вважається, що всі обчислення для керованого острівкування повинні бути виконані за час, менший ніж 500 мс [2]. Багато запропонованих методик острівкування є повільними для оперативного застосування до великих енергосистем. Аналізуючи наявні методи і алгоритми їх острівкування, можна зробити висновок, що їх “вузьким місцем” є велика обчислювальна складність, не завжди достатня якість та можливість врахування бажаних обмежень. Якщо мережа сама не встигає за певний час локалізувати проблему, то кероване острівкування може підвищити її надійність. Погана ситуація наступає, коли нестабільність напруги в мережі разом із дисбалансом між навантаженням мережі та генерацією струму, призведе до непоправимих наслідків. Єдиний вихід з цієї ситуації – це кероване острівкування. Для цього необхідно розробити швидкі, надійні і не дуже затратні методи та програмне забезпечення, які б вирішили проблему острівкування енергетичної мережі під час збою в системі, та інші подібні проблеми. Бажаним з цією метою є застосування методу оптимального згортання схеми [19-23].

1. <http://uk.wikipedia.org/wiki/Острівкування>. 2. Peiravi A., R. Ildarabadi. “Comparison of Computational Requirements for Spectral and Kernel k-means Bisectioning of Power Systems”, *Australian Journal of Basic and Applied Sciences*, 3(3): 2366-2388, 2009. 3. Agematsu, S., S. Imai, R. Tsukui, H. Watanabe, T. Nakamura, T. Matsushima. “Islanding Protection System with Active and Reactive Power Balancing Control for Tokyo Metropolitan Power System and Actual Operational Experiences,” in *Proceedings of the 7th IEE Int. Conf. Developments in Power System Protection*, pp: 351–354. , 2001. 4. Cherng, J., S. Chen, C. Tsai, J. Ho, 1999. An efficient two-level partitioning algorithm for VLSI circuits, *Proceedings of the 1999 Design Automation Conference, ASP-DAC '99, Asia and Pacific*, 1: 69-72, Wanchai, Hong Kong, 18-21 Jan. 1999. 5. Cherng, J., S. Chen, 2003. An efficient multi-level partitioning algorithm for VLSI circuits, *Proceedings of the 16th 70-75. International Conference on VLSI Design (VLSI'03)*, 4-8 January 2003. 6. Dhillon, Inderjit S., Guan, Yuqiang, Kulis, Brian, 2005. “A Fast Kernel Based Multilevel Algorithm for Graph Partitioning,” In the *Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery Data Mining (KDD)*, pp: 629-634. 7. Hagen, L., A.B. Kahng, 1992. “New Spectral Methods for Ratio Cut Partitioning and Clustering,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 11(9): 1074-1085. 8. Holzrichter, M., S. Oliveira, 1999. A graph based method for generating the Fiedler vector of irregular problems, *Proceedings of the 11 IPPS/SPDP'99 Workshops Held in Conjunction with the 13th International Parallel Processing Symposium and 10th Symposium on Parallel and Distributed Processing*, San

Juan, P.R., pp: 978-985. 9. Karypis, G., V. Kumar, 1999. "A Fast and High Quality Multilevel Scheme for Partitioning Irregular Graphs", *SIAM J. Scientific Computing*, 20(1): 359-392. 10. <http://glaros.dtc.umn.edu/gkhome/fetch/papers/mlSIAM SC99>. 11. Kernighan, B.W., S. Lin, 1970. "An efficient heuristic procedure for partitioning graphs," *The Bell System Technical Journal*, 49(2). Peiravi, A., R. Ildarabadi, 2009. A fast algorithm for intentional islanding of power systems using the multilevel kernel k-means approach, *Journal of Applied Sciences*, 9(12): 2247-2255. 12. Rajamani, K., U.K. Hambarde, 1999. "Islanding and Load Shedding Schemes for Captive Power Plants," *IEEE Trans. on Power Del.*, 14(3): 805–809. 13. Rehtanz, C., 2003. *Autonomous systems and intelligent agents in power system control and operation*, Springer. 14. Venkatasubramanian, V., J. Quintero, 2005. "Detection, Prevention and Mitigation of Cascading Events-Part II", *Power System Engineering Research Center, A National Science Foundation Industry/University Cooperative Research Center Since 1996, PSERC Publication 05-60*. 15. Vittal, V., Xiaoming Wang, 2005. "Detection, Prevention and Mitigation of Cascading Events-Part III", *Power System Engineering Research Center, A National Science Foundation Industry/University Cooperative Research Center Since 1996, PSERC Publication 05-61, October 2005*. 16. Wang, X.Z., Z. Yan, W. Xue, 2008. "An adaptive clustering algorithm with high performance computing application to power system transient stability simulation", *3rd International Conference on Deregulation and Restructuring and Power Technologies, DRPT 2008*, art. no. 4523578, pp: 1137-1140. 17. Wang, X., Vittal, V., 2004. "System Islanding using Minimal Cutsets with Minimum Net Flow", *In Proceedings of the IEEE Power Eng. Soc. Power Syst. Conf. Expo., New York, 1*: 379-384. 18. Mohar, B., 1991. "The Laplacian Spectrum of Graphs," in *Graph Theory, Combinatorics, and Applications*, Vol. 2, Ed. Y. Alavi, G. Chartrand, O. R. Oellermann, A. J. Schwenk, Wiley, New York, pp: 871-898. 19. Базилевич П.П. "Декомпозиционные и топологические методы автоматизированного проектирования электронных устройств", Львов, "Вища школа", 1981, 168 С. 20. R.P. Bazylevych, R.A. Melnyk, O.G. Rybak. "Circuit partitioning for FPGAs by the optimal circuit reduction method". In: *VLSI Design*, Vol. 11, No 3, pp. 237-248, 2000. 21. Bazylevych R.P. "The optimal circuit reduction method as an effective tool to solve large and very large size intractable combinatorial VLSI physical design problems". In: *10-th NASA Symp. on VLSI Design*, March 20-21, 2002, Albuquerque, NM, USA, pp. 6.1.1-6.1.14. 22. R. Bazylevych, I. Podolsky and L. Bazylevych. "Partitioning optimization by recursive moves of hierarchically built clusters". In: *Proc. of 2007 IEEE Workshop on Design and Diagnostics of Electronic Circuits and Systems*. April, 2007, Krakow, Poland, pp. 235-238. 23. Roman Bazylevych, Marek Pałasiński, Dmytro Yanush, Lubov Bazylevych. "Partitioning Optimization by Iterative Reassignment of the Hierarchically Built Clusters with Border Elements", *2nd Mediterranean Conference on Embedded Computing, MECO – 2013, Budva, Montenegro*, pp.219-222.

М. Копитчук, В. Передерій

Одеський національний політехнічний університет,
кафедра комп'ютерних інтелектуальних систем та мереж
Миколаївський національний університет ім. О.М. Сухомлинського,
кафедра комп'ютерних систем та мереж

ВИЗНАЧЕННЯ НЕЧІТКИХ СИТУАЦІЙ ВПЛИВУ НАВКОЛИШНЬОГО СЕРЕДОВИЩА НА КОГНІТИВНИЙ СТАН ОСОБИ, ЩО ПРИЙМАЄ РІШЕННЯ В ЕРГАТИЧНИХ СИСТЕМАХ

© Копитчук М., Передерій, В., 2014

Розроблено нечіткі класифікатори, які дають можливість описати з максимальною точністю фактори, для яких немає відомих точних закономірностей і для яких необхідно провести асоціацію між якісною та кількісною оцінками факторів, які впливають на особу при прийнятті рішення.

Ключові слова: особа, що приймає рішення, фактори навколишнього середовища, нечіткі множини, нечіткі класифікатори.

Fuzzy classifiers which allow to accurately describe the factors for which there are no strict regularities and one must associate quantitative and qualitative assessments of factors which influence the individual while decision-making.

Key words: individual making a decision, environmental factors, fuzzy sets, fuzzy classifiers.

Вступ

Особливості людино-машинних систем керування в тому, що вони містять як елементи і технічні системи, і людей, що взаємодіють із цими системами. Для ефективного функціонування схожих систем необхідно вибирати раціональні способи взаємодії когнітивних характеристик користувачів з технікою на підставі висновків ергономіки та впливу зовнішнього середовища. Якщо функціонування ергатичної системи відбувається в умовах невизначеності, то якість керування забезпечується якістю роботи особою, що приймає рішення (ОПР).

При цьому основні труднощі зв'язані не тільки з удосконалюванням технічних і програмних засобів, але і з недостатнім розвитком методів обліку людського фактора і неможливістю вірогідно прогнозувати вплив на систему зміни навколишнього середовища, при створенні та експлуатації складних систем. Внаслідок цього високу актуальність набуває завдання розроблення способів визначення нечітких ситуацій впливу навколишнього середовища на когнітивний стан особи що приймає рішення.

Постановка проблеми

У роботах [4–5] автори проаналізували основні причини, які негативно впливають на якість роботи та порушення працездатності особи, що приймає рішення в людино-машинних системах управління складними об'єктами, де однією з головних причин є людський фактор.

У літературі [5] детально розглянуто питання визначення комфортного робочого середовища користувача під час функціонування системи. У [5,6] описано створення математичних моделей і алгоритмів, що дають змогу оцінити релевантність прийнятих рішень з урахуванням впливу зовнішніх і особистих факторів на безпеку роботи людино-машинних систем. У [6] автори розробили алгоритми формалізації взаємозв'язку зовнішніх факторів і психофункціональних характеристик користувачів на основі теорії нечітких множин і алгоритм для оптимізації

прийняття рішень. Але все це не дає можливості описати з максимальною точністю фактори, для яких немає відомих точних закономірностей і для яких необхідно провести асоціацію між якісною та кількісною оцінками факторів, які впливають на особу при прийнятті рішення.

Мета роботи

Розробити спосіб визначення нечітких ситуацій впливу навколишнього середовища на когнітивний стан особи, що приймає рішення на основі нечітких класифікаторів.

Виклад основного матеріалу

Особливого значення в діяльності особи, що приймає рішення, в ергатичних системах набувають неконтрольовані фактори, які можна поділити на дві групи [6]:

1. Ергономічні та фактори зовнішнього середовища Sc (табл. 1):
 - інтенсивність шуму IN ;
 - інтенсивність вібрацій IV ;
 - освітленість робочого місця E ;
 - температура T ;
 - вологість f ;
 - атмосферний тиск P .
2. Фактори, зумовлені когнітивним станом Sp , (табл.2):
 - рівень інформаційної пропускнуої здатності I_0 ;
 - ступінь втоми користувача F ;
 - обмеженість часу прийняття рішення Tr ;
 - невідповідність ступеня напруженості TS ;
 - сконцентрованість уваги A .

Особливість цих факторів полягає в їхньому нечисловому характері, неможливості представити їх у вигляді дійсного числа, а також взаємної залежності між двома групами факторів.

Задачу розв'язують, знаходячи максимум цільової функції:

$$R_p = F(Sc^+, Sp^+) \rightarrow \max, \quad (1)$$

де стан першої групи факторів

$$Sc = f_1 \left(I_N^-, I_V^-, E^+, T^{+-}, f^+, \left(\frac{\Delta P}{\Delta t} \right)^- \right) \quad (2)$$

стан другої групи факторів

$$Sp = f_2 \left(I_0^+, F^-, T_p^-, TS^-, A^+ \right) \quad (3)$$

Підставивши (2) і (3) до (1), одержимо цільову функцію факторів, які впливають на особу, що приймає рішення в ергатичних системах:

$$R_p = F \left[f_1 \left(I_N^-, I_V^-, E^+, T^{+-}, f^+, \left(\frac{\Delta P}{\Delta t} \right)^- \right)^+, f_2 \left(I_0^+, F^-, T_p^-, TS^-, A^+ \right)^+ \right] \rightarrow \max \quad (4)$$

Застосуємо при описі факторів теорію нечітких множин. Для визначення нечіткої множини розглядається деяка множина U , що складається з елементів u і відображення із цієї множини на відрізок $[0,1]$ - $\mu : U \rightarrow [0,1]$.

Визначення: Нечіткою множиною A , заданою на базовій множині U , є множина пар $A = \{ \mu_A(u) / u : u \in U \}$, де $\mu_A(u)$ називається функцією приналежності нечіткої множини A [1].

У такий спосіб для завдання нечіткої множини A потрібне завдання його функції приналежності $\mu_A(u)$ на деякій базовій множин U . Приклад нечіткої множини “середня інтенсивність шуму” наведено на рис. 1, де U – множина значень інтенсивності шуму.

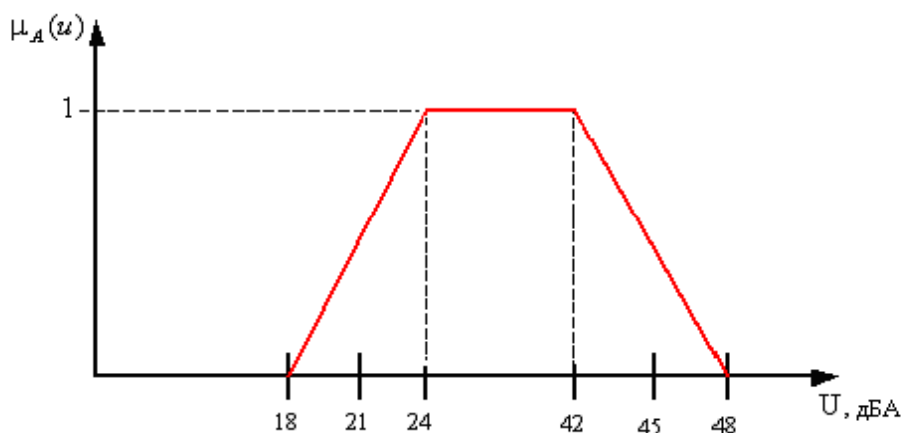


Рис. 1. Нечітка множина “середня інтенсивність шуму”

Використовуючи поняття нечіткої множини, можна перейти під час опису факторів від числових змінних до лінгвістичних змінних (ЛП), значення яких є деяким набором слів природною або формальною мовою. ЛП дозволяють приблизно описувати явища, які з тих або інших причин неможливо описати за допомогою кількісних характеристик.

Визначення: лінгвістичною змінною називається кортеж з п'яти елементів $(X, T(X), U, G, M)$, де X – ім'я ЛП, $T(X)$ – терм-множина ЛП, тобто множина лінгвістичних значень, кожне з яких являє собою нечітку множину на базовій множині U ; G – множина синтаксичних правил, що породжують імена значень ЛП, M – множина семантичних правил, що ставлять у відповідність кожному терму ЛП відповідну нечітку множину [2].

Приклад ЛП, що описує фактор інтенсивність шуму, наведено на рис. 2.

Отже, множина лінгвістичних змін a_i , що описує цю предметну область, визначається множиною нечітких значень $A_i = \{U_i^k\}_{k=1..K_i}$, де K_i – кількість нечітких значень, прийнятих i -м параметром, у вигляді нечітких чисел із трапецеїдальною функцією приналежності m_i^k , яку позитивно визначено на деякому інтервалі (u_{ib}^k, u_{ie}^k) , де $u_{ib}^k, u_{ie}^k \in U_i$ – значення початку й кінця інтервалу відповідно, а U_i – базова множина нечітких значень параметра a_i .

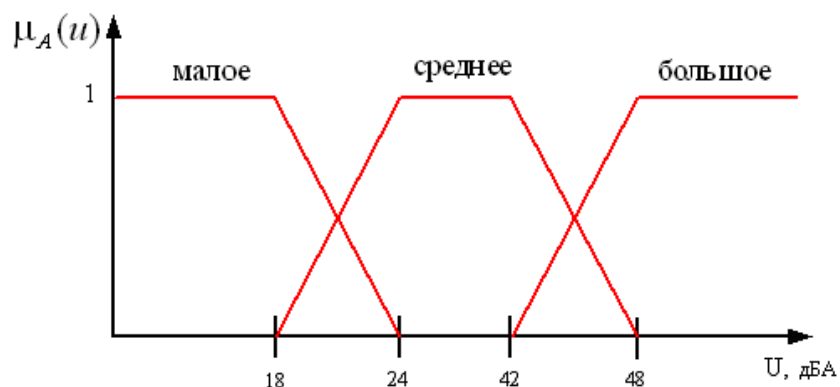


Рис. 2. Лінгвістична змінна “інтенсивність шуму”

Кожне нечітке число $U_i^k \in A_i$ визначимо через функцію приналежності вигляду:

$$U_k^i \Rightarrow m_k^i(u_i) = \begin{cases} 0, & u_i \leq u_{kb}^i, u_i \geq u_{ke}^i \\ \frac{u_i - u_{kb}^i}{u_{kb_1}^i - u_{kb}^i}, & u_{kb}^i < u_i < u_{kb_1}^i \\ 1, & u_{kb_1}^i \leq u_i \leq u_{ke_1}^i \\ \frac{u_i - u_{ke}^i}{u_{ke_1}^i - u_{ke}^i}, & u_{ke_1}^i < u_i < u_{ke}^i \end{cases}, \quad i = 1..N_u, k = 1..3. \quad (5)$$

де u_{kb}^i, u_{ke}^i – початкове і кінцеве значення відповідно інтервалу значень базової множини U_i , на якому функцію приналежності k -го нечіткого значення i -го параметра позитивно визначено; $u_{kb_1}^i, u_{ke_1}^i$ – початкове й кінцеве значення відповідно інтервалу значень базової множини U_i , на якому функція приналежності k -го нечіткого значення i -го параметра дорівнює одиниці.

Функції приналежності термів ЛПП “інтенсивність шуму” у силу своєї трапецеїдальної форми можуть бути описані кортежем із чотирьох значень $(u_{kb}^i, u_{kb_1}^i, u_{ke_1}^i, u_{ke}^i)$, що визначають початкову й кінцеву абсцису нульового й одиничного α -рівнів. Для функцій приналежності, зображених на рис. 2, можна дати опис у вигляді кортежів із чотирьох значень

$$\begin{aligned} m_H &= (0, 0, 18, 24); \\ m_C &= (18, 24, 42, 48); \\ m_B &= (42, 48, 75, 75). \end{aligned}$$

Таблиця 1

Формалізація факторів стану навколишнього середовища в лінгвістичні змінні

Первинний фактор	Границі зміни (U_i)	Терми лінгвістичної оцінки		
		“Н”(низький) m_H ,	“З”(середній) m_C ,	“В”(високий) m_B ,
IN (дБА)	0–75	(0,0,18,24)	(18,24,42,48)	(42,48,75,75)
IV (мм/с)	0–15	(0,0,3,5)	(3,5,9,11)	(9,11,15,15)
E (Лк)	0–600	(0,0,160,240)	(160,240, 360,440)	(360,440, 600, 600)
T (°З)	10–40	(10,10,15,19)	(15,19,23,27)	(23,27,40,40)
f (%)	0–90	(0,0,26,34)	(26,34,56,64)	(56,64,90,90)
$\frac{\Delta P}{\Delta t}$ (мм.рт.ст/сут)	0–15	(0,0,2.5,3.5)	(2.5,3.5,6,8)	(6,8,15,15)

Таблиця 2

Формалізація факторів когнітивного стану користувача в лінгвістичні змінні

Первинний фактор	Границі зміни	Терми лінгвістичної оцінки		
		“Н”(низький) m_H ,	“З”(середній) m_C ,	“В”(високий) m_B ,
I _o (ранг)	0–15	(0,0,4,6)	(4,6,9,11)	(9,11,15,15)
F (ранг)	0–15	(0,0,4,6)	(4,6,9,11)	(9,11,15,15)
Tr (хв)	0–60	(26,34,60,60)	(4,6,26,34)	(0,0,4,6)
TS (ранг)	0–15	(0,0,4,6)	(4,6,9,11)	(9,11,15,15)
A (ранг)	0–15	(0,0,4,6)	(4,6,9,11)	(9,11,15,15)

Нечіткі класифікатори

Розглянемо лінгвістичну змінну (ЛП) $B^{(5)}$ з іменем “Рівень показника” і терм-безліччю значень B_1 “Дуже Низький (ВН), B_2 Низький (Н), B_3 Середній (З), B_4 Високий (В), B_5 Дуже Високий (ОВ)”, які наведено на рис. 3. Базовою множиною цієї змінної є відрізок $[0,1]$, а кожний терм ЛП B_i , $i=1,\dots,5$ описується трапецеїдальною функцією приналежності, що задовольняє співвідношенням (5). За аналогією з [3] розглянемо набір вузлових точок $\alpha = \{0.1, 0.3, 0.5, 0.7, 0.9\}$, які є абсцисами максимумів функцій приналежності термів ЛП $B^{(5)}$ і водночас рівномірно знаходяться один від одного на відрізку $[0,1]$ і симетрично щодо абсциси 0.5. Сукупність ЛП $B^{(5)}$ і множини вузлових точок α назовемо стандартним нечітким п’ятирівневим 01-класифікатором [3].

