Neural Network and Agent Technologies in the Structural-Parametric Modeling of Technological Systems

Yuri A. Ivashkin[1], Ivan G. Blagoveschensky[2], and Marina A. Nikitina[3]

[1] Moscow Tecnical University Communication and Informatics, Moscow, Russia, ivashkin@nextmail.ru

[2] Bauman Moscow State Technical University, Moscow, Russia, igblagov@gmail.com

[3] V.M. Gorbatov Federal Research Center for Food Systems of Russian Academy of Sciences, Moscow, Russia nikitinama@yandex.ru

Abstract. It is offer information technology of identification and forecasting of a complex technological system based on structural and parametric modeling in combination with neural network and agent technologies. The function of the neural network module or intelligent agent is to refine the initially specified coupling coefficients between the monitored state and target parameters and to recognize abnormal situations in the system in order to make optimal decisions. The task of recognizing situations consisted in classifying them based on real-time presentation of the current states of the system by belonging to the areas of decision-making. It is offer variant of the architecture of the Hamming neural network with a multilayer recurrent structure, as a specialized heteroassociative memory device with pairs of interconnected input and output vectors.
Proposed information technology used in problems of identifying the anomalous state of technological systems of food production and making optimal decisions in the management of the quality of products of agro-processing enterprises.

Keywords: Technological system · Structural-parametric analysis · Situational analysis · Information technologies · Neural networks · Multi-agent modeling

1 Introduction

Information technologies for the structurally-parametric and situational analysis of complex chemical-technological and biotechnological systems based on the

In: S. Belim et al. (eds.): OPTA-SCL 2018, Omsk, Russia, published at http://ceur-ws.org

processing of statistical data on a managed object, in conditions of uncertainty and risk, require the development and inclusion of intelligent modules for recognizing complex situations for computer support for the adoption of optimal adequate solutions [1].

The existing direction of the structural-parametric and situational analysis of the state of the technological system [2] is related to their structuring according to the functional principle and the description of the functional relationships between the state and target parameters in the matrix form.

The structural-parametric model (SPM) of the system is represented in the form of a cellular matrix with blocks of indicators placed along the main diagonal and off-diagonal blocks, communication operators between the parameters and their functional groups. The absence of connections is described by the zero-operators $\|0\|$, which determine the non-working domain of interaction.

Initially, the characteristics of the links are determined expertly path and refined in the presence of statistical data with the determination of the correlation coefficients and linear multiple regression P_{ij} of the current deviations $\Delta x_1, ..., \Delta x_n$ of the system state variables x_i from the given norms x_j^0, depending on the deviation of the factors of the controlled set Δx_j, $j = \overline{1,n}$ $j \neq i$

$$\Delta x_i = \sum_{j=1}^{m_i} P_{ij} \Delta x_j, \quad i = \overline{1,n}; \quad j = \overline{1,n}; \quad j \neq i \tag{1}$$

Then follows the transition to the cognitive matrix $\|C_{ij}\|^n$ of relative comparable characteristics of the relationships between the different physical quantities x_i and x_j according to the formula:

$$C_{ij} = P_{ij} \frac{\Delta x_j^0}{\Delta x_i^0}; \quad i,j = \overline{1,n}; \quad j \neq i \tag{2}$$

where $\Delta x_i^0, \Delta x_j^0$ - are the admissible deviations of the variables.

The development of information technology for identifying and forecasting the state of a complex technological system in real time is associated with a rational combination of applied mathematical statistics with the analysis of fuzzy data and self-learning based on methods of artificial intelligence, neural network and agent technologies.

2 Situational Model of Technological System

On the basis of SPM, a situational matrix model of the system $C_{ij} \Delta x_j$, $i,j = \overline{1,n}$ by multiplying $\|C_{ij}\|^n$ by the diagonal matrix of the normalized deviation vector Δx_1, ... , Δx_n :

$$\begin{Vmatrix} 1 & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & 1 \end{Vmatrix} \cdot \begin{Vmatrix} \Delta x_1 & & & \\ & \Delta x_2 & & \\ \dots & \dots & \dots & \dots \\ & & & \Delta x_n \end{Vmatrix} = \begin{Vmatrix} \Delta x_1 & c_{12}\Delta x_2 & \dots & c_{1n}\Delta x_n \\ c_{21}\Delta x_1 & \Delta x_2 & \dots & c_{2n}\Delta x_n \\ \dots & \dots & \dots & \dots \\ c_{n1}\Delta x_1 & c_{n2}\Delta x_2 & \dots & \Delta x_n \end{Vmatrix} \tag{3}$$

where $\Delta x_i = \frac{|x_i - x_i{}^0|}{\Delta x_i{}^0}$; $i = \overline{1,n}$ - normalized deviations of the state parameters from the range of permissible deviations Δx_i^0.

