Запропоновано метод оптимізації класифікаційних нечітких баз знань за критеріями «точність – складність», який дозволяє спростити процес налаштування шляхом переходу до реляційної моделі. Задачу оптимізації бази знань зведено до задачі min-max кластеризації. Суть методу у виборі таких матриць розбиття «входи – вихід», які забезпечують необхідні або екстремальні рівні точності виведення та кількості правил

Ключові слова: оптимізація нечітких баз знань, min-max кластеризація, нечіткі реляційні моделі

Предложен метод оптимизации классификационных нечетких баз знаний по критериям «точность – сложность», который позволяет упростить процесс настройки путем перехода к реляционной модели. Задача оптимизации базы знаний сведена к задаче min-max кластеризации. Суть метода в выборе таких матриц разбиения «входы – выход», которые обеспечивают необходимые или экстремальные уровни точности вывода и количества правил

Ключевые слова: оптимизация нечетких баз знаний, min-max кластеризация, нечеткие реляционные модели

UDC 681.5.015:007053.81+004.91
DOI: 10.15587/1729-4061.2017.95870

OPTIMIZATION OF KNOWLEDGE BASES ON THE BASIS OF FUZZY RELATIONS BY THE CRITERIA "ACCURACY – COMPLEXITY"

H. Rakytyanska
PhD, Associate Professor
Department of software design
Vinnytsia National Technical University
Khmelnytske shose str., 95,
Vinnytsia, Ukraine, 21021
E-mail: h_rakit@ukr.net

1. Introduction

The tuning of expert fuzzy knowledge bases involves maximum approximation to experimental data for a given level of complexity or maximum simplification without losing accuracy of inference [1]. The number of output terms or classes of output [2] determines the quality of a fuzzy classification knowledge base. The optimization of such knowledge base implies: a search for the minimum inference error with the limitation to the complexity of a model (the number of input terms, output classes, and rules); search for the minimum of rules (classes) at the assigned level of accuracy. A transition to the relational model makes it possible to simplify the design process by presenting the rules in the form of a matrix of fuzzy relations "input terms – output classes" [1]. In this case, a multi-dimensional matrix of relations $\mathbf{R}(\mathbf{X})$ is presented in the form of projections $\mathbf{R}_1(x_1)$,..., $\mathbf{R}_n(x_n)$ [3]. The number of input and output terms is set in advance, and the tuning of the model implies selection of the elements of a matrix of relations [4, 5]. However, relational models leave open the problem on the optimal choice of the number of output classes. At the same time, the problem on the optimization of a fuzzy knowledge base is the task of fuzzy clustering [6]. In addition, it requires a partition of the space of input variables into such number of classes that provides the required or extreme levels of inference accuracy and the number of rules.

2. Literature review and problem statement

Methods of relational clustering, which conduct the partition of objects by similarity measures, are limited by the assigned number of classes [6, 7]. If the number of classes is unknown, the methods of min-max clustering are used, which imply the generation of easily understandable rules-hyperboxes [8]. Hyperboxes learn using supporting vector machines (SVM) [9, 10] through extension/compression. Balancing between the inference accuracy and the number of rules (classes) is achieved by combining/partition of hyperboxes. To restore nonlinear boundaries between classes and avoid excessive coverage density, the mode of learning in the min-max neural networks must reduce the number of hyperboxes without compromising the recognizing capacity [11, 12]. There remains a problem in the adaptation of maximum size of the hyperbox, which determines how many rules can be generated. Classes overlapping and classification errors render this parameter very important. If the value of this parameter is small, unnecessary hyperboxes (classes) are formed [13].

A general problem of the min-max clustering methods is the selection of the number of output classes and the minimization of the number of input terms without compromising the inference accuracy. The method for the optimization of output classes of fuzzy knowledge base was proposed in papers [14, 15]. In contrast to the heuristic procedures of rules (classes) selection [8–13], the generation of fuzzy knowledge bases is reduced to the problem on discrete optimization of indicators of algorithm reliability [14, 15]. For the selection of output classes, the gradient method was used. The number of classes is defined under the offline mode [14]. Clarification of class boundaries is carried out by adaptive adding/removing classes in arrangement vectors [15]. For the current output classes, interval rules are generated by solving the problem on inverse logical inference [2]. This solves the problem of control and adaptation of the hyperbox size [16]. The structure of the model is determined by parameters of interval rules that are connected to the coordinates of the maximum of a membership function.