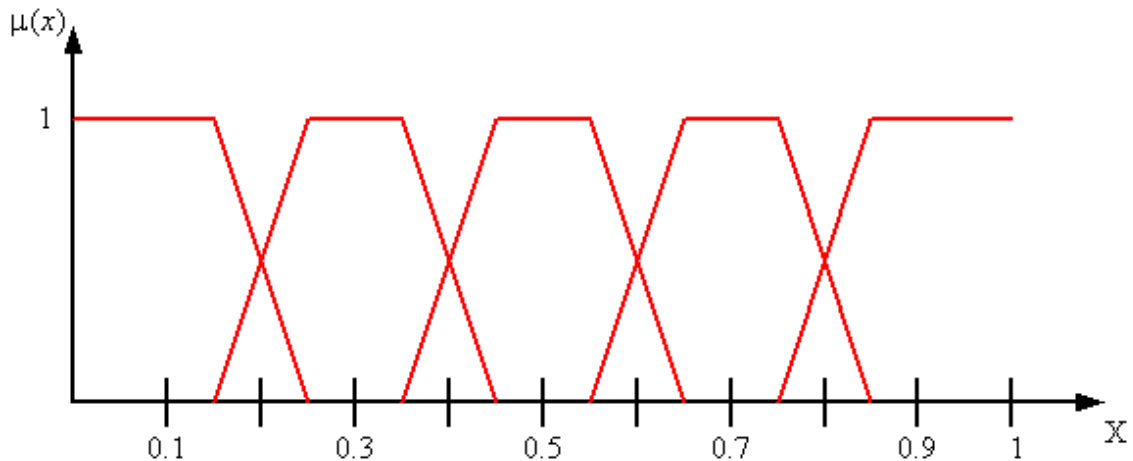


Рис. 3. Графік функцій приналежності терм-множини стандартного нечіткого п’ятирівневого 01-класифікатора

Функції приналежності термів B_i , $i=1,\dots,5$ через їхню трапецеїдальну форму можна описати кортежем із чотирьох значень, що визначають початкову й кінцеву абсцису нульового й одиничного α -рівнів. Враховуючи це, терм-множину стандартного нечіткого п’ятирівневого 01-класифікатора можна визначити так:

$$\begin{aligned} \mu_{ОН} &= (0, 0, 0.15, 0.25); \\ \mu_{Н} &= (0.15, 0.25, 0.35, 0.45); \\ \mu_{З} &= (0.35, 0.45, 0.55, 0.65); \\ \mu_{В} &= (0.55, 0.65, 0.75, 0.85); \\ \mu_{ОВ} &= (0.75, 0.85, 1, 1). \end{aligned} \tag{6}$$

Викладений спосіб можна застосувати і для загальнішого випадку – побудови n -рівневого нечіткого класифікатора, за умови, що $n > 1$ і $n \in \mathbb{Z}$, де \mathbb{Z} – множина цілих чисел. Кількість рівнів n вибирають залежно від вимог, висунутих до моделі. Дворівневий класифікатор не представляє інтересу, тому що не містить серединного нечіткого значення, навколо якого найчастіше групуються стани реальних об’єктів.

Тому доцільно розглянути класифікатор при $n=3$. Стандартний трирівневий нечіткий 01-класифікатор введемо через лінгвістичну змінну $B^{(3)}$ з терм-множиною значень B_1 “Низький (Н), B_2 Середній (З), B_3 Високий (В)”. Функції приналежності відповідних термів описуються співвідношеннями:

$$\begin{aligned}\mu_H &= (0,0,0.2,0.4); \\ \mu_C &= (0.2,0.4,0.6,0.8); \\ \mu_B &= (0.6,0.8,1,1);\end{aligned}\tag{7}$$

Графік функцій приналежності, описаних співвідношеннями (7), наведено на рис. 4. Множина вузлових точок набуде значення $\alpha = \{0.1,0.5,0.9\}$.

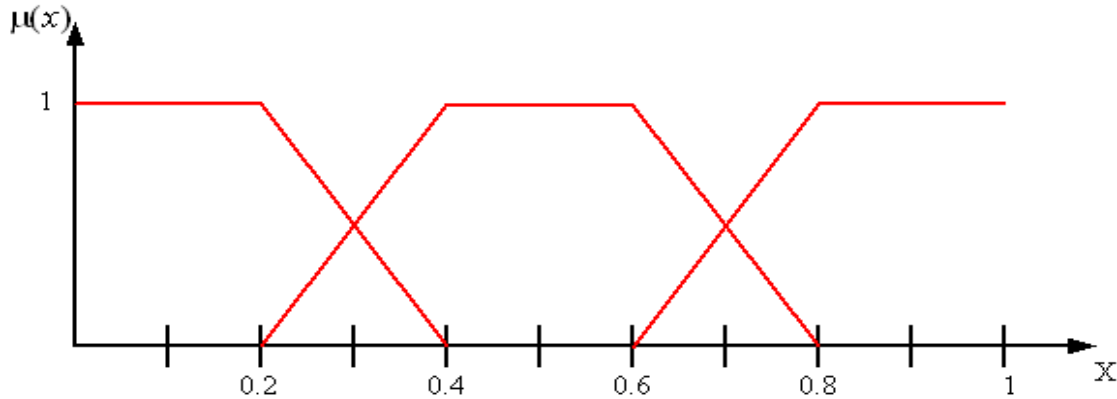


Рис. 4. Графік функцій приналежності терм-множини стандартного нечіткого трирівневого 01-класифікатора

Зазначимо, що область значень деяких заданих у табл. 1, 2 факторів за допомогою лінійних перетворень можна привести до носія нечіткого класифікатора, а терми, що описують нечіткі значення до терм стандартного нечіткого трирівневого 01-класифікатора.

Розглянемо наступний рівень ієрархії розглянутої моделі, для якої визначено цільову функцію (4). Введемо в розгляд вектора

$$\begin{aligned}x_1 &= \{x_{1i}\}_{i=1}^{N_1} = [I_N^-, I_V^-, E^+, T^{+-}, f^+, \left(\frac{\Delta P}{\Delta t}\right)^-], N_1 = 6, \\ x_2 &= \{x_{2i}\}_{i=1}^{N_2} = [I_0^+, F^-, T_p^-, TS^-, A^+], N_2 = 5, \\ r &= [r_1, r_2] = [S_c, S_p].\end{aligned}\tag{8}$$

Тоді у векторному вигляді залежності (2) і (3) можна переписати у вигляді

$$r_i = f_i(x_i), \quad i = 1, 2.\tag{9}$$

Для кожного з окремих факторів $x_{mi}, m = \overline{1, 2}, i = \overline{1, N_m}$ побудовані свої трирівневі класифікатори, що не залежать від виду параметра. Тоді для опису залежностей (9) можна скористатися формулою

$$\begin{aligned}r_m &= \sum_{i=1}^{N_m} z_{m,i} v_{m,i} \sum_{k=1}^3 a_k^m m_{ki}^m(x_{mi}) \\ \sum_{i=1}^{N_m} v_{m,i} &= 1.\end{aligned}\tag{10}$$

де α_k – вузлові точки стандартного класифікатора; $v_{m,i}$ – вага m-го фактора для i-го параметра; $m_{ki}^m(x_{mi})$ – значення функції приналежності k-го терма i-го фактора, m – і залежності; $z_{m,i} \in \{1, -1\}$ – коефіцієнт, що визначає позитивний або негативний напрямок впливу i-го параметра.

Для оцінювання залежності (10) необхідно, своєю чергою, побудувати нечітку базу знань, яка дасть змогу оцінити значення функції приналежності k-го терма i-го фактора.

Висновки

Параметри навколишнього середовища і стану системи можуть описувати деяку нечітку ситуацію або не становити інтересу з погляду опису нечітких ситуацій. Тому за набором фактів, зібраних на деякому інтервалі спостереження, можна визначити нечіткі ситуації, що найчастіше зустрічаються, і ввести набір еталонних ситуацій.

Нечіткі класифікатори мають велике значення, оскільки дають можливість описати з максимальною точністю фактори, для яких немає відомих точних закономірностей і для яких необхідно провести асоціацію між якісною й кількісною оцінками факторів. Якщо при побудові класифікатора є деяка додаткова інформація про поведінку фактора, то в загальному випадку класифікатор не буде стандартним, тому що вузлові точки не будуть симетричні щодо середини носія відповідного фактора.

1. Борисов В. В. *Нечеткие модели и сети* / В. В. Борисов, В. В. Круглов, В. В. Федюлом. – М. : Горячая линия – Телеком, 2007. – С. 284. 2. Рыжов А. П. *Элементы теории нечетких множеств и ее приложений* / А. П. Рыжов. – М., 2003. – С. 81. 3. Недосекин А. О. *Методологические основы моделирования финансовой деятельности с использованием нечетко-множественных описаний* : дис. ... д-ра экон. наук // СПб., Санкт-Петербургский государственный университет экономики и финансов, 2003. – С. 280. 4. Шеридан Т. Б. *Системы человек-машина: Модели обработки информации, управления и принятия решений человеком-оператором* / Т. Б. Шеридан, У. Р. Феррел; пер. с англ.; под ред. К. В. Фролова. – М.: Машиностроение, 1980. – 400 с. 5. *Человеческий фактор. В 6-ти т. Т. 2. Эргономические основы проектирования производственной среды: Пер. с англ. / Д. Джоунз, Д. Бродбент, Д. Е. Вассерман и др.* – М.: Мир, 1991. – 500 с. 6. *Передерий В.И., Еременко А.П. Математические модели и алгоритмы определения релевантности принимаемых решений с учетом психофункциональных характеристик пользователей при управлении автоматизированными системами. [Текст] / В.И. Передерий, А.П. Еременко // Автоматика. Автоматизация. Электротехничні комплекси та системи.* – 2007, №2. – С. 34–40. 7. Кофман А. *Введение в теорию нечетких множеств* / А. Кофман. – М.: Радио и связь, 1982. – 432 с. 8. Ricardo J.G.V. Campello, Wagner Caradori do Amaral. *Hierarchical Fuzzy Relational Models: Linguistic Interpretation and Universal Approximation (2006), IEEE Transactions on Fuzzy Systems, vol. 14, No. 3, pp. 446–453, doi: 10.1109/TFUZZ.2006.876365.* государственный технический университет. 30.06.2006. – Обнинск: 2006. – 159 с. 9. *Справочник по инженерной психологии / Под ред. Б. Ф. Ломова.* – М.: Машиностроение, 1982. – 368 с. 10. Koji Shimojima, Toshio Fukuda, Yasuhisa Hasegawa. *Self-tuning fuzzy modeling with adaptive membership function, rules, and hierarchical structure based on genetic algorithm (1995), Fuzzy Sets and Systems, vol. 71, pp. 295-309.*

SYNTHESIS OF RBF-NETWORK FOR PREDICTION OF SECONDARY PROTEIN STRUCTURE

© Lytvynenko V., 2014

Запропоновано методологію синтезу радіально-базисних мереж для вирішення проблеми білкового середнього передбачення структури за допомогою алгоритму вибору клонів. Щоб вирішити цю проблему було використано метод “один проти всіх”. Обчислювальні експерименти щодо випробуваного зразка показали, що точність прогнозування сягає до 72 %, що вказує на високу точність запропонованого способу.

Ключові слова: алгоритм клонального відбору, радіальна базисна функція, метод “один проти всіх”, прогнозування, вторинна структура білка.

In this paper we propose the methodology of team radial-basis networks synthesis for solving the problem of protein secondary structure prediction using clonal selection algorithm. To solve such problem the method of “one against all” have been used. The carried out computational experiments on test sample have shown that the prediction accuracy allows to achieve up to 72 %, indicating a high accuracy of the proposed method.

Key words: clonal selection algorithm, radial basis function, “one against all” method, predicting, the secondary structure of protein.

1. Introduction

Proteins are large biological molecules with complex structures and constitute to the bulk of living organisms: enzymes, hormones and structural material [1]. The function of a protein molecule in a given environment is determined by its 3-dimensional (3-D) structure [1]. Protein 3-D structure prediction directly from amino acid sequences still remains as an open and important problem in life sciences. The bioinformatics approach first predicts the protein secondary structure (PSS) which represents an 1-D projection of the very complicated 3-D structure of a protein [3]. Secondary structures are regular structural elements which are formed by hydrogen bonds between relatively small segments of the protein sequence. Often the driving force for the formation of a secondary structure is the saturation of backbone hydrogen donors (NH) and acceptors (CO) with intramolecular hydrogen bonds. This saturation allows the protein to bury hydrophobic side chains in its interior (hydrophobic core) without conflicting with the polar backbone. There are three common secondary structures in proteins, namely *a*-helix, *b*-strand, and coil [12]. An *a*-helix is formed from a connected stretch of amino acids. The *a*-helix is characterized by hydrogen bonds along the chain, which are almost coaxial. The *a*-helix is the most abundant helical conformation found in globular proteins. The average length of an *a*-helix is around 10 residues. A *b*-strand is the principal component of a *b*-sheet. The *b*-sheet is characterized by hydrogen bonds crossing between chains. Each participating *b*-strand in a *b*-sheet is not continuous in terms of the primary sequence and does not even have to be close to another *b*-strand in the sequence. A *b*-strand has a sequence of 5-10 residues in a very extended conformation. Approximately one-third of all residues in globular proteins are contained in coils. The coils in a protein serve to reverse the direction of the polypeptide chain. Coils vary in length.

The goal of secondary structure prediction is to classify a pattern of residues in amino acid sequences to a corresponding secondary structure element: an α -helix (H), β -strand (E) or coil (C, the remaining type). Many computational techniques have been proposed in the literature to solve the PSS prediction problem, which can be broadly fallen into three categories: (a) statistical methods, (b) neural network approaches, and (c) nearest neighbor methods. The statistical methods are mostly based on likelihood techniques [4, 5, 6]. Neural network approaches use residues in a local neighborhood or window to predict the secondary structure at a particular location of an amino acid sequence [7, 8]. The nearest neighbor method often uses the k -nearest neighbor techniques [9, 10]. SVMs have been earlier applied to PSS prediction [11]. One of the drawbacks in this approach is that the method does not capture the global information of the amino acid sequence due to the limited size of the local neighborhood. Additionally, the method only constructs a multi-class classifier by combining several binary classifiers.

Despite the existence of many approaches, the current success rates of existing approaches are insufficient; further improvement of the accuracy is necessary. Most existing secondary structure techniques are single-stage approaches, except the PHD [8] and PSIPRED [9] methods which combined two multi-layer perceptron (MLP) networks. Single-stage approaches are unable to find complex relations (correlations) among different elements in the sequence. This could be improved by incorporating the interactions or contextual information among the elements of the output sequence of secondary structures. We argue that it is feasible to enhance present single-stage approaches by augmenting with another prediction scheme at their outputs and propose to use SVMs as the second-stage.

This paper investigates the use of multi-class RBF neural networks which we synthesize using clonal selection algorithm for PSS prediction. We present new multi-class techniques based on binary classifiers to PSS prediction.

2. Data and methods

2.1. Problem definition

One main sub problem of the domain is ‘*Protein Secondary Structure Prediction*’. The primary sequence of a protein can be represented as [13] $\{A, R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y, V\}^n$, where the letters are the one letter codes of the amino acid residues (total 20 possible amino acids) and n is the length of the protein to be predicted. The secondary structure of the sequence having length n is $\{a, b, c\}^n$, where a, b, c are different secondary structure classes. So, the problem of secondary structure prediction can be represented as a mapping problem as follows [13]:

$$\{A, R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y, V\}^n \rightarrow \{a, b, c\}^n$$

2.2. Orthogonal encoding of amino acid

Orthogonal encoding of amino acid types has been used in many bioinformatic neural network models: 20 input units are assigned to describe one protein residue. In the 20-dimensional space, the vector $[1, 0, 0, 0, \dots, 0, 0, 0]$ represents alanine, and $[0, 0, 0, \dots, 0, 0, 0, 1]$ stands for valine. With this encoding, a typical input window of 13 residues requires 260 (13×20) input units. It can easily lead to large input layers, many connecting weights, and hence complex models. Without sufficient data to support training, over-complex models are prone to overfitting. Unfortunately, in many bioinformatic problems, huge data sets can be simply unavailable. Even when they are available, analysing them is often very computationally demanding. Simplified encoding schemes use less input units to describe a given amino acid sequence; thus, we can use smaller models to describe the same phenomena. By introducing these simplified models, we can reduce the reliance on huge data sets and improve performance. To increase the level of neural network generalization, in work [14] defined a 10-unit input scheme for representation of amino acid type. Each amino acid was described using ten numbers. In work [15] their representation was based on the amino acid features described by: each unit corresponds to one biochemical feature; amino acids sharing many features have similar codes. In work [16] suggested two differing properties, “sequence-derived hydrophobicity” and “sequence-derived polarity”, based on correlations in protein sequences. In work

[17] applied an adaptive encoding neural network to find automatically a classifier with a low-dimensional encoding matrix. Their encoding scheme was tested on the prediction of cleavage sites in human signal peptides of secretory proteins.

2.3. Data Set

The set 126 nonhomologous globular protein chains used in the experiment of Rost and Sander [18], referred to as the RS126 set, was used to evaluate the accuracy of the classifiers. The dataset contained 23349 residues with 32% α -helix, 23% β -strand, and 45% coil. Many current generation secondary structure prediction methods have been developed and tested on this dataset.

The RS126 set is available at <http://www.compbio.dundee.ac.uk/~www-jpred/data/>. The single stage approaches and second-stage approaches were implemented, with multiple sequence alignments, and tested on the dataset, using a sevenfold cross validation technique to estimate the prediction accuracy. With sevenfold cross validation approximately one-seventh of the database was left out while training and, after training, the left one-seventh of the dataset was used for testing. In order to avoid the selection of extremely biased partitions, the RS126 set was divided into seven subsets with each subset having similar size and content of each type of secondary structure.

2.4. Synthesis of radial-basis network

As the classifier, in general, is called a function that for object attributes vector makes the decision: to which of classes it belongs:

$$F : \mathcal{R}^n \rightarrow Y. \quad (1)$$

The function F reflects the vector characteristics space in the space of class labels Y . In the case of two classes $Y = \{0,1\}$, '1' corresponds to case of event you are looking for; '0' - an event not found. We consider the option of training with a teacher (*supervised learning*), when for the classifier training available a set of vectors $\{x\}$ for which is known their true identity to one of the classes.

In binary classification the class identifiers can be interpreted as states of the system (active or passive, normal or abnormal), which are presented by number of properties.

The vector of properties determines the system state. Each state of $x^i = (x_1^i, \mathbf{K}, x_n^i) \in [0,1]^n$ is represented by set $U \subseteq [0,1]^n$. The properties vector elements can be scaled or normalized in the interval $[0,1]$.

The properties vector set of $Positiv \subseteq U$ represents the normal state of system. Its complement is called $Negativ$ and determined as $Negativ = U - Positiv$. In some cases, we will define a set $Positiv$ (or $Negativ$), using its characteristic function [26] $c_{Positiv} : [0,1]^n \rightarrow \{0,1\}$:

$$c_{Positiv}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in Positiv \\ 0 & \text{if } \mathbf{x} \in Negativ \end{cases}. \quad (2)$$

For a given set of positive examples $Positiv' \subseteq Positiv$ we have to evaluate the characteristic function ($c_{Positiv}$) of normal space, which should have the ability to solve whether the observed state of positive or negative.