As a result, the elements of the main diagonal of the situational matrix reflect the current normalized deviations of the x_i controlled factors from the given values $x_i{}^0$, and the off-diagonal $C_{ij} \cdot \Delta x_j;\ i,j = \overline{1,n};\ (i \neq j)$ - the contributions of the deviations $\Delta x_j;\ j = \overline{1,n}$, to the deviation $\Delta x_i;\ i = \overline{1,n}$ in accordance with the system of equations

$$\Delta x_i = \sum_{j}^{N} C_{ij} \cdot \Delta x_j;\ i,j = \overline{1,n};\ i \neq j \tag{4}$$

with ordering by rows of all a priori known causes of the deviation of Δx_i, and by columns - of the possible investigative effects of the deviation of Δx_i on other parameters.

In the general case, the situational matrix $\|c_{ij} \cdot \Delta x_j\|^n$ with a multitude of functional elements $\{x_1 ... x_n\}$ and the links between them $\|c_{ij}\|^n$ describes a structurally complex situation of cause-effect interaction of elements in the current state of the system, by combining an a **priori knowledge base** on the structure of links to current information Δx.

A formalized algorithm for identifying an abnormal situation in a technological system is as follows.

In the line of the maximum, diagonal element corresponding to the maximum deviation from the norm $\Delta x_i{}^0$ in the observed set of state parameters, the maximum nondiagonal element corresponding to the main cause that caused this deviation. Then, on the found column, need to go to the new element of the main diagonal, after which in the new line founds the main cause of the anomaly on this cause-and-effect step. The search continues until a diagonal deviation are founds, in the line of which all nondiagonal elements will be zero, which means finding the original cause of the anomalous situation.

The registration of current situations in real time complements the original database with the subsequent recalculation of regression coefficients.

However, SPM in the mode of passive observation and accumulation does not always ensure the necessary speed and accuracy of decision making in problems of identification and forecasting due to the inadequacy of statistics and the inadequacy of regression bounds. The methods, uses for the passive accumulation of data and active experiment in real time for an operating technological system are practically not realizable, because require a long period of observation or an active experiment with a sufficient number of repetitions and verification of reproducibility.

In this case, the IT technology of situational modeling of technological systems in real time under conditions of uncertainty and fuzzy data requires further intellectualization based on neural network and agent technologies.

3 Neural Network Situation Model

The intellectual function of a neural network module or self-learning agent is in refine and correct the originally defined coupling coefficients between the states and target parameters, as well as in recognize and classify anomaly situations in the system as belonging to decision-making classes based on present real-time conditions [2].

For situation analysis in conditions of fuzzy and inadequate information, a variant of architecture of the artificial neural network (ANN) Hamming with a multilayer recurrent structure is proposed as a specialized heteroassociative memory device with a predefined training sample of reference situations and associated pairs of input and output layer vectors (figure 1).

The inputs of the network receive values, n components of the current situation vector $x_1, ..., x_n$ and the network problem is to find the minimum hemming distance between the input vector and the reference vectors training samples, coded in the network structure.

First ANN layer (neurons 1-3) with unidirect propagation of its output signals to the neurons of the output layer (11-13) has fixed values of weights corresponding to the components of the vector of the observed situation (image) so that $w_{ij}^{(1)} = x_j^{(i)}$ for $i = \overline{1, p}$ (p is the number of neurons of the first layer).