© H. Rakytyanska, 2017

This paper proposes a method for the optimization of output classes and input terms of a fuzzy knowledge base. If the number of terms is set in advance, the problem of min-max clustering may be solved by relational partition of the space of input variables [1]. The number and location of hyperboxes is determined by the matrix of relations [17] and the sizes of hyperboxes are determined as a result of adjusting the triangular membership functions [1]. Then the optimization of a relational fuzzy knowledge base lies in the selection of such partition matrices "inputs – output", which provide the required or extreme levels of inference accuracy and the number of rules. Following [14, 15], the selection of number of input and output terms in the partition matrices may be performed both under the offline mode and by adaptive adding/removing of terms.

3. The aim and tasks of the study

The aim of present work is to develop an approach to the optimization designing of relational fuzzy knowledge bases by the criteria "inference accuracy – complexity". This approach should simplify the process of the knowledge bases tuning based on fuzzy relations for both the assigned and the unknown output classes.

To achieve the set goal, the following tasks were to be solved:

– development of a relational fuzzy model that matches a fuzzy classification knowledge base;

– development of a method for the optimization of knowledge base on the basis of fuzzy relations under offline and online modes.

4. Models and methods for the optimization of knowledge bases on fuzzy relations

4. 1. Fuzzy relational model

Consider an object of the form $y=f(x_1,...,x_n)$ with n inputs $\mathbf{X}=(x_1,...,x_n)$ and output y, for which the relation "inputs – output" may be represented in the form of a system of fuzzy classification IF-THEN rules [2]:

$$\bigcup_{p=\overline{1,z_j}}\Big[\bigcap_{i=\overline{1,n}}\{x_i=a_i^{jp}\}\Big]\rightarrow y=d_j,\quad j=\overline{1,m}, \qquad (1)$$

where a_i^{jp} is the fuzzy term for the evaluation of variable x_i in line jp, $j=\overline{1,m}$, $p=\overline{1,z_j}$; d_j is the fuzzy term for the evaluation of variable y; z_j is the number of rules in class d_j; m is the number of terms of the output variable.

Let $\{c_{i1},...,c_{ik_i}\}$ be a set of input terms for the evaluation of variable x_i, $i=\overline{1,n}$.

We designate

$$\{C_1,...,C_N\}=\{c_{11},...,c_{1k_1},...,c_{n1},...,c_{nk_n}\},$$

where $N=k_1+...+k_n$.

Then the system of one-dimensional matrices of fuzzy relations corresponds to a fuzzy knowledge base (1):

$$\mathbf{R}_i\subseteq c_{il}\times d_j=[r_{il,j},\ i=\overline{1,n},\ l=\overline{1,k_i},\ j=\overline{1,m}],$$

that is equivalent to a multi-dimensional matrix:

$$\mathbf{R}\subseteq C_I\times d_J=[r_{IJ},\ I=\overline{1,N},\ J=\overline{1,m}].$$

Given matrices $\mathbf{R}_i$, $i=\overline{1,n}$, dependence "inputs – output" is described using the extended compositional rule of inference [1]:

$$\boldsymbol{\mu}^d(y)=\boldsymbol{\mu}^{A_1}(x_1)\circ\mathbf{R}_1\cap...\cap\boldsymbol{\mu}^{A_n}(x_n)\circ\mathbf{R}_n, \qquad (2)$$

where $\boldsymbol{\mu}^{A_i}(x_i)=(\mu^{c_{i1}},...,\mu^{c_{ik_i}})$ and $\boldsymbol{\mu}^d(y)=(\mu^{d_1},...,\mu^{d_m})$ are the vectors of membership degrees of variables x_i and y to terms c_{il}, $i=\overline{1,n}$, and d_j, $j=\overline{1,m}$, respectively.