The entire set of neural networks Δ be divided into subset A caused by the chosen topology [27] of RBF-neural network (number of RBF-neuron network is Γ). Within each class $RBFN_i \subset RBFN$ the neural networks are characterized by an additional set of parameters: the number of inputs n ; the set of synaptic weights of the output layer $W = \{w^i, i=1, \mathbf{K}, p\}$; the number of RBF-neurons of network $\Gamma = \{g^i, i=1, \mathbf{K}, p\}$; the centers of RBF $C = \{c^i, i=1, \mathbf{K}, p\}$; the parameters of RBF $\Sigma = \{s^i, i=1, \mathbf{K}, p\}$; the parameter of the output neuron activation function of i -th network $A = \{a^i, i=1, \mathbf{K}, p\}$; in the case of RBF-neural network scale $S = \{s^i, i=1, \mathbf{K}, p\}$ of RBF shifts to time axis $T = \{t^i, i=1, \mathbf{K}, p\}$. Thus, parameters vector RBF-neural networks $q = \{\Gamma, W, C, \Sigma, A\}$ is formed. The natural criterion for selecting RBF-neuron network will function defined by the standard RMS for any input.

Thus, the task of RBF-neural network synthesis can be reduced to optimization problems such as:

$$F^* = F(\mathbf{q}^*) = \min F(\mathbf{q}), \quad (3)$$

$$a_1 \leq x_1 \leq b_1, \dots, a_n \leq x_n \leq b_n,$$

where the function F is not imposed any restrictions, such as differentiation, Lipschitz condition, continuity, etc.

The problem solution of multiparametric function optimization of the form (3) could be used the appropriate operators of clonal selection.

RBF-network consists of input, single hidden (radial-basis) and linear (output) layer. Input layer consists of sensors that connect the network with the external environment. The hidden layer neurons act by the principle of centering on the elements of training sample. As the center stands the weight matrix (W^r). Box ($dist$) the Euclidean distance between input vector (X) and the corresponding center is calculating. Around each center there is a region, called the radius. Radius (sensitivity of the network) is adjusted by means of the smoothing coefficients vector (S_1, \dots, S_m).

The conversion function (usually Gauss - $f(x) = e^{-\frac{(x-c)^2}{2s^2}}$), which varies from 0 to 1, determines the output of hidden layer. Output layer contains the usual linear or sigmoidal neurons and by adjustment of their weights (W^l) determines the output network.

The behavior of RBF-network depends largely on the number and position of radial basis function of hidden layer. Indeed, for any real n -dimensional input vector $x = (x_1, x_2, \dots, x_n)$, where $x \in X \subset \mathfrak{R}_n$, the network output will be determined as follows:

$$y_i = \sum_{k=1}^m w_{ik}^l f_k(dist(x, w_k^r), s_k), \quad (4)$$

where $w_{ik}^l \in W^l$ - is the weight of linear layer; $w_k^r \in W^r$ - centers of radial-basis functions. If as basic functions used Gauss-function one, then

$$f_k(x) = -\frac{dist(x, w_k^r)^2}{2s_k^2}, \quad k = \overline{1, m}. \quad (5)$$

Architecture of RBF-neural network used by us for solving the classification tasks presented in Fig. 1. The hidden layer neurons are the RBF-neurons. As the RBF parameters used its scale (s) and shift (t) in terms of the time axis.

Output layer contains the usual linear or sigmoidal neurons and by adjustment of their weights (W^l) determines the output network.

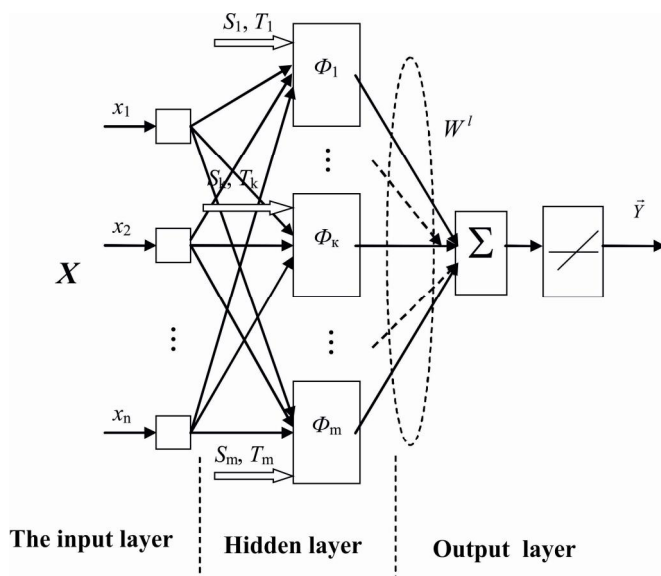


Fig. 1. Generalized RBF-neural network architecture for solving classification problems [19]

Based on the architecture, and at $p = 1$ (for time series), the output of the neural network will be determined as follows:

$$y_i = \sum_{k=1}^m w_k^l \Phi_k(x, S_k, T_k), \quad (6)$$

where $x = (x_1, x_2, \dots, x_n)$, $x \in X \subset \mathfrak{R}_n$ - an arbitrary n -dimensional vector of input variables; $w_k^l \in W^l$ are the weights of linear layer; S_k - the scale parameters for the hidden layer neurons; T_k is the displacement parameters for the hidden layer neurons.

Based on the architecture of neural network (Fig. 2), as parameters are customizable, one can distinguish the following [19]:



Fig. 2. Encoding configurable parameters RBF neural network in the form of antibodies (chromosomes)

where: n – number of inputs; m – number of neurons; f - ON / OFF neuron (0 or 1); c – RBF centers;

w – synaptic weights of output layer; S – RBF parameter; a – parameter of function activate the output neuron.

In general, the procedure of the synthesis of each neural network for each class, are performed in accordance with the procedure presented in Fig. 3.

In this chapter the synthesis RBF-neural networks, that aimed at solving classification tasks. The classifier, in general, is a function that on attributes vector base of the object makes a decision, which is a class it belongs, respectively (1). In the binary classification the class identifiers can be interpreted as states of the system (active or passive, normal or abnormal), which presented by the number of properties. Types of basis functions, their number, type and parameters of activation function in the linear layer, are setting up as the parameters of AIS.

Study and synthesis of neural networks groups is carried out by the scheme shown in Fig. 3

As an objective function and the function affinity selected the RMS error of network on the training data. The training procedure has the following characteristics [19, 20]:

- selection is implemented on the tournament selection, which allows to control convergence and to maintain the diversity of population at the required level;
- because of the binary encoding specific, scheme of mutation has been proposed, whereby the probability of single bit line changing depends not only on antibody affinity in general, but also on the significance of this bit.

In the case of RBF-neural networks, the network behavior is largely dependent on the number and position of radial basis function of hidden layer. Indeed, for any real n -dimensional input vector $x = (x_1, x_2, \dots, x_n)$, where $x \in X \subset \mathfrak{R}_n$, the network output will be determined as follows:

$$y_i = \sum_{k=1}^m w_{ik}^l f_k \left(\text{dist} \left(x, w_k^r \right), S_k \right), \quad (7)$$

where $w_{ik}^l \in W^l$, $i = \overline{1, p}$ are the linear layer weights; $w_k^r \in W^r$ are the center of radial-basis functions.

If as basic function is used the Gauss function, then

$$f_k(x) = -\frac{\text{dist}(x, w_k^r)^2}{2S_k^2}, \quad k = \overline{1, m}. \quad (8)$$

In the context of the classification problem debugging the network is to find functions $y : \mathfrak{R}_n \rightarrow \mathfrak{R}$ that satisfy the equation (7.5) at $p = 1$. Let we have the sample is composed of S training data points X_1, \dots, X_S , $X_i \in \mathfrak{R}_n$. If you know the output value for each of these points d_1, \dots, d_S , $d_i \in \mathfrak{R}$,

then each basis functions can be centered on one of the points X_i . Thus, in the limiting case the number of centers and, accordingly, the hidden layer neurons is equal to the number of data points in training sample $m = S$.

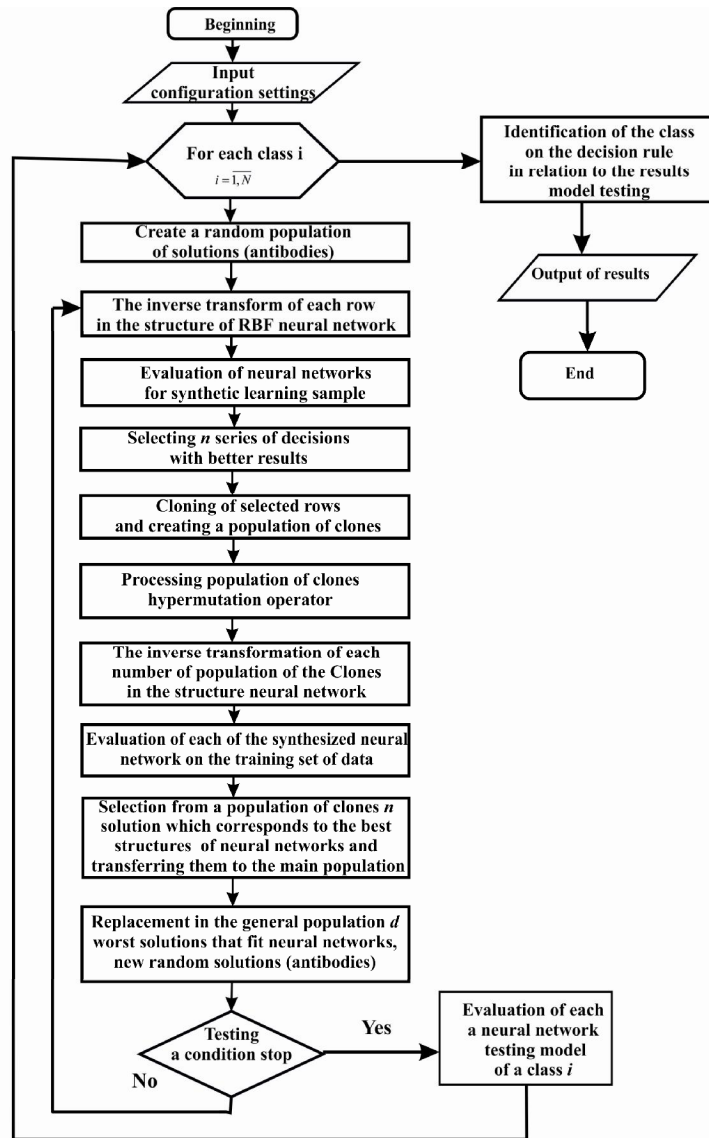


Fig. 3. The procedure of the RBF-neural network synthesis using the clonal selection

Synthesis of collective neural networks, where each neural network recognizes only a single class, is similar to the procedure described for RBF-neural networks according to the adjustment parameters presented in Fig. 4.

Many discriminative methods, including Support vector machine, neural network and classifiers based on the artificial immune systems, are often most accurate and efficient when dealing with two classes only (they can deal with more classes, but usually at reduced accuracy and efficiency). For large number of classes, higher-level multi-class methods are developed that utilize these two-class classification methods as the basic building blocks.

To solve the problem we used the strategy of one-versus-all based on multiparameter optimization function of the form (3) use the corresponding operators clonal selection algorithm [21]. Types of basis functions and the activation function of the linear layer are defined as parameters to clonal selection algorithm. Learning and synthesis collectives of neural networks is performed according to the scheme shown in Fig. 4.

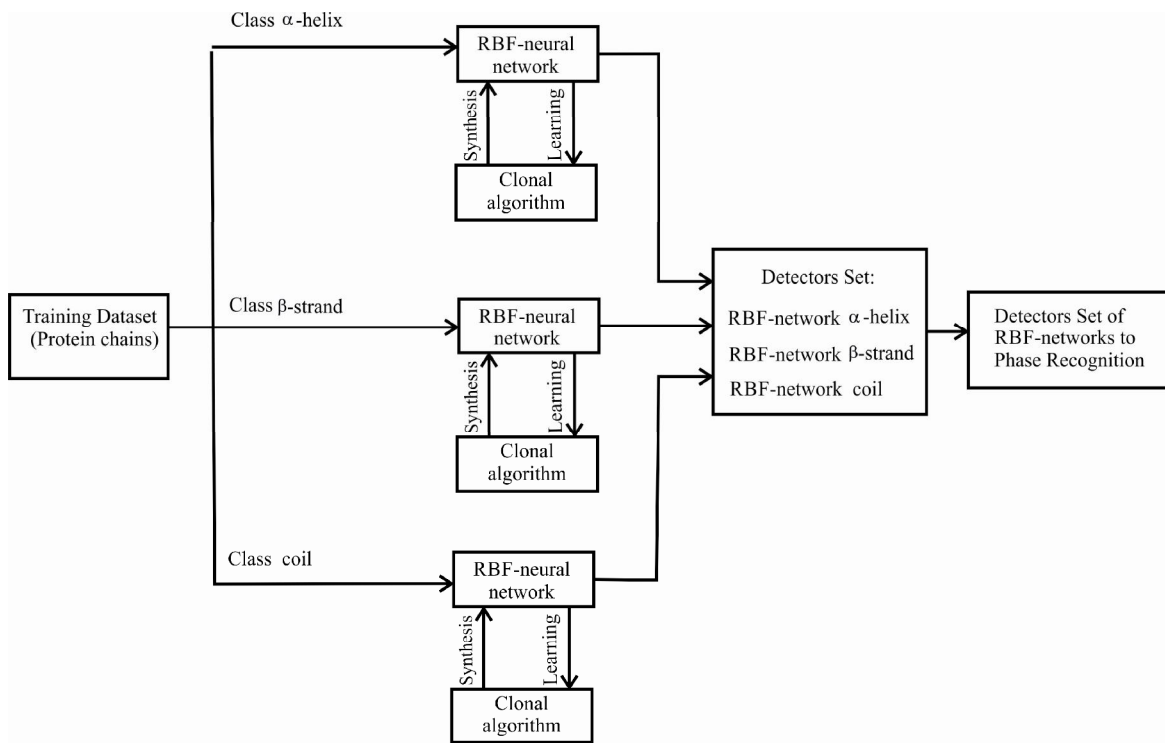


Fig. 4. Synthesis of binary classifiers collective RBF-neural networks

The simplest approach is to reduce the problem of classifying among K classes into K binary problems, where each problem discriminates a given class from the other $K - 1$ classes [22]. For this approach, we require $N = K$ binary classifiers, where the k th classifier is trained with positive examples belonging to class k and negative examples belonging to the other $K - 1$ classes. When testing an unknown example, the classifier producing the maximum output is considered the winner, and this class label is assigned to that example. In work [22] state that this approach, although simple, provides performance that is comparable to other more complicated approaches when the binary classifier is tuned well [23].

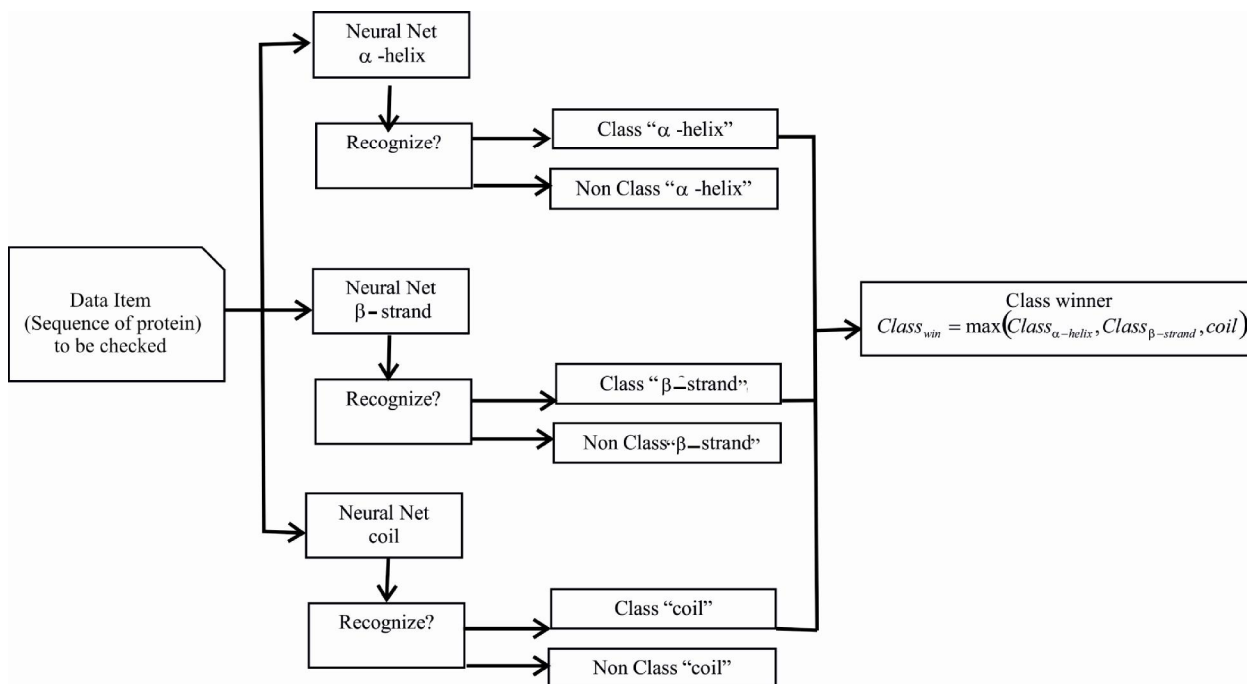


Fig. 5. Testing of synthesized neural networks

The identification of the system state performed by the expression (9):

$$Class_{win} = \max(Class_{a=helix}, Class_{b-s\ tan\ d}, Class_{coil}) \quad (9)$$

2.5. Measurements of accuracy

The most commonly reported measure of secondary structure prediction accuracy is the success rate, or Q_3 . This is the overall percentage of correctly predicted residues of all 3 types [24], i.e.,

$$Q_3(\%) = \frac{R_{helix} + R_{sheet} + R_{coil}}{N} \times 100 \quad (10)$$

Here, R_i , is the number of correctly predicted residues of type i , and N is the total number of residues.

Although the Q_3 score provides a quick measure of the accuracy of the algorithm, it does not account for differing success rates on different types of secondary structure. We therefore also calculated the correlation coefficients [25] for prediction of helix C_H , sheet C_E , and coil C_C .

$$C_a = \frac{(p_a \cdot n_a) - (u_a \cdot o_a)}{\sqrt{(n_a + u_a)(n_a + o_a)(p_a + u_a)(p_a + o_a)}} \quad (11)$$

where p_a is the number of positive cases that were correctly predicted, n_a is the number of negative cases that were correctly rejected, o_a is the number of overpredicted cases (false positives), and u_a is the number of underpredicted cases (misses). Similar expressions hold for C_b , and C_{coil} . The Q_3 measure will be used to assay the overall success rate of network models during learning, although it is not as good an indicator as the individual correlation coefficients.

3. Results and discussion

We studied the dependence of testing success rate on the size of the input window using a standard network with 15 hidden units (RBF-functions). The results shown in Table 1 indicate that when the size of the window was small the performance on the testing set was reduced, probably because information outside the window is not available for the prediction of the secondary structure. When the size of the window was increased, the performance reached a maximum at around 20 groups (6 on either size of the center residue).

Table 1

Dependence of testing success rate on window size

Window size	$Q_3(\%)$	C_H	C_E	C_C
2	51	0.49	0.50	0.48
4	55	0.50	0.52	0.54
6	54	0.51	0.53	0.52
8	56	0.52	0.52	0.54
10	56	0.54	0.53	0.53
12	58	0.54	0.55	0.57
14	60	0.57	0.58	0.57
16	64	0.61	0.66	0.59
18	65	0.61	0.65	0.65
20	73	0.69	0.68	0.68
22	69	0.67	0.65	0.68
44	50	0.51	0.46	0.49
54	48	0.41	0.55	0.45
64	45	0.49	0.38	0.48
74	39	0.44	0.35	0.40

We used 12 RBF-neurons in hidden layer, window size 20, and different number of epochs. Results are in Table 2

Table 2.