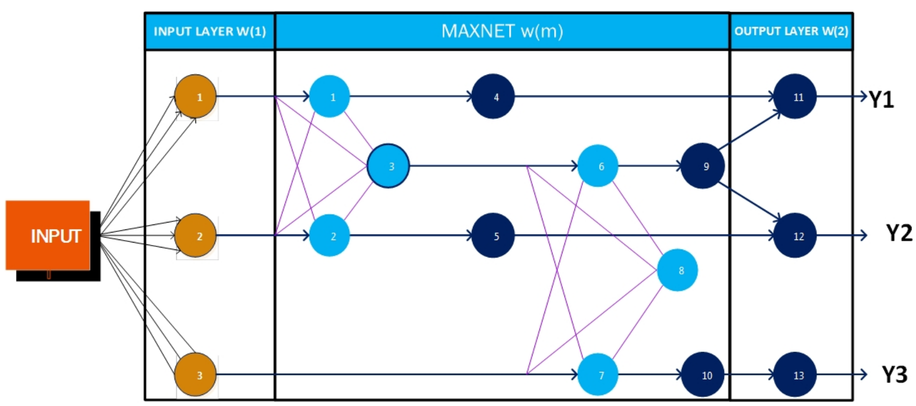

Fig. 1. The Architecture of the ANN of Hamming

Similarly, the weights of the output layer (neurons 11-13) correspond to the next vectors of reference situations $y^{(i)}$, related to $x^{(i)}$:

$$w_{ij}^{(2)} = y_j^{(i)} \tag{5}$$

The hidden layer, MAXNET, consists of neurons with feedbacks on the principle of “everyone with each”. In this case, with a proper output, the neuron is connected by a positive (exciting) feedback with a weight equal to +1, and with other neurons - negative (overwhelming) feedback with a weight inversely proportional to the number of neurons p.

Neurons of the MAXNET layer (1-3) function in WTA (Winner Takes All) mode so that the network weights should amplify the neuron's own signal and weaken the others. To achieve this effect - $w_{ii}^m = 1$, and $-\frac{1}{p-1} < w_{ij}^{(m)} < 0$ for $i \neq j$.

To ensure absolute convergence of the weight algorithm w_{ii}^m should differ from each other:

$$w_{ij}^{(m)} = -\frac{1}{p-1} + \xi \tag{6}$$

where ξ - random variable with a sufficiently small amplitude.

Neurons of different layers of ANN are function differently. Neurons of the first layer calculate the Hamming distances between fed on input N- dimensional vectors x and the vectors of the weights $w^{(i)} = x^{(i)}$ individual neurons of this layer $(i = 1, 2, ..., p)$, applied to the input of the network. The values of the output signals of these neurons are determined by the formula:

$$\widehat{y_i} = 1 - \frac{d_h(x^i, x)}{N} \tag{7}$$

where $d_H(x^i, x)$ denotes the Hamming distance between the input vectors x and $x^{(i)}$, i.e. the number of bits by which these two vectors differ.

The output signals $\widehat{y_i}$ of the neurons of the first layer become the initial states of the MAXNET layer neurons in the second phase of the network functioning. The task of the neurons of this layer is to determine the winner, i.e. a neuron, whose excitation level is closest to 1 by the **recurrence** formula:

$$y_i(k) = f(y_i(k-1) + \sum_{j \neq i} w_i^{(n)} y_j(k-1)) \tag{8}$$

at the initial value $y_j(0) = \widehat{y_i}$.

Such a neuron points to an image vector with a minimum Hamming distance to the input vector x.

The activation function $f(y)$ of neurons in the MAXNET layer is given by expression:

$$f(y) = \begin{cases} y \text{ for } y \geq 0 \\ 0 \text{ for } y < 0 \end{cases} \tag{9}$$

The iterative process terminates at a time when the state of the neurons is stabilized and the activity continues to manifest only one neuron, while the rest are in the zero state. The active neuron becomes the winner and, through the weights $w_{ij}^{(2)}$ of the linear neurons of the output layer, represents the vector $y^{(i)}$, which corresponds to the vector $x^{(i)}$, recognized by the MAXNET layer as nearest to the input vector x.

In the process of network operation, we can distinguished three phases. In the first of them an N-element vector x is fed to its input. After the presentation of this vector, the signals that define the initial states of the neurons of the second layer are generates at the outputs of the neurons of the hidden layer, i.e. MAXNET.

In the second phase, the MAXNET-initiated signals are deletes, and the iterative process (8) within this layer are starts from the initial state formed by them. The iterative process terminates at a time when all the neurons, except for the winner with an output signal equal to 1, goes to zero state. A neuron-winner with a non-zero output signal becomes a representative of the data class to which the input vector belongs.

In the third phase, the same neuron, by means of weights connecting it with the neurons of the output layer, forms a response at the output of the network in the form of a vector y, corresponding to the exciting vector x.