From ratio (2), hence follows the system of fuzzy logical equations, which connects membership functions of fuzzy input and output terms:

$$\mu^{d_j}(y)=\min_{i=\overline{1,n}}\{\max_{l=\overline{1,k_i}}[\min(\mu^{c_{il}}(x_i),r_{il,j})]\},\quad j=\overline{1,m}. \qquad (3)$$

Ratio (3) defines a fuzzy model of an object as follows:

$$y=f(\mathbf{X},N,m,\boldsymbol{\Psi}_r), \qquad (4)$$

where $\boldsymbol{\Psi}_r=(\mathbf{R},\underline{\mathbf{B}}_C,\overline{\mathbf{B}}_C,\mathbf{H}_C,\underline{\mathbf{B}}_d,\overline{\mathbf{B}}_d,\mathbf{H}_d)$ is the vector of parameters of fuzzy relations, which includes:

$$\underline{\mathbf{B}}_C=(\underline{\beta}^{C_1},...,\underline{\beta}^{C_N}),\quad \overline{\mathbf{B}}_C=(\overline{\beta}^{C_1},...,\overline{\beta}^{C_N}),\quad \mathbf{H}_C=(h^{C_1},...,h^{C_N}),$$

$$\underline{\mathbf{B}}_d=(\underline{\beta}^{d_1},...,\underline{\beta}^{d_m}),\quad \overline{\mathbf{B}}_d=(\overline{\beta}^{d_1},...,\overline{\beta}^{d_m}),\quad \mathbf{H}_d=(h^{d_1},...,h^{d_m}).$$

– vectors of lower and upper bounds, as well as vectors of coordinates of the maximum of triangular membership functions of fuzzy terms C_I and d_j; f is the operator of connection "inputs – output", which corresponds to formula (3).

4. 2. Problems on the optimization of knowledge base based on fuzzy relations

For a fuzzy knowledge base (1), the interrelation between the mean root square error and the number of rules depends on the number and bounds of output classes. Then the problem on the optimization of a fuzzy knowledge base (1) is reduced to the problem on the min-max clustering and lies in selecting such a partition matrix **R** that provides the required or extreme levels of inference accuracy and the number of rules.

Let the training sample be assigned as P pairs of experimental data:

$$\left\langle\hat{X}_s,\hat{y}_s\right\rangle,\quad s=\overline{1,P},$$

where $\hat{X}_s=(\hat{x}_1^s,...,\hat{x}_n^s)$; $\hat{y}_s$ are the vectors of values of input and output variables in the experiment number *s*.

Optimization of the number of input terms and output classes is carried out under the offline mode. In this case, the preliminary boundaries of d_j classes are assigned by an expert.

We shall evaluate the complexity of a fuzzy model (4) based on the number of rules Z(N, m, **R**), which are associated with relation matrix **R**. We shall assess the quality of a fuzzy model (4) based on the root mean square error:

$$E=\sqrt{\frac{1}{P}\sum_{s=1}^{P}\left[f(\hat{X}_s,N,m,R)-\hat{y}_s\right]^2}.$$

Then the problem of selecting the optimal number of input terms and output classes may be formulated in the direct and dual statement.

Direct statement. Find such a number of input terms N, output classes m and fuzzy partition matrix **R** that provide the minimum number of rules for a permissible inference error: $Z(N,m,\mathbf{R}) \to \min$ and $E(N,m,\mathbf{R}) \le \overline{E}$, where $\overline{E}$ is the maximum permissible root mean square error.

Dual statement. Find such a number of input terms N, output classes m and fuzzy partition matrix **R**, which provide minimum inference error for the assigned number of rules: $E(N,m,\mathbf{R}) \to \min$ and $Z(N,m,\mathbf{R}) \le \overline{Z}$, where $\overline{Z}$ is the maximum permissible number of rules.

Optimization of boundaries of output classes is performed under the *online* mode. In this case, clarification of the partition method is made by adaptive adding/removing of terms.