Design of neural networks for different number of epoch

	Number of training epochs	Test	
		$Q_3\%$ (training set)	$Q_3\%$ (testing set)
1	100	45.5	43.4
2	200	50.1	48.2
3	300	53.2	49.9
4	400	59.6	55.8
5	500	60.1	57.2
6	600	63.0	60.8
7	700	64.9	61.6
8	800	65.6	61.2
9	900	67.1	63.1
10	1000	68.2	64.1
11	1100	68.4	64.6
12	1200	72.1	70.0
13	1300	74.5	72.2
14	1400	75.8	73.2
15	1500	75.7	72.1

4. Conclusion

The paper shows the results of research carried out by the authors of the combined classification algorithm based on group RBF-networks for solving the problem of classification of mass spectra. Analysis of the problem solutions demonstrates the effectiveness of this algorithm that uses parallel-distributed organization of calculations. Feasibility of using it explains their high flexibility, the ability to search for parallel, resistant to noise, associative memory, self-organizing, structural flexibility and high adaptive capacity.

1. Nguyen M.N., Rajapakse J.C. *Multi-Class Support Vector Machines for Protein Secondary Structure Prediction*, *Genome Informatics* 14, 218-227, 2003. 2. Clote, P. and Backofen, R., *Computational Molecular Biology*, Wiley and Sons, Ltd., Chichester, 2000. 3. Mount, D.W., *Bioinformatics: Sequence and Genome Analysis*, Cold Spring Harbor Laboratory Press, 2001. 4. Garnier, J., Osguthorpe, D.J., and Robson, B., *Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins*, *Journal of Molecular Biology*, 120:97-120, 1978. 5. Garnier, J., Gibrat, J.F., and Robson, B., *GOR method for predicting protein secondary structure from amino acid sequence*, *Methods Enzymol*, 266:541-553, 1996. 6. Gibrat, J.F., Garnier, J., and Robson, B., *Further developments of protein secondary structure prediction using information theory*, *Journal of Molecular Biology*, 198:425-443, 1987. 7. Jones, D.T., *Protein secondary structure prediction based on position-specific scoring matrices*, *Journal of Molecular Biology*, 292:195-202, 1999. 8. Rost, B. and Sander, C., *Prediction of protein secondary structure at better than 70% accuracy*, *Journal of Molecular Biology*, 232:584-599, 1993. 9. Salamov, A.A. and Solovyev, V.V., *Prediction of protein secondary structure by combining nearest-neighbor algorithms and multiple sequence alignments*, *Journal of Molecular Biology*, 247:11-15, 1995. 10. Salamov, A.A. and Solovyev, V.V., *Protein secondary structure prediction using local alignments*, *Journal of Molecular Biology*, 268:31-36, 1997. 11. Hua, S. and Sun, Z., *A novel method of protein secondary structure prediction with high segment overlap measure: support vector machine approach*, *Journal of Molecular Biology*, 308:397-407, 2001. 12. Lipo Wang and Xiuju Fu *Data Mining With Computational Intelligence Berlin: Springer-Verlag, 2005, pp. 276, (ISBN 3-540-24522-7)*. 13. Sudipta Saha *Protein Secondary Structure Prediction by Fuzzy Min-Max Neural Network with*

Compensatory Neuron / Thesis submitted in partial fulfillment of the requirements for the degree of Master of Technology In Computer Science & Engineering/ Department of Computer Science & Engineering Indian Institute of Technology, Kharagpur-721302, West Bengal, India May, 2008, 126 p.

14. Skolnick, J., A. Kolinski, and A. R. Ortiz. 1997. MONSSTER: a method for folding globular proteins with a small number of distance restraints. *J. Mol. Biol.* 265:217–241.

15. Taylor, W. R. & Thornton. J. M. (1984). Recognition of super-secondary structure in proteins. *J. Mol. Biol.* 173. 487-514.

16. O. Weiss and H. Herzog. Measuring Correlations in Protein Sequences. *Z. Phys. Chem.*, 204, 183-197 (1998).

17. Jagla, B. and Schuchhardt, J. 2000. Adaptive encoding neural networks for the recognition of human signal peptide cleavage sites. *Bioinformatics* 16: 245–250.

18. Rost, B. and Sander, C., Prediction of protein secondary structure at better than 70% accuracy, *Journal of Molecular Biology*, 232:584-599, 1993.

9. Литвиненко В.І., Фефелов А.О., Дідик О.О. Методологія синтезу колективу радіально-базисних мереж для розв'язування задач класифікації за допомогою алгоритму клонального відбору // Наукові праці ЧДУ ім. Петра Могили. Науково-методичний журнал. Серія "Комп'ютерні науки", – ЧДУ ім. Петра Могили. – 2009. – Вип. 93. – Том.106. – С.111–123.

20. Литвиненко В.І. Искусственные иммунные системы как средство индуктивного построения оптимальных моделей сложных объектов// Проблемы управления и информатики. – 2008. – №3. – С. 30–42.

21. Chris H.Q.Ding and Inna Dubchak Multi-class protein fold recognition using support vector machines and neural networks / *Bioinformatics*, Vol.17 no. 4, 2004 p. 349-358.

22. Ryan Rifkin and Aldebaro Klautau. Parallel networks that learn to pronounce english text. *Journal of Machine Learning Research*, pages 101–141, 2004.

23. Mohamed Aly. Survey on Multi-Class Classification Methods. Technical Report, Caltech, USA, 2005.

24. N. Qian T. J. Sejnowski Predicting the Secondary Structure of Globular Proteins Using Neural Network Models/ *J. Mol. Biol.* (1988) 202, p. 865-884.

25. Matthews, B.W., Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta* 1975, 405, 442–451.

26. Fabio Gonzalez. A Study of Artificial Immune Systems Applied to Anomaly Detection [Ph.D. thesis]. USA The University of Memphis; 2003. -184 p.

27. Тюмерев В.В. Метод эволюционного накопления признаков для автоматического построения нейронных сетей // Вычислительные методы и программирование.- 2001.- Т.2.- С.88-108.

КОМП'ЮТЕРНА СИСТЕМА АНАЛІЗУ АЕРОІОННОГО РОЗПОДІЛЕННЯ НА ПОХИЛІЙ ПЛОЩИНІ

© Строкань О., 2014

На основі теоретичних і практичних результатів дослідження поширення концентрації від'ємних аероіонів від штучних джерел аероіонного випромінювання у виробничих і побутових приміщеннях запропоновано комп'ютерну систему аналізу аероіонного поширення на похилій площині.

Ключові слова: від'ємні аероіони, аероіонізатор, кут нахилу, похила площина, комп'ютерна система, середовище математичного процесора Maple.

On the basis of theoretical and practical research results of spreading of concentration of negative ions from artificial sources of air ionic radiation in industrial and domestic premises the proposed computer system of analysis of aeroion distribution on an inclined plane.

Key words: negative aeroions, airionizer, angle of incline, inclined plane, a computer system, the environment-above mathematical processor Maple.

Вступ

За умов стрімкої урбанізації більшість мешканців міст проводять 90 % свого часу у закритих приміщеннях: вдома і на роботі, вдихаючи повітря з високим рівнем забруднення, з недостатньою кількістю кисню. Зокрема пил є колосальним джерелом інфекцій, єдиним засобом переміщення бактерій, вірусів і пилових кліщів, які приносять у приміщення алергію і багаточисленні хвороби. Через це виникає проблема очищення повітря, тобто створення комфортних умов, у виробничих і побутових приміщеннях. Одним із засобів наближення стану повітря у таких приміщеннях до оптимального є насичення його від'ємними іонами, тобто іонізування повітря [2, 3]. Крім очищувальної дії від'ємні аероіони мають лікувальну і профілактичну дію: активізують імуну систему, нормалізують дихальний обмін і величину артеріального тиску, мають антистресову, антимікробну і антивірусну дії [3].

Для насичення повітря приміщень від'ємними іонами застосовують так звані аероіонізатори. Поширення концентрації від'ємних аероіонів від аероіонізаторів визначається за допомогою спеціальних карт із нанесеними на них лініями однакової концентрації аероіонів – ізолініями. Відомо, що зі зміною кута площини, для якої ведуться розрахунки, змінюються і координати точок, які інцидентні заданим ізолініям концентрації аероіонів [4]. З метою прогнозування і візуалізації процесу розподілу концентрації від'ємних аероіонів при зміні положення розрахункової площини необхідно розробити комп'ютерну систему аналізу цього розподілу.

Аналіз останніх досліджень і публікацій

Починаючи з другої половини двадцятого століття поширений напрям науки і техніки, основою якого є використання силової дії електричних полів на електрзаряджені частки – так звана електронно-іонна технологія. Результатами численних досліджень того часу є закономірності розподілення від'ємних аероіонів на горизонтальній площині від одного і більше аероіонізаторів [3]. Сьогодні збільшився інтерес до дослідження розподілу концентрації аероіонів від штучних джерел аероіонного випромінювання. Зокрема постають питання моделювання і прогнозування цього процесу та проектування розміщення аероіонізаторів у заданих приміщеннях. Визначенню картини поширення від'ємних аероіонів від аероіонізаторів на розрахунковій площі присвячено наукові роботи [1,2]. У цих роботах розроблено спосіб геометричного моделювання розподілу

від'ємних аероіонів, який дає змогу визначити зони аероіонного комфорту і дискомфорту у виробничому приміщенні від одного і більше іонізаторів на горизонтальній площині. Також в роботі [4] розглянуто розподіл концентрації аероіонів на площині, яка розташована під деяким гострим кутом до горизонтальної площини (похила площина), а саме розроблено алгоритм визначення ізоліній від'ємних аероіонів.

Формулювання цілі статті

Пропонується на основі результатів дослідження розподілення концентрації від'ємних аероіонів від штучного джерела аероіонного випромінювання розробити алгоритм моделювання поширення концентрації від'ємних аероіонів на похилій площині з метою візуалізації і прогнозування цього процесу в середовищі математичного процесора Maple.

Виклад основного матеріалу

Геометрична модель розподілу концентрації аероіонів з одного джерела випромінювання на горизонтальній площині являє собою ізолінії у вигляді кіл як результат перетину з горизонтальною площиною сфер як просторових ізоповерхонь [2]. У реальних умовах існують приміщення, в яких підлога (а отже, і розрахунковий рівень – рівень дихання людини) знаходиться під деяким гострим кутом до горизонтальної площини (наприклад, глядацька зала у кінотеатрі, лекційна аудиторія тощо). На рис. 1: Σ_1 – горизонтальна площина; Σ'_1 – похила площина; α – кут нахилу площини, відносно якої ведуться розрахунки (похила площина); N – центр аероіонізатора; N_1 – проекція центру аероіонізатора на розрахункову площину; A'_i – точка, інцидентна заданій ізолінії концентрації від'ємних аероіонів.

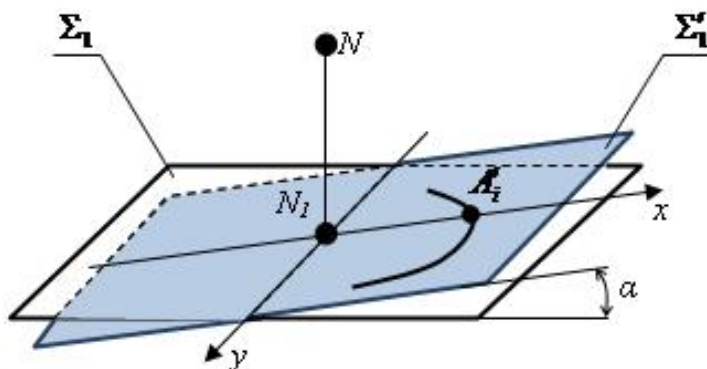


Рис. 1. Положення розрахункової площини

Враховуючи закон розподілу концентрації від'ємних аероіонів на горизонтальній площині від розсіювального джерела аероіонів [2] і кут нахилу α розрахункової площини відносно горизонтальної, запишемо рівняння для моделювання розподілу концентрації від'ємних аероіонів на похилій площині:

$$n' = \frac{1}{a(x^2 + y^2) + b} 2^{-xtga}, \quad (1)$$

де n' – рівень концентрації від'ємних аероіонів у розрахунковій точці на похилій площині, іон/см³; a, b – коефіцієнти, які залежать від сили випромінювання джерела [2]; α – кут нахилу розрахункової площини відносно горизонтальної, 0 ; x, y – координати проекції центру аероіонізатора на похилій площині, м.

На підставі отриманої закономірності (1) складемо алгоритм моделювання розподілу концентрації від'ємних аероіонів на похилій площині від одного аероіонізатора в середовищі математичного процесора Maple. Алгоритм містить такі кроки: спочатку задаємо значення

коєфіцієнтів a і b , які характеризують випромінювальну здатність аероіонізатора; далі задаємося координатами джерела аероіонного випромінювання відносно прийнятої системи координат; потім визначаємо необхідний рівень концентрації від'ємних аероіонів $n_{зад}$ і кут нахилу розрахункової площини α ; після задання всіх необхідних параметрів задаємося законом аероіонного поширення; далі виводимо на екран зображення отриманої кривої – ізолінії заданої концентрації від'ємних аероіонів. Для виведення на екран зображення вихідної ізолінії концентрації аероіонів скористаємося оператором *plot* і *implicitplot*.

Як особливий випадок на рис. 2 наведено лістинг програми (а) і результат роботи цієї програми (б).

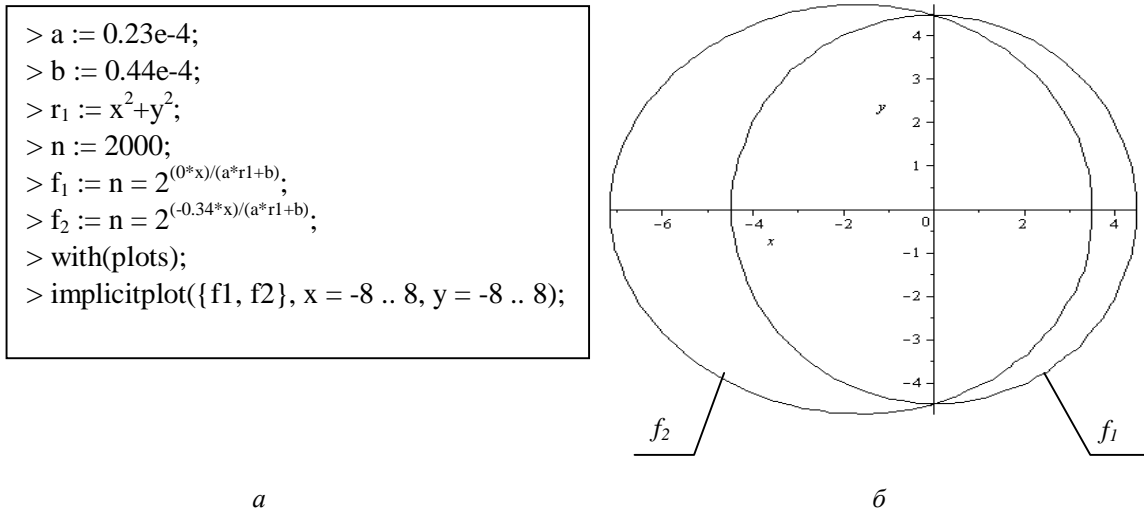


Рис. 2. Комп'ютерна система аналізу аероіонного розповсюдження на похилій площині

На рис. 2, б наведено дві криві: f_1 – крива при $\alpha=0^\circ$; f_2 – крива при $\alpha=20^\circ$. Аналіз отриманих ізоліній показує, що при зміні кута нахилу розрахункової площини відбувається зсув координат відповідних точок заданої кривої.

На рис. 3 наведено ізолінії концентрації аероіонів від одного джерела аероіонного випромінювання на похилій площині. Кожна ізолінія відповідає заданим значенням концентрації від'ємних аероіонів: 2000, 4000, 6000 і 10000 іон/см³ при куті нахилу розрахункової площини $\alpha=5^\circ$ і $\alpha=10^\circ$.

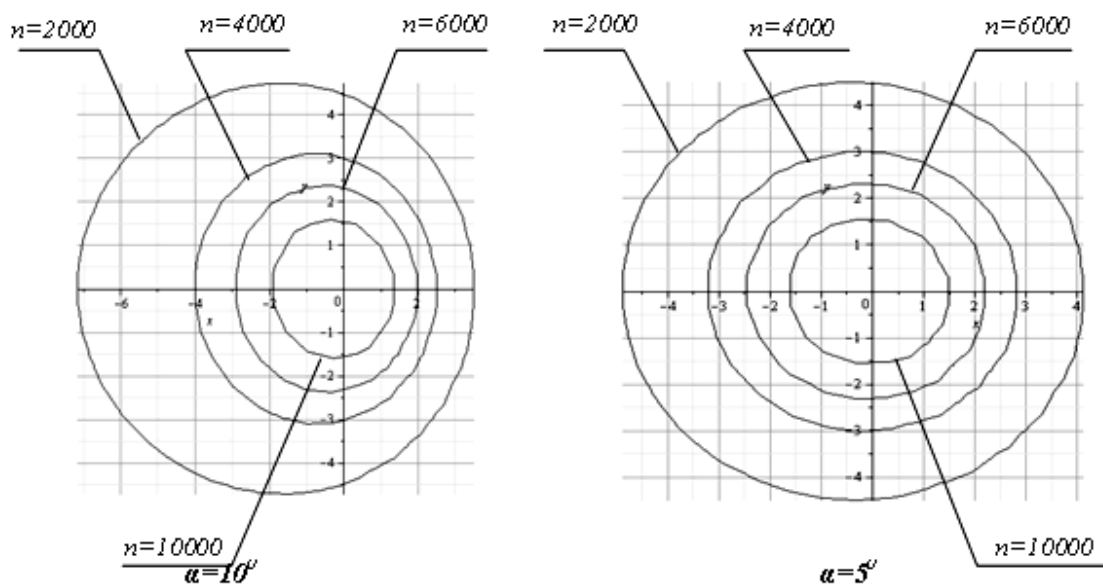


Рис. 3. Ізолінії концентрації від'ємних аероіонів на похилій площині при $\alpha=10^\circ$ і $\alpha=5^\circ$

Порівнюючи отримані ізолінії, можна зробити висновок, що зі збільшенням кута нахилу розрахункової площини відносно горизонтальної посилюється викривлення ізоліній. При певних параметрах ізолінія розривається, наприклад, для рис. 4 ($\alpha=20^\circ$) - $n=2000$ іон/см³ і $n=4000$ іон/см³. Чим більше значення заданої (нормованої) концентрації від'ємних аероіонів, тим більша ймовірність існування ізоліній концентрації аероіонів.

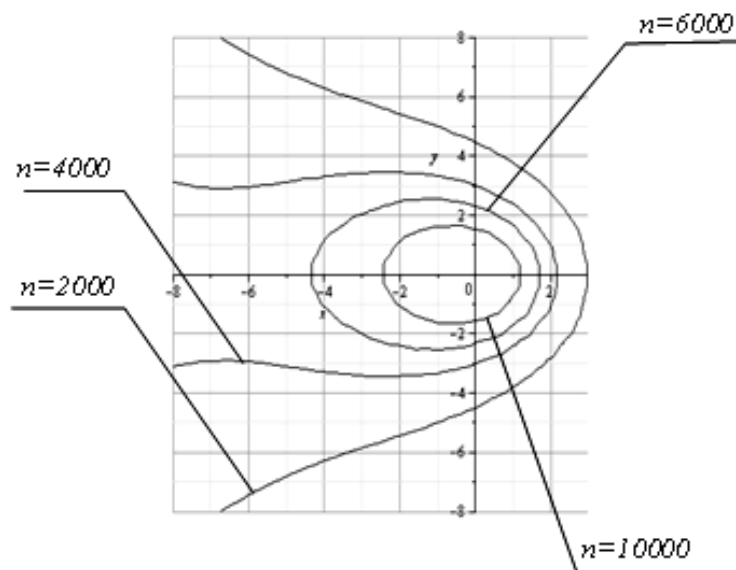


Рис. 4. Ізолінії концентрації від'ємних аероіонів на похилій площині при $\alpha=20^\circ$

Запропонована комп'ютерна система дозволяє візуалізувати процес поширення концентрації від'ємних аероіонів від аероіонізаторів у заданому просторі і проаналізувати отриману картину. Результати аналізу можна використати при проектуванні місця розташування систем іонізації повітря приміщень, які володіють похилою розрахунковою площиною – лекційних аудиторій, кінотеатрів тощо, а також при розміщенні робочого персоналу у таких приміщеннях.