The data, receives from the real-time monitoring system and fed to the input of the neural network should be normalized by delta coding, with a pixelated calculation of the difference in the values of the object's state parameter in the current and previous control cycles, which can significantly reduce the dynamic range of the data.

For n parameters, the difference frame are represents as a column vector x dimensionality n:

$$X = (x_1, x_2, ..., x_j, ..., x_n)^T$$

at which on the output of the ANN will form a column-vector Y, are formed as:

$$Y = (y_1, y_2, ..., y_j, ..., y_n)^T$$

Trained network before the start of functioning with working technological parameters need to check for the quality of training and the ability to generalize the acquired knowledge. In this case, are found out, whether the results that the network gives out at the outputs are within the permissible error when the sample is fed into the network with predetermined values of the output parameters, but different from the training sample.

In the test operation of the network in the food industry of the agro-industrial complex (AIC), the minimum error in training was 1.04%, which corresponds to an allowable error of 1.5%, agreed with the technologists responsible for the quality of food products.

In a particular test implementation, the INS learning algorithm was reduced to the following sequence of actions:

1. Formation of a matrix of reference samples $\overline{\overline{X}}$ of the size $k \times n$ of the Hamming network (Table 1):

Table 1. Matrix of reference samples $\overline{\overline{X}}$ of the Hamming network

Number image	Input binary variables					
	1	2	...	i	...	n
1	x_{11}	x_{12}	...	x_{1i}	...	x_{1n}
2	x_{11}	x_{12}	...	x_{1i}	...	x_{1n}
...	...	...	...	...	...	...
j	x_{j1}	x_{j2}	...	x_{ji}	...	x_{jn}
...	...	...	...	...	...	...
p	x_{p1}	x_{p2}	...	x_{pi}	...	x_{pn}

with weight coefficients of neurons of the first layer $w_{ij} = 0.5x_{ij}$.

2. Determine the setting of the activation function of the linear threshold function:

$$f(s) = \begin{cases} 0, \text{ for } s \leq 0 \\ s, \text{ for } 0 < s \leq T \\ T, \text{ for } s \geq T \end{cases} \tag{10}$$

where $T = 0.5p$ so that the outputs of the neural network can take values within $[0, T]$.

3. Entering the synapse values of feedbacks of neurons of the hidden layer in the form of elements of a square matrix of size $p \times p$:

$$\varepsilon_{jp} = \begin{cases} 1 \text{ for } j = p \\ -\varepsilon \text{ for } j \neq p \end{cases} \tag{11}$$

where $\varepsilon \in [0, \frac{1}{p}]$, or in the matrix form:

$$\overline{\overline{E}} = \begin{cases} 1 \; for \; diagonal \; elements \\ -\varepsilon \; exclude \; diagonal \; elements \end{cases}$$

Synapses of feedbacks of a Hamming neural network with negative weights are inhibitory.

4. Setting an allowable difference of output vectors on two consecutive iterations, $E_{max} = 0.1$, for estimating the stabilization of the solution found.

The neural network algorithm for classifying situations in the observable system to support the making of managerial decisions in real time under conditions of certainty reduces to the following.

An unknown binary vector are feds on the network inputs $\overrightarrow{x}$ signals of the current state of the system parameters:

$$x_{ij} = \begin{cases} 1 \; if \; the \; parameter \; is \; within \; the \; norm \\ -1 \; if \; case \; of \; deviation \; from \; the \; norm \end{cases}$$

In the case of deviation of the state and output values of the neurons of the first layer are calculated by the formula:

$$s_{1j} = \sum_{i=1}^{n} w_{ji} x_i^* + T \tag{12}$$

The activation linear-threshold function (10) uses to calculate the outputs of the neurons of the first layer - $\overrightarrow{y}_1$. The outputs of the neurons of the second layer are assigned the values of the outputs of the neurons of the first layer, obtained at the previous step: $\overrightarrow{y}_2^{(0)} = \overrightarrow{y}_1$, after which the first layer of neurons is practically not involved.