We shall introduce a limitation on the volume of relations matrix in the following way: $k_i \le \overline{k_i}$, $m \le \overline{m}$, where $\overline{k_i}$ and $\overline{m}$ are the maximum number of input terms and output classes.

Assume:

$$\mathbf{U} = (u_1,...,u_{\overline{N}}), \quad \mathbf{V} = (v_1,...,v_{\overline{m}}),$$

are the vectors of arrangement of input terms and output classes, where u_I=1(0) or v_J=1(0) correspond to the addition (removal) of term C_I or d_J, respectively.

We shall evaluate a complexity of fuzzy model (4) based on the number of rules **Z (U, V, R)**, which are associated with relations matrix **R**. We will assess the quality of fuzzy model (4) based on root mean square error

$$E = \sqrt{\frac{1}{P}\sum_{s=1}^{P}\left[f(\hat{X}_s,U,V,R) - \hat{y}_s\right]^2}.$$

Then the problem on the selection of optimum boundaries of output classes may be formulated in direct and dual statement.

Direct statement. Find vectors of arrangement of input terms **U**, output classes V and fuzzy partition matrix **R**, for which under condition of limitation on the knowledge base volume $Z(\mathbf{U},\mathbf{V},\mathbf{R}) \to \min$ and $E(\mathbf{U},\mathbf{V},\mathbf{R}) \le \overline{E}$.

Dual statement. Find vectors of arrangement of input terms **U**, output classes **V** and fuzzy partition matrix **R**, for which under condition of limitation on the volume of knowledge base $E(\mathbf{U},\mathbf{V},\mathbf{R}) \to \min$ and $Z(\mathbf{U},\mathbf{V},\mathbf{R}) \le \overline{Z}$.

4. 3. Method for the optimization of relational fuzzy knowledge base

To select the values of controlling variables, the gradient method is used, which was proposed in [14] for the solution of problems on discrete optimization of fuzzy knowledge base. This method implies a coordinate-wise rise along the surface of objective function in the direction of gradient. Algorithms for solving the optimization problems have a unified structure, consisting of two iteration sections [14]. In the first of them, the first permissible solution by successive adding of terms with the highest gradients is determined; in the second, an improvement of the found solution by decreasing the complexity of the model is accomplished. For the current output classes, fuzzy relations are tuned by the methods proposed in [2].

4. 3. 1. Algorithms of the optimization under offline mode

Gradients:

$$\gamma_x^i(k_i), \; i = \overline{1,n} \text{ and } \gamma_y(m),$$

will be defined as the ratio of infallibility increment $\Delta E(k_i+1, \mathbf{\Psi}_r)$ or $\Delta E(m+1, \mathbf{\Psi}_r)$ to the increment in the number of rules $\Delta Z(k_i+1, \mathbf{\Psi}_r)$ or $\Delta Z(m+1, \mathbf{\Psi}_r)$ at increasing the number of input or output terms in partition matrices:

$$\gamma_x^i(k_i) = \frac{\Delta E(k_i,\mathbf{\Psi}_r)}{\Delta Z(k_i,\mathbf{\Psi}_r)} = \frac{E(k_i,\mathbf{\Psi}_r) - E(k_i+1,\mathbf{\Psi}_r)}{Z(k_i+1,\mathbf{\Psi}_r) - Z(k_i,\mathbf{\Psi}_r)},$$

$$\gamma_y(m) = \frac{\Delta E(m,\mathbf{\Psi}_r)}{\Delta Z(m,\mathbf{\Psi}_r)} = \frac{E(m,\mathbf{\Psi}_r) - E(m+1,\mathbf{\Psi}_r)}{Z(m+1,\mathbf{\Psi}_r) - Z(m,\mathbf{\Psi}_r)}.$$

We designate the solution vector, obtained at the tth step of the optimization algorithm as:

$$\mathbf{\Psi}^{(t)} = (k_i^{(t)}, m^{(t)}, \mathbf{\Psi}_r^{(t)}).$$

The algorithm for solving the problem in direct statement is performed in the following sequence:

1. Set the zero-option of a fuzzy model:

$$t=0 \;\; \mathbf{\Psi}^{(0)} = (k_i^{(0)}, m^{(0)}, \mathbf{\Psi}_r^{(0)}).$$

If $E(\mathbf{\Psi}^{(0)}) < \overline{E}$, proceed to step 4.