Висновки

Запропонована у статті комп'ютерна система аналізу аероіонного розподілення на похилій площині використовує середовище математичного процесора Maple і дозволяє візуалізувати і прогнозувати процес розподілення концентрації від'ємних аероіонів у приміщеннях, які мають такого типу площини. Запропонований алгоритм моделювання розподілу концентрації від'ємних аероіонів в середовищі математичного процесора Maple дозволяє отримати картину розподілу аероіонів на похилій площині за різного її положення (за зміни кута нахилу до горизонтальної площини від 0° до 90°). Отримані результати можна використати при проектуванні місця розташування системи іонізації повітря з метою забезпечення нормованих показників концентрації від'ємних аероіонів у виробничих і побутових приміщеннях.

1. Спосіб визначення однакового рівня концентрації аероіонів від двох розсіювальних джерел аероіонного випромінювання [Текст]: пат. 48805 Україна: МПК9 А62L/9/22 , А61N 1/44 / О.В. Строкань, О.В. Івженко, А.Я. Чураков; заявник і патентовласник Таврійський державний агротехнологічний університет. - №2009 04661; заявлено 12.05.2009; опубл. 12.04.2010, Бюл. №7. – 4 с.
2. Строкань О.В. Геометричне моделювання процесу розподілення негативних аероіонів у закритому просторі [Текст]: дисс.... канд. техн.наук /О.В. Строкань.– Мелітополь, 1010. – 172 с.
3. Чижевский А.Л. Аэроионы и жизнь. Беседы с Циолковским [Текст] / А.Л. Чижевский. – М.:

Мысль, 1999. – 716 с. 4. Чураков А.Я. Визначення ізоліній концентрації аероіонів на похилій площині [Текст]/ А.Я. Чураков, О.В. Строкань, О.І. Морозова // Прикладна геометрія та інженерна графіка. - Мелітополь, 2011. –Т.50. – С. 85–89.

УДК 004.519.7 (045)

А. Вавіленкова

Національний авіаційний університет

ЗАСТОСУВАННЯ ФОРМАЛЬНИХ АЛГОРИТМІВ У СТРУКТУРНІЙ ЛІНГВІСТИЦІ

© Вавіленкова А., 2014

Проаналізовано концепцію роботи формальних алгоритмів, наведено приклад використання команд машини Тюрінга для аналізу речення природною мовою. Запропоновано алгоритм автоматизованої лінгвістичної обробки електронних документів на основі побудови логіко-лінгвістичних моделей речень.

Ключові слова: природна мова, формальні алгоритми, семантична структура, логіко-лінгвістична модель, текстова база.

The article analyzes the concept of formal algorithms; it demonstrates an example of using the commands of Turing machine to analyze natural language sentence. The research proposes the algorithm of automated linguistic processing of electronic documents based on constructing logico-linguistic models of sentences.

Key words: natural language, formal algorithms, semantic structure, logico-linguistic model, text base.

Постановка проблеми

Основним предметом досліджень структурної лінгвістики є структура природної мови, яка являє собою мережу відношень між елементами мовної системи, що впорядковані і знаходяться в ієрархічній залежності у межах певних рівнів. Структурний опис мови передбачає такий аналіз тексту, який дасть змогу виділити узагальнені інваріантні одиниці та співвіднести їх з конкретними мовними сегментами на основі чітких правил реалізації [1].

Сьогодні структурна лінгвістика займається розробленням точних методів вивчення природних мов, використовуючи математичні підходи. Зокрема особливої уваги заслуговує створення таких формальних алгоритмів, які дозволили б аналізувати текстові документи за змістом.

Сформуванню уявлення про структуру тексту можна, здійснивши його лінгвістичний аналіз, в результаті якого буде створена текстова база, що містить семантичне представлення тексту, а у вигляді взаємопов'язаних пропозицій фіксує те, про що було сказано у тексті. Тестова база повинна відповідати критерію локальної та глобальної зв'язності. Перший передбачає те, що пропозиції, які зустрічаються у тексті, повинні бути пов'язані між собою. Глобальна ж зв'язність забезпечує розуміння та виявлення основної теми тексту і роль кожної фрази у ньому з погляду розкриття основної теми.

Автоматизований структурний аналіз текстів можна здійснити лише застосовуючи формалізовані правила виявлення міжфразових зв'язків. Це, своєю чергою, можливо при виявленні та систематизації трансформацій, що виникають при глибинних лексико-синтаксичних перетвореннях [2].

Отже, сьогодні актуальною є задача формалізації процесу виявлення глибинних зв'язків між простими та складними частинами текстів з метою вилучення знань з текстової інформації.

Аналіз останніх досліджень та публікацій

Над можливістю вилучення змісту з текстової інформації сьогодні працюють зарубіжні та вітчизняні вчені. Процес вилучення знань з текстових документів не можливий без залучення основ теорії текстів та математичного апарату, який дасть змогу формалізувати та автоматизувати такий процес (власне те, чим і займається структурна лінгвістика).

Так, в області філології з'являються роботи з теорії організації текстів. Наприклад, робота Ю.Н. Земської "Теорія тексту" узагальнює існуючі сьогодні гіпотези конструктивної, семантичної та комунікативно-прагматичної організації текстів [3]. Дж. Лайонз [4] знаходить відповідність між формою, структурою складових та лексичним значенням слів і речень природною мовою. Також спроби неявної формалізації структури текстів зроблено у роботі Філіппова [5], де автор окреслює чіткі способи організації тематичних та змістовних зв'язків, проте без використання математичних підходів.

Величезним внеском у автоматизацію процесу розуміння електронних документів став лексикографічний словник української мови [6]. Він є тезаурусом, що містить лексичні зв'язки між словами, надає вичерпну інформацію щодо тлумачення слів, їх синонімів, граматичних форм, фразеологізмів і т.д. Роль організації концептуальних зв'язків у текстах за допомогою використання лексичних мереж висвітлюють автори роботи "Cognitive linguistics: basic readings" [7]. Особливу увагу звернено на технології автоматичного аналізу тексту, що базується на лінгвістичних моделях, у межах проведення міжнародної конференції "Діалог" [8], коли було апробовано методи верифікації та оцінювання результатів лінгвістичних досліджень, а також різноманітних систем аналізу текстів. Питання використання якісних формальних алгоритмів аналітичної обробки текстової інформації піднімають на таких щорічних конференціях та симпозіумах, як International Multiconference on Computer Science and Information, International Conference "Knowledge - Dialogue – Solution" та ін.

Незважаючи на численні підходи до лінгвістичного аналізу текстів і маси формальних алгоритмів структуризації текстової інформації, змістовна обробка електронних документів є недосконалою, а всі теоретичні дослідження у цій сфері знаходяться на стадії розроблення.

Формулювання цілі статті

Синтаксичні одиниці, з яких будуються речення природною мовою, численні, але їх множина скінченна. Довільна мовна система має кінцевий словник простих одиниць, а також скінченний набір правил, які встановлюють зв'язок між двома структурними рівнями та дозволяють виявити, які послідовності одиниць є реченнями, а які – ні. Саме на створення та застосування формальних правил виявлення лексичних зв'язків між одиницями текстів є метою подальшого дослідження.

Метою статті є аналіз існуючих формальних алгоритмів обробки текстової інформації та створення нових алгоритмів, які здатні були б опрацьовувати текстову інформацію на семантичному рівні.

Виклад основного матеріалу

Для досліджень в області автоматизованої аналітичної обробки текстової інформації важливо, щоб довільна програма обчислення функції, область визначення якої утворено реченнями природної мови, була програмою, яка моделюється автоматом, не менш потужним, ніж машина Тюрінга. Адже до формальних алгоритмів, відомих сьогодні, зараховують насамперед машину Тюрінга, машину Поста, нормальні алгоритми Маркова, рекурсивні функції та стохастичні алгоритми. Якщо розглядати аналізатор текстів як деякий автомат з визначеним числом станів, то процес його роботи можна поділити на:

- операцію розбору – виконує функцію перетворення тексту природною мовою з неформалізованого вигляду на формалізоване внутрішнє представлення;
- аналізу – функції перетворення даних, що існують у внутрішньому представленні та виведення на їх основі нових даних також у формалізованому вигляді;
- синтезу – функції формування відповіді природною мовою, адекватній внутрішньому формалізованому представленню [9].

Спробуємо на прикладі складного речення природною мовою довести неможливість застосування існуючих формальних алгоритмів для вилучення знань із текстової інформації.

Машина Тюрінга – це впорядкована п’ятірка:

$$M = \langle A, Q, P, q_1, q_0 \rangle,$$

де $A = \{a_0, a_1, \dots, a_n\}$ – кінцевий зовнішній алфавіт, завжди припускаємо, що $n \geq 1, a_0 = 0, a_1 = 1$.

$Q = \{q_0, q_1, \dots, q_m\}$ – кінцевий алфавіт внутрішніх станів; $P = T\{i, j\}, 0 \leq i \leq m, 0 \leq j \leq n$ – програма, що складається з команд $T(i, j)$, кожна з яких є словом вигляду: $q_i a_j \rightarrow a_l d q_k$ ($0 \leq k \leq m, 0 \leq l \leq n$), де q_i – внутрішній стан машини; a_j – символ, що зчитується; q_k – новий внутрішній стан; a_l – новий символ, що записується; d – напрям руху пристрою зчитування; L – ліворуч; R – праворуч; E – на місці; q_0 – початковий стан; q_m – кінцевий стан.

Передбачається, що кожній парі $q_i a_j$ відповідає точно одна команда. Множина таких команд називається програмою машини, а отже, у програмі є $m \times (n + 1)$ команд.

Тобто застосування формального алгоритму – машини Тюрінга – для аналізу речення природною мовою означає, що має бути розроблений такий набір команд, який дав би змогу за скінченну кількість кроків зчитати набір слів речення саме у такій послідовності, в якій їх вжито у реченні. Це означає, що для речення, яке матиме іншу структуру, тип чи навіть кількість слів, потрібно створювати новий алгоритм.

Наприклад, нехай є речення природною мовою: “У випадках, коли чинним законодавством передбачено можливість проведення митних процедур без подання декларації, застосовується законодавство”. Для зчитування даного речення пропонується набір команд:

- T(0,п)=(0,П,1);
- T(1,)=(0,П,2);
- T(2,i)=(0,П,1);
- T(1,,)=(0,П,1);
- T(2,c)=(0,П,1);
- T(2,пр)=(0,П,1);
- T(2,д)=(0,П,1);
- T(2,діє)=(0,П,1).

Тут слова замінено першими літерами частин мови, якими вони є.

Нульовим станом вважають знаходження маркера над першим словом речення природною мовою. Речення вважається зчитаним, якщо маркер знаходиться над останнім словом.

Процес виконання команд та поточні результати для цього прикладу наведено в табл. 1.

Таблиця 1

Результати роботи машини Тюрінга

№ стану	Поточний рядок символів	Команда	Отриманий рядок символів
1	2	3	4
0	У випадках, коли чинним законодавством...	T(0,п)=(0,П,1)	У_випадках, коли чинним законодавством...
1	У_випадках, коли чинним законодавством...	T(1,)=(0,П,2)	У випадках , коли чинним законодавством...
2	У випадках , коли чинним законодавством...	T(2,i)=(0,П,1)	У випадках, коли чинним законодавством...
1	У випадках, коли чинним законодавством...	T(1,,)=(0,П,1)	У випадках, _ коли чинним законодавством...
1	У випадках, _ коли чинним законодавством...	T(1,)=(0,П,2)	У випадках, коли чинним законодавством...
2	У випадках, коли чинним законодавством...	T(2,c)=(0,П,1)	У випадках, коли_чинним законодавством...
1	У випадках, коли_чинним законодавством...	T(1,)=(0,П,2)	У випадках, коли чинним законодавством...
2	У випадках, коли чинним законодавством...	T(2,пр)=(0,П,1)	У випадках, коли чинним_ законодавством...

Задані команди дали змогу пройти шлях від першого до останнього слова речення, отже, запропонований алгоритм правильний. Не важко помітити, що зміни в структурі речення, зокрема вживання знаків пунктуації, спричинять необхідність написання додаткових команд в інструкції. Також зчитування речень природною мовою за допомогою машини Тюрінга приводить до значних затрат часу, крім того, не дає можливості аналізувати речення за змістом.

Машина Поста складається з нескінченної стрічки, розділеної на однакові комірки, що можуть бути або пустими – 0, або містити мітку 1, та зчитувального пристрою (каретки), здатної пересуватися по стрічці, перевіряти наявність мітки, стерти та записувати мітку.

Стан машини Поста описується станом стрічки та положенням каретки. Кареткою керує програма, що складається із команд, кожна з яких має такий синтаксис: iKj , де i – номер команди, K – дія каретки, j – номер наступної команди.

Для машини Поста існує шість типів команд:

V_j – поставити мітку і перейти до j -го рядка програми;

X_j – стерти мітку і перейти до j -го рядка програми;

$< -j$ – переміститися ліворуч і перейти до j -го рядка програми;

$- > j$ – переміститися праворуч і перейти до j -го рядка програми;

$? j_1; j_2$ – якщо в комірці немає мітки, то перейти до j_1 -го рядочку програми, інакше перейти до j_2 рядка програми;

! – завершення програми.

Застосування машини Поста для аналізу речення природною мовою дає результат, аналогічний машині Тюрінга.

До формальних алгоритмів також належать **нормальні алгоритми Маркова**, що являють собою систему послідовних застосувань підстановок, які реалізують певні процедури отримання нових слів з базових, побудованих із символів деякого алфавіту. Нормальні алгоритми є вербальними, тобто призначені для застосування до слів у різноманітних алфавітах. Визначення довільного нормального алгоритму складається з двох частин: визначення алфавіту алгоритму та визначення його схеми. Простими формулами підстановки називають слова вигляду $L \rightarrow D$, де L і D – два довільні слова в алфавіті алгоритму (відповідно ліва і права частини формули). $L \rightarrow \cdot D$ – кінцева формула.

Існуючі формальні алгоритми орієнтовані на чітке виконання команд за заздалегідь визначеним шаблоном. Для природномовних текстів необхідні гнучкіші алгоритми, які дозволять працювати з реченнями різної складності, дадуть змогу аналізувати їх за змістом та вилучати з текстів знання.

Алгоритмом, що використовує для структурного аналізу тексту формальні моделі, є алгоритм автоматизованої лінгвістичної обробки електронних документів на основі побудови логіко-лінгвістичних моделей речень.

Весь електронний документ можна представити як сукупність логіко-лінгвістичних моделей речень природною мовою:

$$t = \bigwedge_{g=1}^{N(t)} L_g(S_g), \quad (1)$$

де $L_g(S_g)$ – логіко-лінгвістична модель речення S_g , $g = \overline{1, N(t)}$; $N(t)$ – кількість речень у тексті t .

Логіко-лінгвістична модель речення має вигляд [монографія]:

$$L(S) = \bigwedge_{\mu=1}^{v(S)} L_\mu(S), \quad (2)$$

де $L_\mu(S)$ – простий предикат, що описує частину речення S , яка відображає закінчений зміст; $\mu = \overline{1, v(S)}$, $v(S)$ – кількість частин речення S , які відображають закінчений зміст.

Алгоритм автоматизованої лінгвістичної обробки електронних документів на основі побудови логіко-лінгвістичних моделей речень полягає у виконанні таких кроків.

1. Побудова схематичної структури документа, тобто розбиття тексту на розділи, частини, абзаци, речення. На цьому етапі формується множина складних частин електронного документа $F = \{f_1, \dots, f_j, \dots, f_m\}$, де $j = \overline{1, m}$, m – кількість складних синтаксичних частин та множина абзаців $A = \{a_1, \dots, a_k, \dots, a_q\}$ – множина абзаців тексту, $k = \overline{1, q}$, q – кількість абзаців.

2. Формування стійких словосполучень. Цей етап зводиться до послідовної перевірки необхідних умов у вигляді формалізованих правил формування словосполучень природної мови [10]. Якщо правило виконується, то утворюється вузол, для якого ініціалізуються всі синтаксичні характеристики; службові частини мови, що були пов'язані зі словами, які увійшли до словосполучення, стають атрибутами вузлів.

3. Побудова логіко-лінгвістичних моделей речень природної мови $L(S)$, що входять до електронного документа t . На цьому етапі кожному реченню природної мови відповідає логічна формула [11].

4. Встановлення відношень між вузлами. Відбувається послідовний пошук змістовних зв'язків між логіко-лінгвістичними моделями, що виконується завдяки послідовному застосуванню формалізованих правил використання засобів когезії у природно мовних текстах. Це такі прийоми як семантичне повторення, дискурсивні слова, еліпсис, синтаксичний паралелізм, анафоричні зв'язки та ін. Результатом виконання цього етапу алгоритму є ідентифікація та уніфікація у логіко-лінгвістичних моделях речень всього тексту синонімів, спільнокореневих та похідних слів, відновлення займенників та перетворення сполучних слів на відповідні логічні операції.

5. Створення лінгвістичного корпусу документа передбачає аналіз використаних формальних правил застосування засобів когезії, логічних зв'язків між абзацами тексту, встановлення виду тематичних прогресій та рем, вжитих у тексті. Усі отримані дані зіставляються із даними словникових статей, що містять заздалегідь прописані граматичні та синтаксичні особливості текстів різних стилів. Внаслідок такого аналізу у відповідність кожному тексту можна поставити сукупність логіко-лінгвістичних моделей, що відображають змістовні зв'язки, а також кортеж лінгвістичних характеристик тексту:

$$t' = \langle C, F, B, A \rangle, \quad (3)$$

де $t' \in T$ – конкретний електронний текст із всієї множини текстів; $C = \{c_1, \dots, c_i, \dots, c_n\}$ – множина існуючих типів текстів, $i = \overline{1, n}$, n – кількість типів; $F = \{f_1, \dots, f_j, \dots, f_m\}$ – множина складних синтаксичних частин тексту, $j = \overline{1, m}$, m – кількість складних синтаксичних частин;

B – текстова база, що складається з набору ключових слів тексту та взаємопов'язаних пропозицій;

$A = \{a_1, \dots, a_k, \dots, a_q\}$ – множина абзаців тексту, $k = \overline{1, q}$, q – кількість абзаців.

Нехай аналізується такий фрагмент тексту: *“При здійсненні митного контролю та митного оформлення товарів і транспортних засобів, що переміщуються через митний кордон України, застосовуються виключно нормативно-правові акти, чинні на день прийняття митної декларації митним органом України.*

У випадках, коли чинним законодавством передбачена можливість проведення митних процедур без подання декларації, застосовується законодавство, чинне на день здійснення таких процедур”.

1. Будується схематична структура документа. Множина складних частин тексту складається із одного елемента $F = \{f_1\}$, $m = 1$, а множина абзаців – з двох елементів $A = \{a_1, a_2\}$, $q = 2$. Це суто технічний етап, на якому не встановлюються жодні типи зв'язків між структурними елементами тексту.

2. Формування стійких словосполучень:

- митний контроль;
- митне оформлення товарів;
- оформлення транспортних засобів;
- митний кордон України;
- нормативно-правові акти;
- чинне законодавство;
- митні процедури;
- подання декларації.

3. Побудова логіко-лінгвістичних моделей речень, що становлять фрагмент тексту.

L_1 =Застосовуються (акти, нормативно-правові, виключно, здійснення, контролю, митного),

L_2 =Застосовуються (акти, нормативно-правові, виключно, здійснення, оформлення, митного, товарів),

L_3 =Застосовуються (акти, нормативно-правові, виключно, здійснення, оформлення, митного, засобів, транспортних),

L_4 =Переміщуються (засоби, транспортні, кордон, митний, України).