For each q-th iteration in hidden layer calculates new values of states and outputs of neurons by recurrent ratio:

$$s_{2j}^{(q+1)} = y_{2j}^{(q)} - \varepsilon \sum_{i=1, i \neq j}^{n} y_{2i}^{(q)} \tag{13}$$

The new output values $\overrightarrow{y}_2^{(j+1)}$ are determines using the linear threshold activation function for processing the corresponding states of the neurons - $\overrightarrow{s}_2^{(j+1)}$. This cycle repeats until the output vector stabilizes in accordance with the condition:

$$\|\overrightarrow{y}^{j(q+1)} - \overrightarrow{y}^{j(q)}\| \leq E_{max} \tag{14}$$

In the ideal case, after stabilization, there is an output vector with a single positive element with the remaining zero elements, the index of which indicates the class of the unknown input image of the situation.

If the input image data is very noisy or there is no suitable standard in the training sample, several positive outputs corresponding to the accuracy condition (14) can be obtained as a result. In this case, it is concluded that the input image can not be assigned to a certain class, but the positive output indices indicate the most similar standards.

On an example of classification according to three reference situations (Table 2) in the technological system of confectionery production [8] the neural network includes 9 input variables and 3 neurons in the first and second (output) layers (figure 1).

In accordance with the learning stage algorithm, a matrix (3×9) is formed to configure the Hamming neural network for 3 reference images with 9 inputs (Table 2).

Table 2. Matrix of the Hamming network

Number image	Number input variable								
	1	2	3	4	5	6	7	8	9
1	1	−1	1	−1	1	−1	1	−1	1
2	−1	1	−1	1	1	1	−1	1	−1
3	1	1	1	1	−1	1	1	1	1

Based on the template of the reference images, calculates weighting coefficients of neurons of the first layer (Table 3).

Table 3. Matrix the weighting coefficients of neurons of the first layer

Number neuron of first layer	Number input variable								
	1	2	3	4	5	6	7	8	9
1	0.5	−0.5	0.5	−0.5	0.5	−0.5	0.5	−0.5	0.5
2	−0.5	0.5	−0.5	0.5	0.5	0.5	−0.5	0.5	−0.5
3	0.5	0.5	0.5	0.5	−0.5	0.5	0.5	0.5	0.5

If use $T = \frac{p}{2}$, we determine the threshold of the activation function T = 1.5.

With restriction are $\varepsilon \in (0, \frac{1}{3})$, the absolute value of the weight of each inhibitory synapse is $\varepsilon = 0.3$ and $E_{max} = 0.1$ and the matrix of the inverse synapse weighting coefficients (11).

$$\varepsilon_{jp} = \begin{cases} 1 \text{ for } j = i \\ -0.3 \text{ for } j \neq i \end{cases}$$

The test vector feds to the network inputs: $\overrightarrow{X}^T = [1, 1, -1, -1, 1 - 1, 1, 1, 1]$ and the condition (14) determines the column vector of the states of the neurons of the first layer, and at the output of the activation function of the state (10) is the vector-column of the output values of the neurons of the first layer:

$$\overrightarrow{s_1} = \begin{matrix} 8.00 \\ 2.00 \\ 3.00 \end{matrix} \qquad \overrightarrow{y_1} = \begin{matrix} 4.50 \\ 2.00 \\ 3.00 \end{matrix}$$

The ANN outputs assigned the corresponding output values of neurons of the first layer. Then, using the ratio (13), a series of output vectors calculates iteratively until the stabilization condition is satisfied. ANN signals obtained in iteration cycle q when the test situation feds to its inputs, represents in Table 4.

Table 4. ANN signals

number of iteration	State vector			Vector of outputs			$\|\overline{y}^{(q+1)} - \overline{y}^{(q)}\|$
	$s_{21}(q)$	$s_{22}(q)$	$s_{23}(q)$	$y_{21}(q)$	$y_{22}(q)$	$y_{23}(q)$	
1	8.00	2.00	3.00	4.50	2.00	3.00	–
2	3.00	-0.25	1.05	3.00	0.00	1.05	10.5
3	2.69	-1.22	0.15	2.69	0.00	0.15	0.91
4	2.64	-0.85	-0.66	2.64	0.00	0.00	0.02

As we can see from the table, the criterion for stopping the feedback loop of the signal after feedback mades after the 4-th iteration. The positive output value of the i-st neuron indicates that the input vector should be assigned to the i-st class.

4 Agent-based Situational Modeling of Systems

The presented neural network technology for recognizing and classifying situations in real time are suggest to use in describing the dynamics of agent behavior in complex situations. An intelligent agent is understood as [4] an imitation model of an active element capable of performing the functions assigned to it by a certain living or cybernetic organism, depending on the behavior of other agents and environmental influences.