2. If $E(\mathbf{\Psi}^{(t)}) > \overline{E}$, proceed to step 3, otherwise – to step 4.

3. For models

$$\mathbf{\Psi}_i' = (k_i^{(t)}+1, m^{(t)}, \mathbf{\Psi}_r') \text{ and } \mathbf{\Psi}'' = (k_i^{(t)}, m^{(t)}+1, \mathbf{\Psi}_r'')$$

identify gradients γ_x^i and γ_y relative to solution $\mathbf{\Psi}^{(t)}$. Find the coordinate, for which $\gamma = \max\{\gamma_x^i, \gamma_y\}$, t:=t+1. For vector $\mathbf{\Psi}^{(t)}$, assign:

$$k_i^{(t)} := k_i^{(t-1)} + 1, \;\; \mathbf{\Psi}^{(t)} := \mathbf{\Psi}_i', \text{ if } \gamma = \gamma_x^i;$$

$$m^{(t)} := m^{(t-1)} + 1, \;\; \mathbf{\Psi}^{(t)} := \mathbf{\Psi}'', \text{ if } \gamma = \gamma_y.$$

Proceed to step 2.

4. Decrease the complexity of model $\mathbf{\Psi}^{(t)}$ by decreasing the number of input or output terms at maintaining permissible inference accuracy. Check the conditions for models $\mathbf{\Psi}_i' = (k_i^{(t)}-1, m^{(t)}, \mathbf{\Psi}_r')$ and $\mathbf{\Psi}'' = (k_i^{(t)}, m^{(t)}-1, \mathbf{\Psi}_r'')$:

$$E(\mathbf{\Psi}_i') \le \overline{E}; \quad (5)$$

$$E(\mathbf{\Psi}'') \le \overline{E}. \quad (6)$$

If conditions (5) and (6) are not fulfilled for any coordinate, consider vector $\mathbf{\Psi}^{(t)}$ as the result of solving the problem, otherwise proceed to step 5.

5. For the coordinates that satisfy conditions (5) and (6), find the magnitude, by which the number of rules ΔZ will decrease. Find the coordinate for which:

$$\Delta = \max\{\Delta Z(k_i^{(t)}-1, m^{(t)}), \;\; \Delta Z(k_i^{(t)}, m^{(t)}-1)\}.$$

t:=t+1. For vector $\mathbf{\Psi}^{(t)}$, assign:

$$k_i^{(t)} := k_i^{(t-1)} - 1, \;\; \mathbf{\Psi}^{(t)} := \mathbf{\Psi}_i', \text{ if } \Delta = \Delta Z(k_i);$$

$$m^{(t)} := m^{(t-1)} - 1, \;\; \mathbf{\Psi}^{(t)} := \mathbf{\Psi}'', \text{ if } \Delta = \Delta Z(m).$$

Proceed to step 4.

The algorithm of solving the problem in the dual statement is performed in the following sequence.

1. Set the zero-option of a fuzzy model:

$$t:=0,\ \mathbf{\Psi}^{(0)}=(k_i^{(0)},m^{(0)},\mathbf{\Psi}_r^{(0)}).$$

If $Z(\mathbf{\Psi}^{(0)})>\overline{Z}$, proceed to step 4.

2. If $Z(\mathbf{\Psi}^{(t)})<\overline{Z}$, proceed to step 3, otherwise – to step 4.

3. The essence of this step coincides with step 3 of the algorithm for solving the problem in direct statement. Proceed to step 2.