L_5 =Передбачена (можливість, проведення, процедур, митних, подання, декларації, законодавством, чинним),

L_6 =Застосовується (законодавство, чинне, день, здійснення, процедур, таких, випадках, подання, декларації).

$$L_1 = P_{[1]}(x_{1[1]}, c_{1[1]}, x_{2[1]}, x_{3[1]}, x_{4[1]}, c_{4[1]}),$$

$$L_2 = P'_{[1]}(x'_{1[1]}, c'_{1[1]}, x'_{2[1]}, x'_{3[1]}, x'_{4[1]}, c'_{4[1]}, x'_{5[1]}),$$

$$L_3 = P''_{[1]}(x''_{1[1]}, c''_{1[1]}, x''_{2[1]}, x''_{3[1]}, x''_{4[1]}, c''_{4[1]}, x''_{5[1]}, c''_{5[1]}),$$

$$L_4 = P'''_{[1]}(x'''_{1[1]}, c'''_{1[1]}, x'''_{2[1]}, c'''_{2[1]}, x'''_{3[1]}).$$

$$L_5 = P_{[2]}(x_{1[2]}, x_{2[2]}, x_{3[2]}, c_{3[2]}, x_{4[2]}, x_{5[2]}, x_{6[2]}, c_{6[2]}),$$

$$L_6 = P'_{[2]}(x'_{1[2]}, c'_{1[2]}, x'_{2[2]}, x'_{3[2]}, x'_{4[2]}, c'_{4[2]}, x'_{5[2]}, x'_{6[2]}, x'_{7[2]}).$$

4. Встановлення відношень між вузлами передбачає заміну структурних компонентів логіко-лінгвістичних моделей речень.

$$L_1 = P_{[1]}(x_{1[1]}, c_{1[1]}, x_{2[1]}, x_{3[1]}, x_{4[1]}, c_{4[1]}),$$

$$L_2 = P_{[1]}(x_{1[1]}, c_{1[1]}, x_{2[1]}, x_{3[1]}, x'_{4[1]}, c'_{4[1]}, x'_{5[1]}),$$

$$L_3 = P_{[1]}(x_{1[1]}, c_{1[1]}, x_{2[1]}, x_{3[1]}, x'_{4[1]}, c'_{4[1]}, x''_{5[1]}, c''_{5[1]}),$$

$$L_4 = P'''_{[1]}(x'''_{5[1]}, c'''_{5[1]}, x'''_{2[1]}, c_{4[1]}, x'''_{3[1]}).$$

$$L_5 = P_{[2]}(x_{1[2]}, x_{2[2]}, x_{3[2]}, c_{4[1]}, x_{4[2]}, x_{5[2]}, x_{6[2]}, c_{6[2]}),$$

$$L_6 = P'_{[2]}(x_{6[2]}, c_{6[2]}, x'_{2[2]}, x_{3[1]}, x_{3[2]}, c'_{4[2]}, x'_{5[2]}, x_{4[2]}, x_{5[2]}).$$

5. Створюється лінгвістичний корпус фрагмента, формальному запису якого відповідає кортеж: $t' = \langle \{4\}, \{1\}, B, \{a_1, a_2\} \rangle$ [10].

Висновки

Для автоматизованої обробки текстової інформації на семантичному рівні необхідно використовувати формальні алгоритми, здатні опрацьовувати тексти з урахуванням логічних зв'язків як на рівні речень, так і на рівні фраз природною мовою. Таким алгоритмом є запропонований у статті алгоритм автоматизованої лінгвістичної обробки електронних документів на основі побудови логіко-лінгвістичних моделей речень. Лінгвістичний корпус документа враховує міжфразові зв'язки та граматико-синтаксичні особливості конкретного тексту, а масив

логіко-лінгвістичних моделей дозволяє простежити зв'язок між реченнями природною мовою. Головною особливістю такого формального алгоритму є застосування формалізованих правил використання засобів когезії у природно мовних текстах.

1. Виноградов В.А. Структурная лингвистика // Лингвистический энциклопедический словарь. – М.:1990.
2. Апресян Ю.Д. Лексическая семантика: в 2-х т. Т. 1. / Ю.Д. Апресян. – М.: “Восточная литература”, 1995. – 422 с.
3. Земская Ю.Н. Теория текста: учебное пособие / Ю.Н. Земская, И.Ю. 4. Качесова, Л.М. Комиссарова, Н.В. Панченко, А.А. Чувакин. – М.: Наука, 2010. – 132 с. УРСС, 2004. – 320 с.
5. Филиппов К.А. Лингвистика текста. Курс лекций. / К.А. Филиппов. – Спб.: Издательство С.-Петербургского университета, 2008. – 336с.
6. Український лінгвістичний портал, словники України/ <http://lcorp.ulif.org.ua/dictua/>.
7. Geeraerts Dirk. Cognitive linguistics: basic readings research / Dirk Geeraerts, Rene Dirven, John R. Taylor. – Berlin-New York: Mouton de cruyster, 2006. – 486 p.
8. Международная конференция по компьютерной лингвистике Диалог 2014 / <http://www.dialog-21.ru/dialog2014/>.
9. Шулкин Д.Е. Морфологический и синтаксический разбор текстов как конечный автомат, реализованный семантической нейронной сетью, имеющей структуру синхронизированного линейного дерева / Д.Е. Шулкин // Новые информационные технологии. – 2002. – С.74-85.
10. Вавіленкова А.І. Теоретичні основи аналізу електронних текстів: [монографія] / А.І. Вавіленкова, Д.В. Ланде, О.Є. Литвиненко. – К.: НАУ, 2014. – 250с.
11. Вавіленкова А.І. Логіко-лінгвістична модель як засіб відображення синтаксичних особливостей текстової інформації / А.І. Вавіленкова // Математичні машини та системи. – 2010. – № 2. – С. 134–137.

УДК 616.3-008.1

М. Тимчак, В. Дозорський, Г. Шадріна

Тернопільський національний технічний університет ім. І. Пулюя,
кафедра біотехнічних систем

КОМПОНЕНТНИЙ МЕТОД ОПРАЦЮВАННЯ ЕЛЕКТРОГАСТРОСИГНАЛУ

© Тимчак М., Дозорський В., Шадріна Г., 2014

На основі енергетичної теорії стохастичних сигналів використано компонентний метод для статистичного опрацювання електрогастросигналу, який дасть змогу розширити можливості електрогастросистем.

Ключові слова: електрогастросигнал, шлунково-кишковий тракт, компонентний метод.

The theory of stochastic signal energy based on component method used for statistical processing elektrohastrosignal that will help empower elektrohastrosystem

Key words: elektrohastrosignal, gastrointestinal tract, component method.

Постановка проблеми

За статистичними даними Міністерства охорони здоров'я України, спостерігається тенденція до зростання кількості людей із захворюваннями шлунково-кишкового тракту (ШКТ), що пов'язані з харчовими отруєннями та гострими кишковими інфекціями, зокрема

станом на кінець 2013 року виявлено 430 випадків на 10000 людей, з них 297 – це дорослі, і 153 – діти віком до 18 років. Прогнозовано [1], що до 2020 року кількість спалахів такого роду захворювань подвоїться. Тому завчасна діагностика стану шлунково-кишкового тракту (ШКТ) є актуальною задачею сучасної медицини.

Патологічні стани ШКТ проявляються у зміні функціонування його органів, зокрема у змінах в моториці ШКТ. Ці зміни добре проявляються в електрогастросигналах (ЕГС), що відображають електричну активність органів ШКТ (зміну електричних потенціалів, що виникають в гладких м'язах окремих відділів ШКТ). Належне опрацювання таких сигналів забезпечить можливість завчасної діагностики захворювань ШКТ та дасть змогу виявити функціональні зміни на ранній стадії їх розвитку і провести профілактичні заходи, а у випадку наявності патології запобігти розвитку хвороби відповідним лікуванням.

Отже, враховуючи наведену вище статистику захворюваності ШКТ, актуальною науковою та технічною задачею є удосконалення відомих або розроблення нових ефективніших методів опрацювання ЕГС для підвищення достовірності результатів роботи автоматизованих діагностичних систем.

Аналіз останніх досліджень

Сьогодні поширеними в медичній практиці діагностування захворювань ШКТ є автоматизовані електрогастроентерологічні системи, наприклад, системи “Гастроскан-ГЭМ” (Росія), Digitapper EGG” (Швеція, США), в яких власне і здійснюють відбір та опрацювання ЕГС. В основу роботи програмного забезпечення у таких системах покладено два підходи до подання ЕГС, а саме детерміністський (ЕГС розглядається як суміш періодичних функцій) та ймовірнісний (ЕГС розглядається як стаціонарний процес). При цьому застосовуються методи гармонічного та спектрально-кореляційного аналізу ЕГС. Однак в таких моделях не враховують стохастичного характеру електрогастросигналу (ЕГС) як відображення функціонального стану ШКТ. Однак ці методи мають такі недоліки: неможливість оцінювання часово-фазової структури з метою виявлення часових моментів появи змін у роботі ШКТ та прогнозування тенденцій розвитку захворювання у випадку його наявності. Також самі методи не мають засобів або врахування стохастичної складової ЕГС (детерміністський підхід), або опису коливних процесів (ймовірнісний підхід), яким є ЕГС.

Математичний апарат для опрацювання такого виду сигналів дає енергетична теорія стохастичних сигналів, застосування якої до опису ЕГС, зокрема подання його у вигляді періодично корельованого випадкового процесу, обґрунтовано в праці [2]. У цій праці також розглянуто можливість застосування до опрацювання ЕГС таких методів, як синфазний та компонентний. У праці [3] застосовано до опрацювання ЕГС синфазний метод та інтерпретовано отримані результати. Однак сама апаратно-програмна реалізація методу є доволі складною. Простішим в цьому плані є компонентний метод.

Компонентний метод аналізу моторики ШКТ

Характерною властивістю ПКВП є періодичність його математичного сподівання $m_x(t)$ та кореляційної функції $b_x(t, u)$. Компонентний метод базується на тому, що згадані статистичні характеристики ЕГС є періодичними функціями від часу, а тому можуть бути представлені за допомогою розкладів типу:

$$\hat{m}_x(t) = \sum_{k \in Z} \hat{m}_k \exp\left(ik \frac{2p}{T} t\right), \quad (1)$$

$$\hat{b}_x(t, u) = \sum_{k \in Z} \hat{B}_k(u) \exp\left(ik \frac{2p}{T} t\right), \quad (2)$$

де \hat{m}_k та $\hat{B}_k(u)$ – коефіцієнти розкладів, причому \hat{m}_k є випадковими величинами, а $\hat{B}_k(u)$ – випадковими функціями.

Коефіцієнти розкладів (1) та (2) називають також компонентами характеристик та знаходять за виразами:

$$\hat{m}_k = \frac{1}{T} \int_0^T \hat{m}_x(t) \exp\left(ik \frac{2p}{T} t\right) dt, \quad (3)$$

$$\hat{B}_k(u) = \frac{1}{T} \int_0^T \hat{b}_x(t, u) \exp\left(ik \frac{2p}{T} t\right) dt. \quad (4)$$

Для знаходження кореляційних компонент, окрім статистики (4), використано вираз:

$$\hat{B}_k(u) = \frac{1}{T} \int_0^T [x(t)x(t+u) - m_x(t)m_x(t+u)] \exp\left(-ik \frac{2p}{T} t\right) dt. \quad (5)$$

При відомому математичному сподіванні ЕГС оцінки кореляційних компонент для $T = k \cdot dt \in$ незсунутими, тому незсунутими \in і визначені на їх основі оцінки кореляційної функції.

Компонентний метод опрацювання ЕГС у вигляді дискретної послідовності

Сьогодні усі відомі методи опрацювання ЕГС в автоматизованих діагностичних системах реалізовано у вигляді програмного забезпечення, яке оперує безпосередньо із дискретним ЕГС, тому компонентний метод застосовано для опрацювання послідовності ЕГС як періодично корельованої випадкової послідовності із дискретним часом.

Компонентний метод статистичного оцінювання характеристик послідовності ЕГС із дискретним часом \in незначними модифікаціями статистики стаціонарних випадкових процесів. Тому обґрунтування методу статистичного оцінювання характеристик ґрунтується на понятті **h**-ергодичності, яка виражається як властивість випадкових процесів із дискретним часом [4].

Для ергодичності процесу із дискретним часом загалом відносно середнього необхідно і достатньо, щоб векторний процес з дискретним часом був ергодичним відносно середнього:

$$\overset{o}{x}(n\Delta t) = x(n\Delta t) - \hat{m}_x^T(n\Delta t), \quad n = \overline{0, N-1}, \quad (6)$$

де $x(n\Delta t)$ – послідовність ЕГС із дискретним часом; Δt – крок дискретизації ($\Delta t \geq \frac{1}{2\Delta f}$, де Δf – частота дискретизації ЕГС за теоремою Котельникова; n – номер відліку; N – довжина послідовності ЕГС $x(n\Delta t)$; $\hat{m}_x^T(n\Delta t)$ – періодичне продовження математичного сподівання ЕГС як послідовності $x(n\Delta t)$ із дискретним часом:

$$\hat{m}_x^T(n\Delta t) = \sum_{k=1, N} c_{D_k}(n\Delta t) \hat{m}_x(n\Delta t + k\Delta t N_T), \quad n \in \overline{0, N-1}, \quad (7)$$

де k – номер періоду; N_T – кількість точок, в межах одного періоду ЕГС T , $N_T = \frac{T}{\Delta t}$;

$c_{D_k}(n\Delta t) = \begin{cases} 1, \text{ якщо } n\Delta t \in D_k \\ 0, \text{ якщо } n\Delta t \notin D_k \end{cases}$ – індикаторна функція; $D_k = [k\Delta t N_T, (k+1)\Delta t N_T)$ – часовий діапазон тривалості k -го відгуку ЕГС; $\hat{m}_x(n\Delta t)$ – оцінка математичного сподівання ЕГС.

Комп'ютерна реалізація компонентного методу

Блок-схему опрацювання ЕГС у вигляді дискретної послідовності із застосуванням описаних вище виразів наведено на рис. 1.

Розроблена блок-схема комп'ютерного опрацювання ЕГС (рис. 1) дає змогу оцінити його характеристики для виявлення нових в області діагностики ШКТ інформативних ознак на основі математичної моделі у вигляді періодично корельованої випадкової послідовності із дискретним часом.

Результати комп'ютерного опрацювання компонентного методу аналізу моторики ШКТ

На основі описаного алгоритму (рис. 1) в середовищі Matlab було розроблено програму із графічним інтерфейсом для автоматичного опрацювання окремих вибірок з реєстрограм ЕГС. На рис. 2 наведено вибірки з реєстрограм ЕГС, відібраних від пацієнтів, що знаходяться в стані медичної норми (а) та патології (б) (власне сигнали, взяті з інтернет-баз даних медичних сигналів).

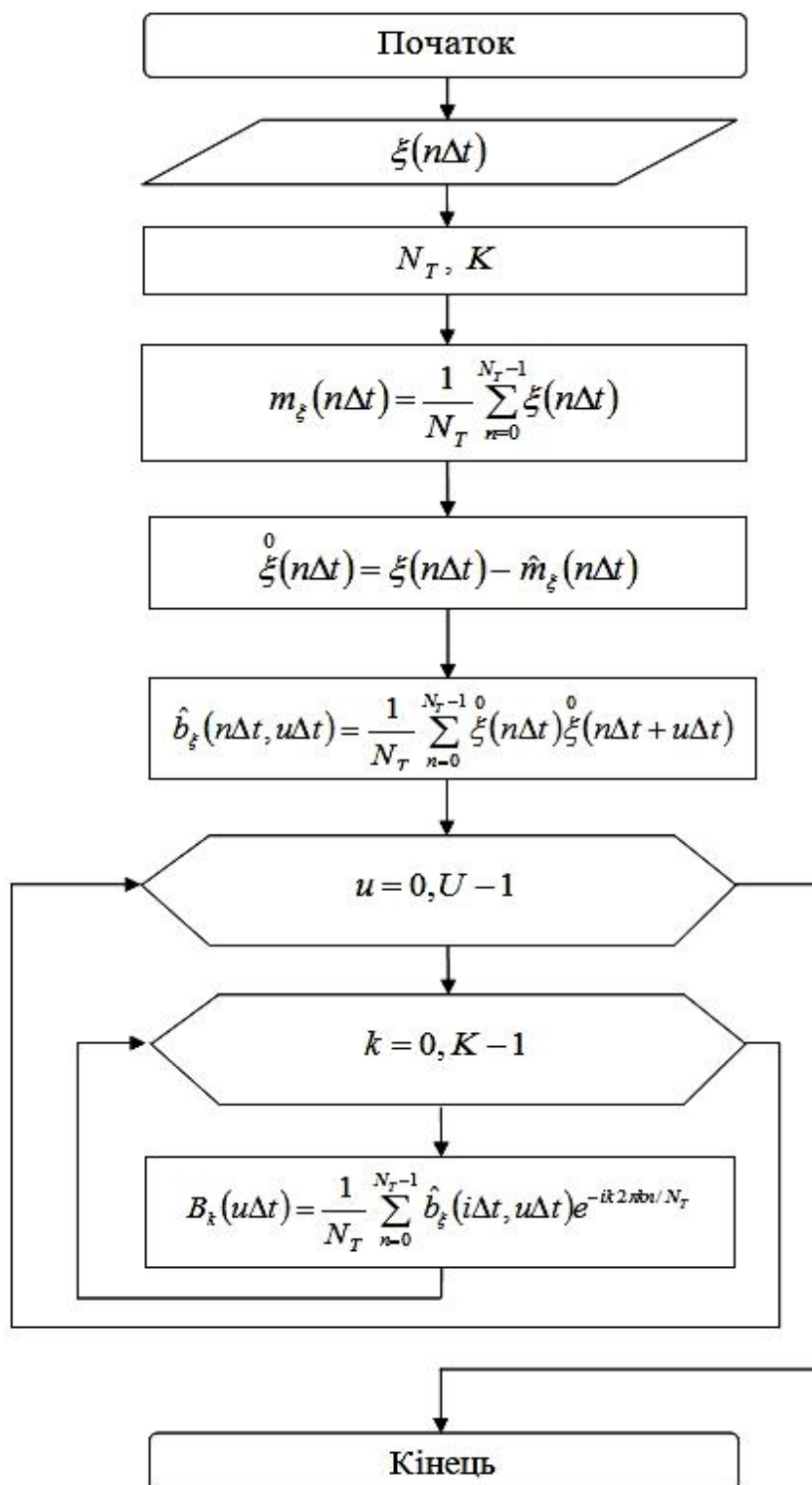
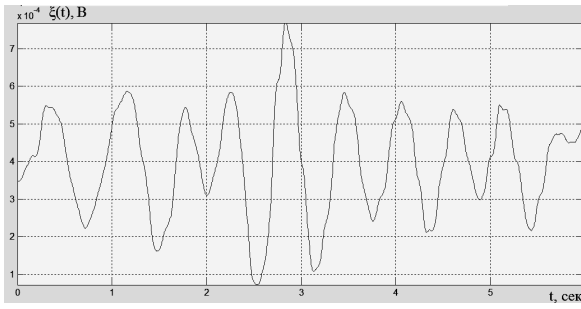
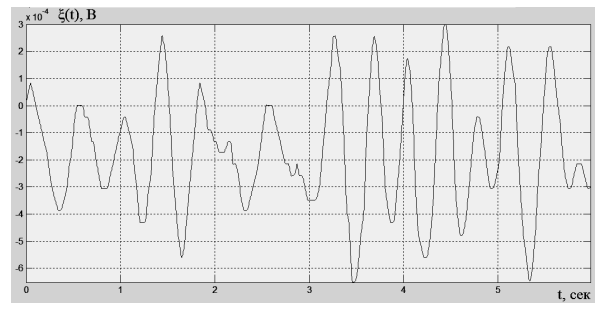


Рис. 1. Блок-схема опрацювання ЕГС компонентним методом



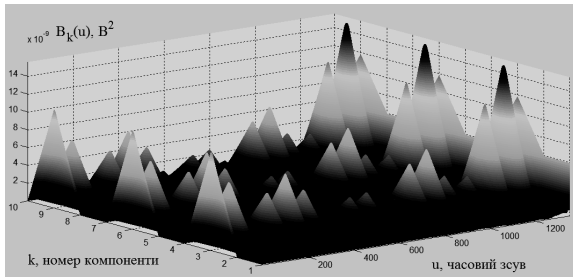
a



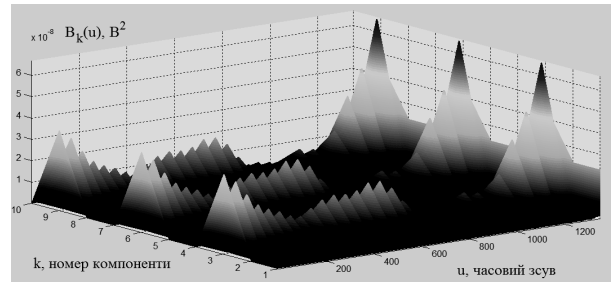
б

Рис. 2. Вибірки з реалізацій ЕГС: а – пацієнт А (норма), б – пацієнт Б (патологія)

На рис. 3 наведено вигляд обчислених оцінок кореляційних компонент для пацієнтів із нормою та патологією.



a



б

Рис. 3. Оцінки кореляційних компонент, обчислені для стану норми (а) та патології (б)

Для оцінювання загального функціонального стану (норма/патологія) виконано усереднення кореляційних компонент (рис. 3) відповідно до виразу (8).

$$M_u \left\{ \hat{B}_k(u) \right\} = \frac{1}{N_k} \sum_{k=1}^{N_k} B_k(u), \quad k = \overline{1, N_k}, \quad (8)$$

де k – номер кореляційної компоненти; N_k – кількість кореляційних компонент.