Self-learning, purposeful agents are able to accumulate knowledge based on large amount of data and ontology of events in the process of interaction with other agents and the environment, adapt to the situation, choose a strategy for achieving the chosen goal and assess the degree of its achievement.

The general algorithm of the behavior of the intellectual agent [6] includes the identification of the situation, the assessment of one's own state and the correction of the goal, followed by a reflexive reaction or intelligent (intelligent) decision-making towards the goal. The criterion of the agent's intelligence is the degree of completeness and depth of a priori knowledge, learning strategies and decision-making algorithms under conditions of uncertainty, risk and conflict.

The parametric description of an agent includes a set of goals and a knowledge base in a specific area, a vector of characteristics of its state; bank of models and strategies of behavior, description of external relations with agents and the environment. Practical implementation of agent technologies is associated with the development of simulation systems that provide an experimentation environment, an agent-oriented language for describing models and software for organizing experiments [5,6]. The methodology of agent modeling of the learning agent reduces the following stages.

1. Parametric description of the external environment of the agent's activity with the formalization of a set of factors of influence on the functional state and objective function of the agent in situational decision-making conditions.

2. Parametric description of functional blocks of the technological system in the form of a set of vectors of input and output factors, state parameters and objective function.

3. Description of the autonomous intelligent agent with a set of state variables, input and sensory variables that communicate with other agents and the environment, as well as dynamics of agent behavior with procedures for learning and identifying current situations and making decisions in the form of discrete-event descriptions and decision-making strategies in conditions of sufficient, incomplete and fuzzy information.

4. Creation of agent-oriented model of real-time management of the technological system, which includes, in accordance with the functional scheme of the system:

- components describing the state and dynamics of agent behavior;
- organizational components that define the structure of interrelations between agents and functional blocks of the system;
- mobile components - to describe messages transmitted through a communication channel between agents and moving objects.

5. Software description of the components of the model of the system under study in a high-level algorithmic language or agent-oriented modeling language in a universal simulation system [5,6].

Agent technologies with neural network algorithms of behavior of learning agents with recognition of current situations open up new possibilities of virtual research of the influence of various technological factors on the abnormal states of the system and the adoption of optimal solutions in the control system.

5 Conclusion

The proposed direction of intellectualization of situational modeling of systems is the basis for constructing intellectual expert systems (IES) for making optimal decisions and operative management of the quality of food products at all stages of its production at processing enterprises of the agro-industrial complex [7,8].

The outlined approach to the development of IT technologies for the identification of multi-factor and weakly formalized technological systems based on artificial intelligence and agent modeling opens new possibilities for computer support for making optimal decisions in conditions of fuzzy information, uncertainty and risk.

References

1. Trakhtengerts, E.A.: Computer Support for Decision-Making. SYNTHEG (1998)

2. Ivashkin, Yu. A.: Structural-parametric modeling and identification of abnormal situations in complex technological systems. Problems of management 3, 39–43 (2004)
3. Savostin, S.D., Blagoveshchenskay, M.M., Blagoveshchensky, I.G.: Automation of Control of Flour Quality Indicators in the Grinding Process Using Intelligent Technologies: The Monograph. Publishing house of Frantera (2016)
4. Ivashkin, Yu.A.: Multiagent Modeling in the Simplex 3 Simulation System: A textbook for higher education. Binom, Laboratory of Knowledge (2016)
5. Schmidt, B.: The Art of Modeling and Simulation: The Textbook. SCS-Europe BVBA, Chent (2001)
6. Karpov, Yu.G.: Simulation of Systems. Introduction to Modeling with AnyLogic 5: The Textbook. BHV-S.-Petersburg (2005)
7. Ivashkin, Yu.A., Nikitina, M.A.: Logistics and structural optimization of material flows in network distribution systems. In: International Scientific and Practical Conference on Logistics and Economics of Resource Saving and Energy Saving in Industry (ISPC "LEREP-11-2017"). pp. 240–244. Russian Federation, Tula (2017)
8. Blagoveshchensky, I.G., Blagoveshchenskaya, M.M., Apanasenko, S.I.: Creation of virtual sensors based on the neural network for determining the main characteristics of confectionery masses. Confectionery and bakery production 11 (154), 37–41 (2014)