4. Decrease the complexity of model $\mathbf{\Psi}^{(t)}$ for the inclusion in the area of permissible solutions by reducing the number of input or output terms. Check the conditions for models

$$\mathbf{\Psi}_i'=(k_i^{(t)}-1,m^{(t)},\mathbf{\Psi}_r') \text{ and } \mathbf{\Psi}''=(k_i^{(t)},m^{(t)}-1,\mathbf{\Psi}_r'')$$

$$Z(\mathbf{\Psi}_i')\le\overline{Z}; \tag{7}$$

$$Z(\mathbf{\Psi}'')\le\overline{Z}. \tag{8}$$

If at least one of the conditions (7) or (8) is fulfilled, then, among permissible solutions, select a model that provides a lower inference error, otherwise proceed to step 5.

5. For the coordinates that do not satisfy limitations (7) and (8), find the increment in deriving error ΔE. Find the coordinate, for which

$$\Delta=\min\{\Delta E(k_i^{(t)}-1,m^{(t)}),\ \ \Delta E(k_i^{(t)},m^{(t)}-1)\}.$$

t:=t+1. For vector $\mathbf{\Psi}^{(t)}$, assign:

$$k_i^{(t)}:=k_i^{(t-1)}-1,\ \ \mathbf{\Psi}^{(t)}:=\mathbf{\Psi}_i',\ \text{if } \Delta=\Delta E(k_i);$$

$$m^{(t)}:=m^{(t-1)}-1,\ \ \mathbf{\Psi}^{(t)}:=\mathbf{\Psi}'',\ \text{if } \Delta=\Delta E(m).$$

Proceed to step 4.

4. 3. 2. Algorithms of optimization under the online mode

Gradients

$$\gamma_x^I(u_I),\ \ I=\overline{1,N} \text{ and } \gamma_y^J(v_J),\ \ J=\overline{1,m},$$

will be defined as the ratio of infallibility increment $\Delta E(u_I=1, \mathbf{\Psi}_r)$ or $\Delta E(v_J=1, \mathbf{\Psi}_r)$ to the increment in the number of rules $\Delta Z(u_I=1, \mathbf{\Psi}_r)$ or $\Delta Z(v_J=1, \mathbf{\Psi}_r)$ as a result of adding the input or output term C_I or d_J:

$$\gamma_x^I(u_I)=\frac{\Delta E(u_I,\mathbf{\Psi}_r)}{\Delta Z(u_I,\mathbf{\Psi}_r)}=\frac{E(u_I=0,\mathbf{\Psi}_r)-E(u_I=1,\mathbf{\Psi}_r)}{Z(u_I=1,\mathbf{\Psi}_r)-Z(u_I=0,\mathbf{\Psi}_r)},$$

$$\gamma_y^J(v_J)=\frac{\Delta E(v_J,\mathbf{\Psi}_r)}{\Delta Z(v_J,\mathbf{\Psi}_r)}=\frac{E(v_J=0,\mathbf{\Psi}_r)-E(v_J=1,\mathbf{\Psi}_r)}{Z(v_J=1,\mathbf{\Psi}_r)-Z(v_J=0,\mathbf{\Psi}_r)}.$$

Designate the solution vector, obtained at the t-th step of the optimization algorithm as $\mathbf{\Psi}^{(t)}=(\mathbf{U}^{(t)},\mathbf{V}^{(t)},\mathbf{\Psi}_r^{(t)})$. The algorithm of solving the problem in direct statement is performed in the following sequence.

1. Assign the zero-option of a fuzzy model:

$$t:=0,\ \mathbf{\Psi}^{(0)}=(\mathbf{U}^{(0)},\mathbf{V}^{(0)},\mathbf{\Psi}_r^{(0)}).$$

If $E(\mathbf{\Psi}^{(0)})<\overline{E}$, proceed to step 4.