Результати усереднення кореляційних компонент наведено на рис. 4.

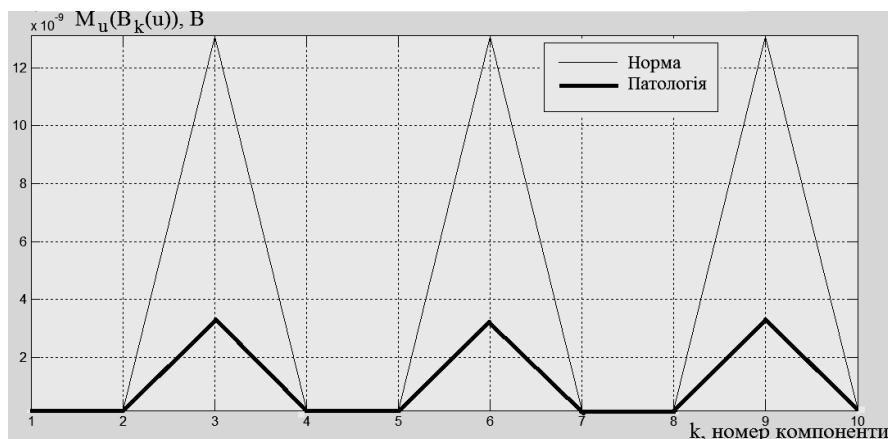


Рис. 4. Реалізації усереднених кореляційних компонент ЕГС (норма і патологія)

На рис. 4 видно, що значення піків оцінок усереднених кореляційних спектральних компонент для пацієнтів з нормою такі патологією зосереджені на тих самих компонентах

(ідентичні за структурою), проте незважаючи на те, що пацієнт А клінічно є здоровим, а Б – хворим, спостерігається незначна зміна значень амплітуди максимумів на 3, 6, та 9-й компоненті, що свідчить про чіткі зміни у функціонуванні ШКТ. Тому обчислені усереднені значення кореляційних компонент ЕГС є інформативними ознаками та можуть бути використані як індикатори захворювань ШКТ, а їх практичне застосування дасть змогу розширити можливості автоматизованих діагностичних систем.

Висновки

У результаті опрацювання ЕГС компонентним методом отримано нові в області діагностики ШКТ інформативні ознаки - кореляційні компоненти, які фактично відповідають функціональному стану органів ШКТ (норма/патологія). Результати досліджень дають змогу застосувати компонентний метод для задач медичної діагностики органів ШКТ на ранніх етапах виникнення та розвитку захворювання.

1. Міністерство охорони здоров'я України [Електронний ресурс]. – Режим доступу : URL : <http://www.moz.gov.ua/ua/portal> / Міністерство охорони здоров'я України. 2. Тимчак М. Математична модель електрогастроентеросигналу для підвищення достовірності електрогастроентеросистем / М. Тимчак, М. Хвостівський, Л. Дедів // Вісник Національного університету "Львівська політехніка", 2012. – №744. – С. 181–186. 3. Тимчак М. Синфазний метод опрацювання електрогастроентеросигналу / М. Тимчак, В. Дунець, О. Гевко, Л. Дедів // Вісник Сумського державного університету. Серія "Технічні науки", 2013. – № 3. С. 38–44. 4. Драган Я. П. Енергетична теорія лінійних моделей стохастичних сигналів / Я. П. Драган. – М.: Львів: Центр стратегічних досліджень екобіотехнічних систем, 1997. –XVI+333с.

РОЗВ'ЯЗУВАННЯ ЗАДАЧІ КОМІВОЯЖЕРА ВЕЛИКИХ РОЗМІРНОСТЕЙ МЕТОДОМ СПІЛЬНИХ РЕБЕР

© Базилевич Р., Кутельмах Р., Томчук А., 2014

Досліджено евристичні алгоритми розв'язання задачі комівояжера: найближчого сусіда, 2-Opt, 3-Opt, Ліна-Кернігана та Ліна-Кернігана-Гельсгауна. Здійснено порівняльний аналіз часу роботи алгоритмів та якості отриманих розв'язків. Запропоновано декомпозиційний метод знаходження розв'язків задачі комівояжера на основі спільних ребер.

Ключові слова: задача комівояжера, NN, 2-Opt, 3-Opt, алгоритм Ліна-Кернігана, алгоритм Ліна-Кернігана-Гельсгауна, евристика, велика розмірність.

Existing heuristic algorithms for solving traveling salesman problem, such as Nearest Neighbor, 2-Opt, 3-Opt, Lin-Kernighan and Lin-Kernighan-Helsgaun have been investigated in this work. The algorithms have been compared in terms of running time and solution quality. Decomposition approach, based on using common edges in multiple solutions, has been proposed.

Key words: traveling salesman problem, NN, 2-Opt, 3-Opt, Lin-Kernighan, LKH, heuristics, large-scale.

Вступ

Проблема побудови оптимальних маршрутів через задану множину точок на площині чи у просторі виникає у багатьох сферах людської діяльності, наприклад, задачах планування та логістики; при виробництві друкованих плат; мінімізації рухів у робототехніці; аналізі структури ДНК та ін. За суттю, усі ці проблеми зводяться до розв'язування задачі комівояжера. Великий інтерес до задачі пов'язаний з використанням її як платформи для дослідження загальних методів комбінаторної оптимізації [1]. Порівняно новим і мало вивченим практичним застосування задачі комівояжера є малювання неперервною лінією. Роберт Бош та Адріан Герман у 2003 році запропонували використання геометричної інтерпретації задачі комівояжера, яка полягає у створенні зображень з допомогою неперервної лінії [2].

Зважаючи на актуальність, існує багато алгоритмів розв'язування задачі комівояжера. Більшість з них незастосовні до задач великих розмірностей, оскільки складність цих алгоритмів зростає експоненційно. Часто доводиться обирати між часом роботи алгоритму та якістю отриманого розв'язку. Компроміс зазвичай знаходять, комбінуючи методи пошуку шляху та його послідовної оптимізації [3].

Постановка задачі

Задачу комівояжера формулюють так: дано множину міст, а також відстань між усіма можливими парами міст. Необхідно знайти шлях, який пролягає через усі міста, та повертається у початкове, окрім того сумарна довжина пройденого шляху має бути мінімальною.

Задачу можна представити у вигляді моделі на графі. Розрізняють різні варіанти задачі, найважливішими з яких є симетрична та асиметрична задачі. У випадку симетричної задачі всі пари ребер між тими самими вершинами мають однакову вагу, тобто для ребер d_{ij} та d_{ji} ваги однакові. В симетричному випадку кількість можливих маршрутів вдвічі менша. Симетрична задача моделюється неорієнтовним графом [1].

Розглянуто симетричну задачу комівояжера, де $V = \{v_1, \dots, v_n\}$ – множина точок, $A = \{(r, s) : r, s \in V\}$ – множина ребер, де $d_{rs} = d_{sr}$ – вартість, що асоціюється з ребром. Для розв'язання задачі комівояжера слід знайти мінімальний замкнутий шлях, що проходить через усі вершини по одному разу [4]. Точки $v_i \in V$ представлені координатами (x_i, y_i) у двовимірному просторі і відстань d_{rs} між будь-якою парою точок представляється як евклідова відстань між r та s . Часто таку задачу ще називають *евклідовою задачею комівояжера*.

Огляд існуючих алгоритмів

Всі існуючі методи розв'язування задачі комівояжера можна поділити на дві групи:

- методи знаходження оптимальних маршрутів;
- наближені методи [5].

Серед алгоритмів пошуку оптимального розв'язку задачі комівояжера відомими є метод гілок та меж, прогресивні методи покращення на основі методів лінійного програмування, комбінації методів гілок та меж та відсікальних площин. Але всі їх поєднує той факт, що вони не застосовні до задач великої розмірності, оскільки обчислювальна складність цих алгоритмів зростає експоненційно. Тому алгоритми цього типу не розглядатимуться.

Одним з найпростіших евристичних алгоритмів є **алгоритм найближчого сусіда** (англ. nearest neighbor). Цей алгоритм, а також йому подібні інколи називають конструктивними, оскільки як їх часто використовують для побудови початкового маршруту, який може піддаватися подальшій оптимізації. Основними перевагами цього алгоритму є його відносно мала обчислювальна складність, простота розуміння і “природність” за принципом дії. Кроки алгоритму:

1. Обрати довільну початкову точку.
2. Знайти найближчу невідвідану точку і перейти до неї.
3. Якщо лишаються невідвідані точки, то перейти до кроку 2.
4. Повернутися до початкової точки.

Обчислювальна складність алгоритму – $O(n^2)$. Якість отриманого розв'язку може бути гіршою від якості оптимального на 25–30 %, проте розв'язок буде отримано з мінімальними часовими затратами [6].

Найчастіше застосовуються такі алгоритми *оптимізації* маршруту, як 2-Opt та 3-Opt. Це сімейство алгоритмів часто називають k-Opt алгоритмами (рис. 1).

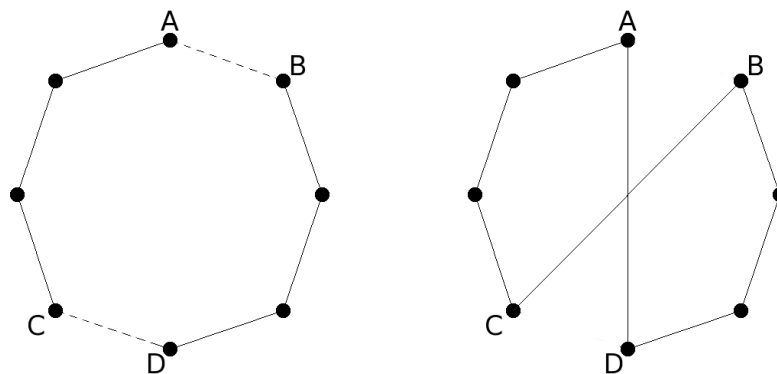


Рис. 1. Перестановка ребер при використанні алгоритму 2-Opt

Алгоритм 2-Opt полягає у видаленні двох ребер з вже існуючого шляху (отриманого за допомогою конструктивного алгоритму) і заміні їх двома новими за умови, що буде отримано

новий коротший шлях. При видаленні ребра АВ та DC і заміною їх новими ребрами AD та BC утвориться новий цикл, і якщо довжина нового шляху менша ніж у попереднього, то перебудовується існуючий шлях. У протилежному випадку аналізується наступне ребро, не змінюючи існуючого маршруту.

Ця дія повторюється для кожної пари ребер. Слід зазначити, що замінити ці два ребра можна лише одним способом, як показано на рис. 1, інакше буде утворено два цикли (рис. 2). Отриманий у результаті оптимізації розв'язок буде гіршим від оптимального на 4–20 %. Обчислювальна складність алгоритму – $O(n^2)$.

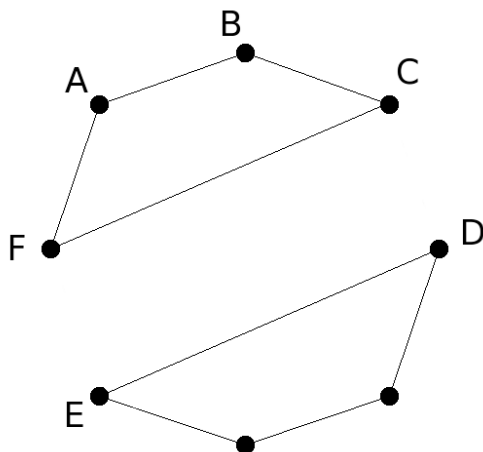


Рис. 2. Можливі утворення циклів у процесі застосування алгоритму 2-Opt

Аналогічно працює алгоритм 3-Opt. Здійснюється заміна трьох ребер. У такому випадку можна утворити дві можливі заміни, як показано на рис. 3. На практиці цей алгоритм часто реалізовується як два або три послідовні застосування алгоритму 2-Opt. Отриманий результат є гіршим від оптимального на 3–4 %. Обчислювальна складність алгоритму $O(n^3)$.

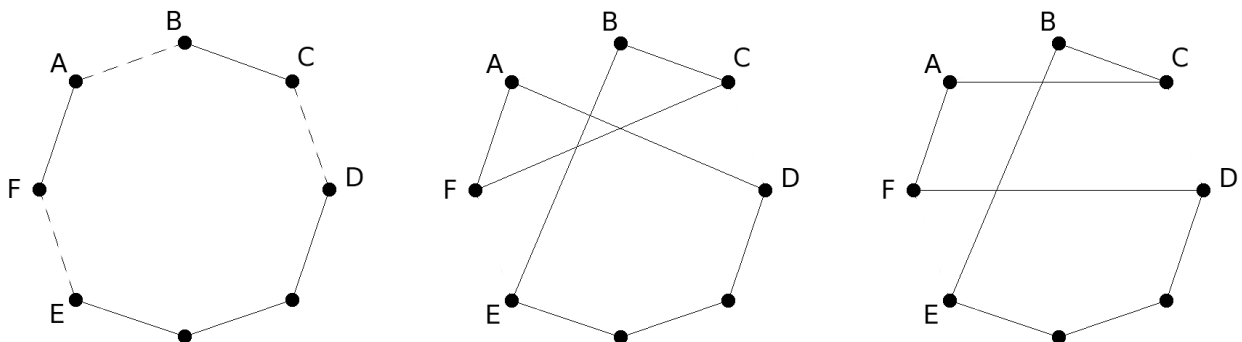


Рис. 3. Можливі перестановки при використанні алгоритму 3-Opt

Логічним продовженням описаних методів оптимізації маршруту є використання “більших k ”, тобто 4, 5 і більше. Зі збільшенням кількості ребер, що видаляються, збільшується кількість можливих перестановок, що, своєю чергою, збільшує час роботи алгоритму і робить його неприйнятним для задач великої розмірності. Вдосконаленням методів типу $k-opt$ є метод Ліна–Кернігана. Цей метод передбачає використання початкового субоптимального маршруту та його оптимізацію. Кількість ребер (число k), які будуть замінені на кожному кроці алгоритму, змінюється динамічно. Особливістю цього алгоритму є те, що він гарантує, що отриманий розв'язок для задачі

розмірністю N точок буде не довшим у $4\sqrt{N}$ разів від оптимального, а також ніколи не є гіршим за початковий маршрут. Метод запропоновано в 1973 році Ліном та Керніганом. Більшість аспектів реалізації описано в оригінальній праці.

Обчислювальна складність методу становить $O(n^{2.2})$. Отриманий результат є гіршим від оптимального на 1–3 %. Сьогодні розроблено багато модифікацій алгоритму Ліна–Кернігана: багаторівневий алгоритм Ліна–Кернігана [8], ланцюговий алгоритм Ліна–Кернігана [9], паралельна модифікація алгоритму Ліна–Кернігана [10]. Найефективнішою модифікацією алгоритму вважається метод Ліна–Кернігана–Гельсгауна, що полягає в удосконаленні правил вибору ребер, що будуть замінені [7].

Експериментальні дослідження методів найближчого сусіда, 2-opt, 3-opt, LK та LKH

З метою порівняння роботи алгоритмів здійснено експериментальні дослідження. Тестові дані містили задачі розмірністю 100, 500, 1000, 5000, 10000, 50000, 100000 точок. Тестові дані було отримано на основі модифікації вхідних даних задачі mona-liza100k. Дослідження проводили на ПК з процесором Intel Celeron M220.

Як еталонний було обрано LKH, оскільки отриманий за його допомогою результат є найближчим до оптимального. Алгоритм Ліна–Кернігана дозволяє отримати результат, що гірший від еталонного менш як на 1%, при цьому максимальне відхилення від цього результату спостерігається під час розв'язання задач розмірністю від 5000 до 50000 точок. Алгоритм 3-Opt дозволяє отримати розв'язок гірший від еталонного у межах до 10 %, залежно від розмірності задачі. Результат роботи алгоритму 2-Opt гірший від еталонного у межах 15 %, і найгірший результат було одержано за допомогою алгоритму найближчого сусіда – він гірший від еталонного у межах 30%. Результати наведено на рис. 4 та в табл. 1.

Таблиця 1

Час роботи алгоритмів найближчого сусіда, 2-opt, 3-opt, LKH для задач розмірністю 10000–100000 точок (секунди)

Розмірність	10000	50000	100000
Алгоритм NN	0,01	0,09	0,2
Алгоритм 2-opt	0,04	0,37	1,3
Алгоритм 3-opt	0,09	0,72	2,5
Алгоритм LKH	15,6	90,3	237,3

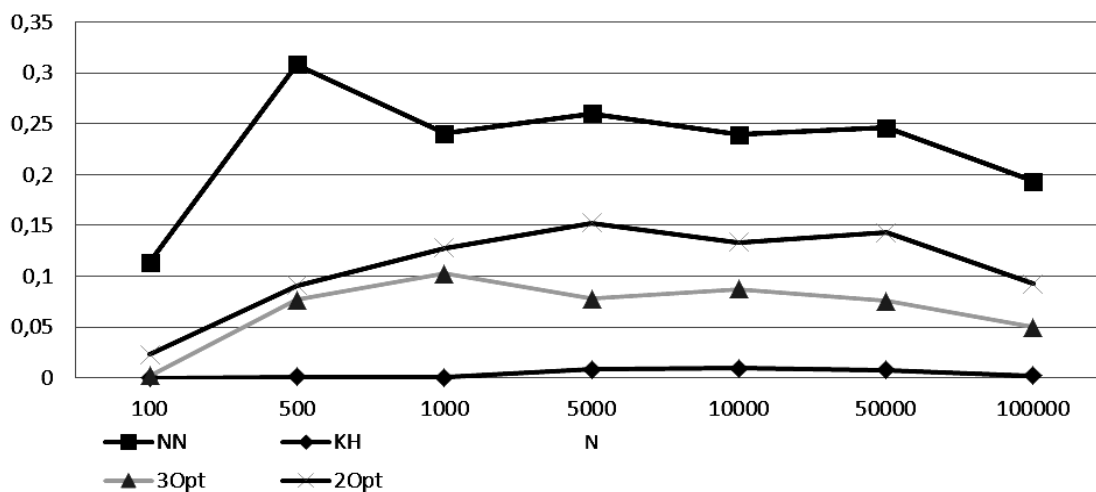


Рис. 4. Якість отриманих розв'язків задач розмірністю 100–100000 точок

Вдосконалення методів знаходження розв'язку задачі на основі спільних ребер

Пропонується метод знаходження розв'язку, що полягає в об'єднанні розв'язків, отриманих за допомогою різних алгоритмів і з меншою обчислювальною складністю. Кроки алгоритму:

1. Знаходження N розв'язків задачі комівояжера за допомогою одного чи різних методів. Допускається застосування методів з різною обчислювальною складністю. В результаті отримується N шляхів S_i (рис. 6).

2. Маючи множину з N шляхів, виділяється множина ребер R , які належать до всіх шляхів, отриманих на кроці 2 (рис. 7).

3. Зшивання отриманих послідовностей ребер у розв'язок задачі.

Для прикладу припустимо, що перед початком роботи алгоритму є деяка множина точок (рис. 5) та декілька розв'язків для цієї множини точок, отримані за допомогою деякого евристичного методу.

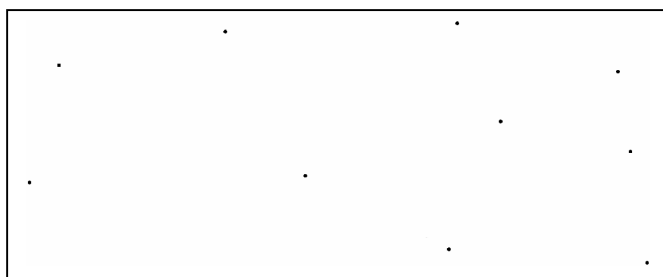


Рис. 5. Початковий набір точок

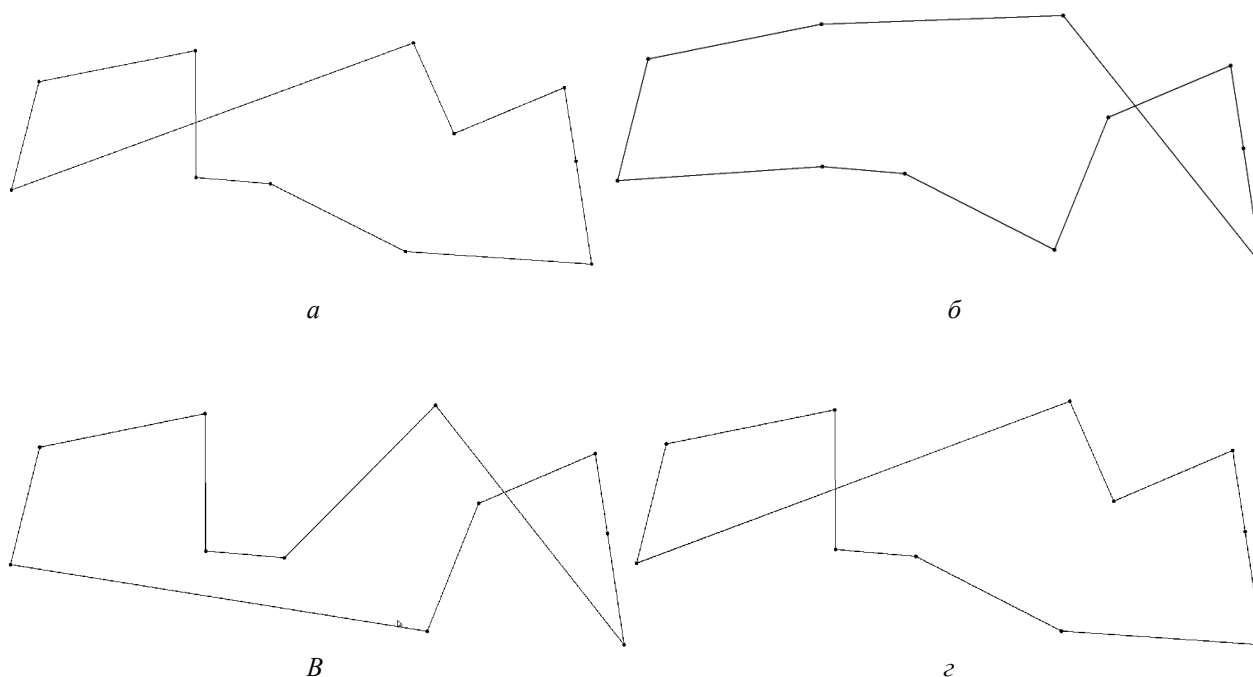


Рис. 6. Розв'язки для даної множини точок

Множину ребер, що входять до всіх розв'язків, показано на рис. 7. Наступним кроком є пошук найкоротших відстаней між не з'єднаними послідовностями ребер (рис. 8).

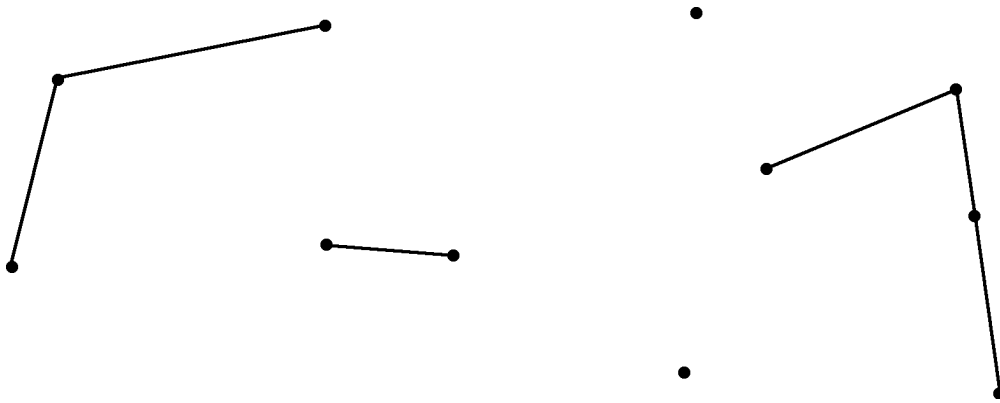


Рис. 7. Ребра, що входять до усіх шляхів

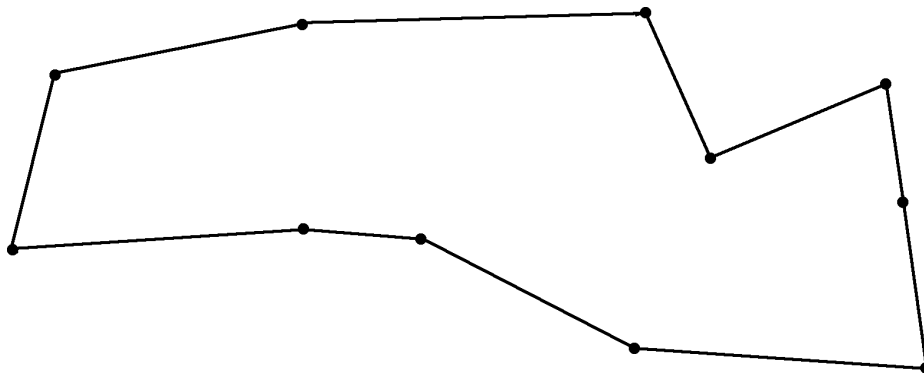


Рис. 8. Результат об'єднання

Проведено експериментальні дослідження запропонованого методу. Як тестову обрано задачу розмірністю 264 точки. Початковими розв'язками задачі було обрано 6 різних маршрутів, одержаних за допомогою методу 2-opt (довжина кожного з яких на 13 % більша від оптимального). Також з метою порівняння з відомими алгоритмами цю задачу було розв'язано за допомогою алгоритмів найближчого сусіда та LKH. Результати наведено в табл. 2.

Таблиця 2

Довжина маршрутів для задачі розмірністю 264 точки, отриманих за допомогою алгоритму спільних ребер та відомих методів відносно оптимального

Алгоритм найближчого сусіда	Алгоритм 2-opt	Алгоритм на основі спільних ребер	Алгоритм LKH
28.7 %	13 %	6,6 %	Оптимальний розв'язок

Отриманий за допомогою алгоритму спільних ребер розв'язок гірший від оптимального на 6,6 % при використанні алгоритму 2-opt для об'єднання фрагментів шляхів. На рис. 9 показано спільні ребра та розв'язок досліджуваної задачі, отриманий за допомогою методу спільних ребер.

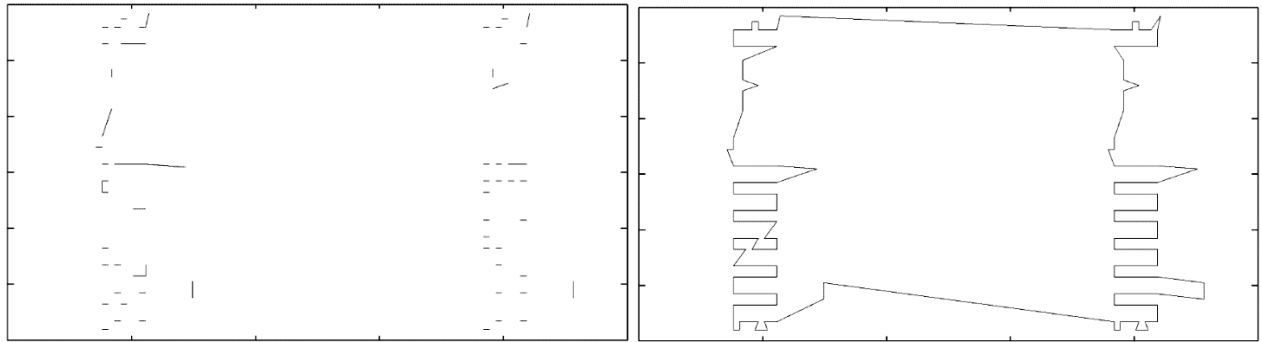


Рис. 9. Спільні ребра та розв'язок, знайдений на їх основі для задачі розмірністю 264 точки

Цю методику доцільно застосувати для розв'язування задач великих розмірностей, зокрема в областях знаходження локальних розв'язків задачі (як додаткову початкову процедуру в методі знаходження розв'язку на основі його нарощення, рис. 10).

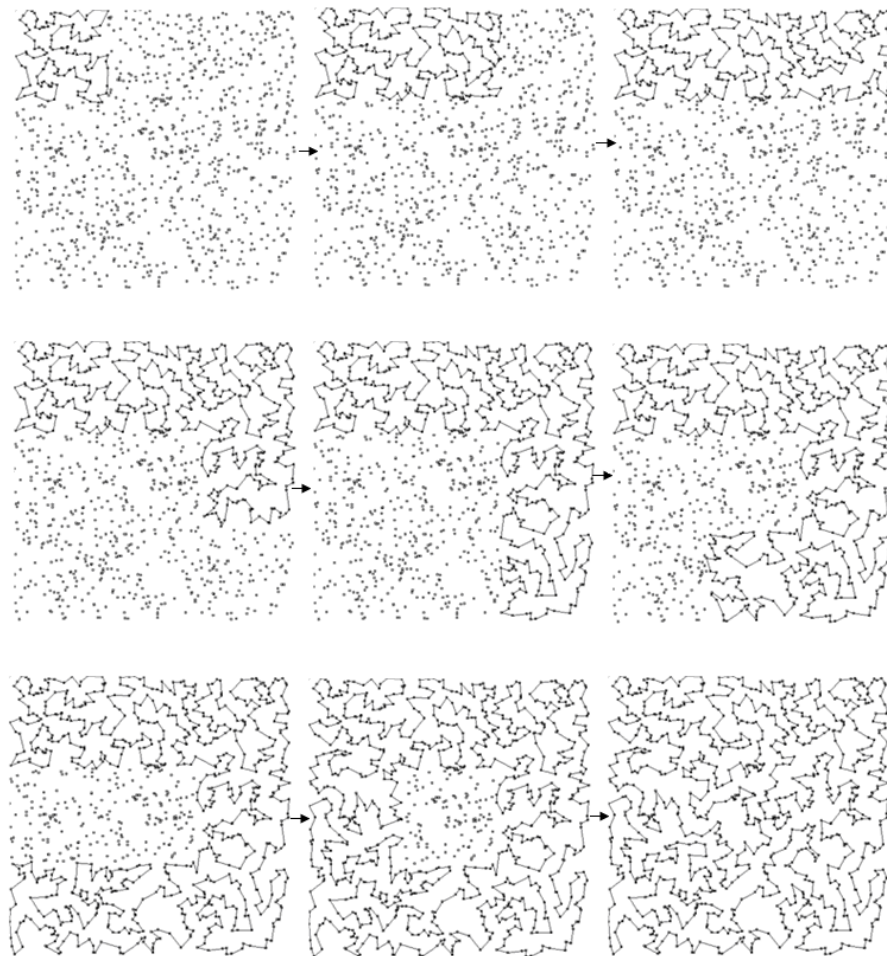


Рис. 10. Знаходження розв'язку задачі на основі нарощення часткового розв'язку

Висновки

У результаті проведених досліджень було проаналізовано алгоритми найближчого сусіда, 2-Орт, 3-Орт, Ліна–Кернігана та Ліна–Кернігана–Гельсгауна. Аналіз отриманих результатів дає змогу зробити припущення про те, що конструктивні алгоритми не є ефективними. У переважній більшості випадків алгоритм Ліна–Кернігана чи Ліна–Кернігана–Гельсгауна є найефективнішими.

Запропоновано методику знаходження розв'язку задачі на основі спільних ребер, що дає можливість широкого розпаралелення та знаходження високоякісних розв'язків задачі. Метод може бути застосованим до задач великих розмірностей. Отриманий за допомогою алгоритму спільних ребер розв'язок гірший від оптимального на 6,6 %. Подальшою роботою буде детальніше дослідження запропонованого методу та його вдосконалення.

1. R. Matai, S.P. Singh, M.L. Mittal, "Traveling Salesman Problem: An Overview of Applications, Formulations, and Solution Approaches", Jaipur, India, 2010, 2. R. Bosch and A. Herman, "Continuous line drawings via the traveling salesman problem," *Operations Research Letters* 3 (2004) 302-303, 3. P. Базилевич, P. Кутельмах // *Комп'ютерні науки та інформаційні технології [Текст] : [зб. наук. пр.] / відп. ред. Ю. М. Рашкевич. – Л. : Видавництво Національного університету "Львівська політехніка", 2009. – 279 с. : іл. – (Вісник / Національний університет "Львівська політехніка"; № 650). – С. 235–244,* 4.E. Nood and J. Been, *An Efficient Transformation of the Generalized Traveling Salesman Problem, October 1991,* 5. D.S. Johnson and L.A. McGeoch, "The Traveling Salesman Problem: A Case Study in Local Optimization", November 20, 1995., 6. Christian Nilsson, *Heuristics for the Traveling Salesman Problem, Linkoping Univercity,* 7. Keld Helsgaun, *General k-opt submoves for the Lin–Kernighan TSP veling heuristic Keld Helsgaun, 1 July 2009,* 8. Chris Walsaw, *A multilevel lin-kernigan-helsgaun algorithm for the treveling salesman problem,* 9. D.Applegate, W.Cook, A.Rohe, *Chained Lin-Kernigan for large traveling salesman problems, Algorithms and Optimization Departament, AT&T Labs-Research* 10.Verhoeven, M.G.A. & Aarts, E.H.L. *A parallel Lin-Kernighan algorithm for the traveling salesman problem, 1994.*

ЗМІСТ

АРХІТЕКТУРА ТА КОМПОНЕНТИ КОМП'ЮТЕРНИХ СИСТЕМ

<i>Tyshchenko O., Pliss I.</i> An evolving reservoir neo-fuzzy network for time series prediction.....	3
<i>Osypenko V., Shvorov A.</i> About some design principles of information-retrieval system and processing of electronic documents in internet	10
<i>Pelech Y., Peleshko D.</i> Marking speech signal based on factor holder smoothness and fast fourier transform	16
<i>Batyuk A., Viter M., Sendzyuk M.</i> Methodological principles of electronic information resources of branch	22
<i>Кунанець Н., Малиновський О.</i> Електронні бібліотеки: обслуговування осіб з особливими потребами	27
<i>Бомба А., Назарук М., Пасічник В.</i> Побудова дифузійноподібної моделі інформаційного процесу поширення знаннєвого потенціалу	35
<i>Огородник І., Огородник Т.</i> Концепція управління якістю міського середовища великого міста.....	45
<i>Скринникова Г.</i> Модель процесу адаптації людино-машинних систем до користувачів	52
<i>Назаркевич М., Троян О.</i> Аналіз сучасних методів та видів графічного захисту друкованих документів	61
<i>Petsko V., Mitsa A., Geche F., Kotsovsky V., Batyuk A.</i> Optimal parameters of four-layer optical structures when the light is incident at an angle	66
<i>Павлюк О.</i> Прогнозування ймовірності роботи та відмови за заданої умови готовності симетричних ієрархічних систем, галужених до 4-го рівня, за допомогою штучних нейронних мереж	72
<i>Обельовська К., Роман В., Ліскевич О.</i> Маршрутизація пакетів з врахуванням пропускних здатностей каналів та кількості проміжних вузлів	78

МОДЕЛЮВАННЯ ПРОЦЕСІВ І СИСТЕМ

<i>Bodyanskiy Ye., Vynokurova O., Mulesa P., Slipchenko O.</i> Self-organizing map and its learning in the fuzzy clustering-classification tasks	83
<i>Басюк Т.</i> Проектування системи автоматизації процесу Seo-оптимізації.....	92
<i>Гожий О., Кобилінський І., Лугінець Д.</i> Підхід до оцінювання ризиків у задачах планування	98
<i>Берко А., Висоцька В., Чирун Л.</i> Теоретико-множинна модель систем електронної контент-комерції.....	105
<i>Седушев О., Буров Є.</i> Форми нечіткостей у базах даних та знань.....	115
<i>Kozhukhivska O., Bidyuk P., Kozhukhivskyi A.</i> Model based decision support system for forecasting financial processes	123
<i>Lytvynenko V.</i> Hybrid swarm negative selection algorithm for dna-microarray data classification	134
<i>Ковальчук А., Цмоць І., Ступень М.</i> Кубічні і лінійні фрактали з елементами алгоритму RSA в шифруванні і дешифруванні зображень	149
<i>Marusenкова Т.</i> Analytic solution for the problem of in-situ calibrating magnetometric 3D-probes based on non-orthogonally placed hall devices	154
<i>Журавчак Л., Федоришин Ю.</i> Розпізнавання провідних та висоомних включень у кусково-однорідному півпросторі при математичному моделюванні усталених коливань електромагнітного поля.....	159

<i>Хома А.</i> Метод реконструкції поверхні з інтерферограми білого світла на основі перетворення Гільберта	168
<i>Ilnytskyi Ja., Holovatch Yu., Kozitsky Yu., Ilnytskyi H.</i> Computer simulations of a stochastic model for the non-immune disease spread.....	176
<i>Пелецишин А., Корж Р.</i> Формування моделі каталогу генераторів інформаційного образу ВНЗ	185
<i>Лотий Л.</i> Математичні та програмні аспекти моделювання вікової структури розвитку двовидового лісу.....	193

МЕТОДИ І АЛГОРИТМИ СУЧАСНИХ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

<i>Гнатчук Є.</i> Дослідження методів формування вимог до експертних систем технічного діагностування.....	200
<i>Юрчик Г.</i> Структурно-алгоритмічний метод автокалібрування термоелектричних вимірювальних каналів температури в умовах їх експлуатації.....	205
<i>Vogucharskiy S., Mashtalir V.</i> On matrix modification of clarans clustering method in large video surveillance databases.....	211
<i>Висоцька В., Чирун Л.</i> Побудова та функціонування систем електронної контент-комерції.....	217
<i>Морозов Ю., Назаревич Р., Назаревич А., Струк Є., Мархивка В.</i> Інтернет-портал геофізичного моніторингу	228
<i>Базилевич Р., Андрієнко В.</i> Дослідження та аналіз алгоритмів острівкування енергетичних мереж.....	238
<i>Копитчук М., Передерій В.</i> Визначення нечітких ситуацій впливу навколишнього середовища на когнітивний стан особи, що приймає рішення в ергатичних системах.....	244
<i>Lytvunenko V.</i> Synthesis of rbf-network for prediction of secondary protein structure	251
<i>Строкань О.</i> Комп'ютерна система аналізу аероіонного розподілення на похилій площині.....	261
<i>Вавіленкова А.</i> Застосування формальних алгоритмів у структурній лінгвістиці	265
<i>Тимчак М., Дозорський В., Шадріна Г.</i> Компонентний метод опрацювання електрогастросигналу	272
<i>Базилевич Р., Кутельмах Р., Томчук А.</i> Розв'язування задачі комівояжера великих розмірностей методом спільних ребер.....	278

Збірник наукових праць

ВІСНИК
Національного університету
“Львівська політехніка”

Видається з 1964 р.

№ 800

**КОМП'ЮТЕРНІ НАУКИ
ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ**

Редактор *Ольга Дорошенко*

Комп'ютерне верстання *Олени Борисюк, Марти Гарасимів*

Художник-дизайнер *Маріанна Рубель-Кадирова*

Здано у видавництво 15.10.2014. Підписано до друку 17.11.2014.

Формат 60×84¹/₈. Папір офсетний. Друк на різнографі.

Умовн. друк. арк. 33,48. Обл.-вид. арк. 26,9.

Наклад 100 прим. Зам. 140850.

Видавець і виготівник: Видавництво Львівської політехніки
Свідоцтво суб'єкта видавничої справи ДК № 4459 від 27.12.2012 р.

вул. Ф. Колесси, 2, Львів, 79013
тел. +380 32 2582146, факс +380 32 2582136
vlp.com.ua, ел. пошта: vnr@vlp.com.ua