2. If $E(\mathbf{\Psi}^{(t)})>\overline{E}$, proceed to step 3, otherwise – to step 4.

3. For the models where $u_I^{(t)}=0$ and $v_J^{(t)}=0$, add an input or output term as follows:

$$\mathbf{\Psi}_I'=(u_I^{(t)}+1,\mathbf{V}^{(t)},\mathbf{\Psi}_r^I) \text{ or } \mathbf{\Psi}_J''=(\mathbf{U}^{(t)},v_J^{(t)}+1,\mathbf{\Psi}_r^J).$$

Determine gradients $\gamma_x^I(u_I)$ and $\gamma_y^J(v_J)$ relative to solution $\mathbf{\Psi}^{(t)}$. Find the term, for which $\tilde{\gamma}=\max\{\gamma_x^L,\gamma_y^M\}$, where:

$$\gamma_x^L(u_L^{(t)})=\max_I\{\gamma_x^I\},$$

$$\gamma_y^M(v_M^{(t)})=\max_J\{\gamma_y^J\}.$$

t:=t+1. For vector $\mathbf{\Psi}^{(t)}$, assign:

$$u_L^{(t)}=1,\ \ \mathbf{\Psi}^{(t)}:=\mathbf{\Psi}_L',\ \text{if } \tilde{\gamma}=\gamma_x^L;$$

$$v_M^{(t)}=1,\ \ \mathbf{\Psi}^{(t)}:=\mathbf{\Psi}_M'',\ \text{if } \tilde{\gamma}=\gamma_y^M.$$

Proceed to step 2.

4. Improve model $\mathbf{\Psi}^{(t)}$ by attaining the required level of inference accuracy with fewer terms. For the models for which $u_I^{(t)}=1$ and $v_J^{(t)}=1$, decrease the complexity by reducing the number of terms in the following way:

$$\mathbf{\Psi}_I'=(u_I^{(t)}-1,\mathbf{V}^{(t)},\mathbf{\Psi}_r^I);\ \ \mathbf{\Psi}_J''=(\mathbf{U}^{(t)},v_J^{(t)}-1,\mathbf{\Psi}_r^J).$$

For the inputs and outputs, find such sets of terms $Q_x^{(t)}$ and $Q_y^{(t)}$, for which the conditions are fulfilled:

$$E(\mathbf{\Psi}_I')\le\overline{E}; \tag{9}$$

$$E(\mathbf{\Psi}_J'')\le\overline{E}. \tag{10}$$

If $Q_x^{(t)}$ and $Q_y^{(t)}$ are empty sets, consider vector $\mathbf{\Psi}^{(t)}$ as the result of solving the problem, otherwise proceed to step 5.

5. For terms $C_I\in Q_x^{(t)}$ and $d_J\in Q_y^{(t)}$, which satisfy conditions (9) and (10), find the magnitude, by which the number of rules ΔZ decreased. Find the term, for which

$$\tilde{\Delta}Z=\max\{\Delta Z_x^L,\Delta Z_y^M\},$$

where

$$\Delta Z_x^L(u_L^{(t)})=\max_I\{\Delta Z(u_I^{(t)}-1)\};$$

$$\Delta Z_y^M(v_M^{(t)})=\max_J\{\Delta Z(v_J^{(t)}-1)\}.$$

t:=t+1. For vector $\mathbf{\Psi}^{(t)}$, assign:

$$u_L^{(t)}=0,\ \ \mathbf{\Psi}^{(t)}:=\mathbf{\Psi}_L',\ \text{if } \tilde{\Delta}Z=\Delta Z_x^L;$$

$$v_M^{(t)}=0,\ \ \mathbf{\Psi}^{(t)}:=\mathbf{\Psi}_M'',\ \text{if } \tilde{\Delta}Z=\Delta Z_y^M.$$

Proceed to step 4.

The algorithm of solving the problem in the dual statement is performed in the following sequence.

1. Set the zero-option of a fuzzy model:

$$t:=0,\ \mathbf{\Psi}^{(0)}=(\mathbf{U}^{(0)},\mathbf{V}^{(0)},\mathbf{\Psi}_r^{(0)}).$$

If $Z(\mathbf{\Psi}^{(0)})>\overline{Z}$, proceed to step 4.

2. If $Z(\mathbf{\Psi}^{(t)})<\overline{Z}$, proceed to step 3, otherwise – to step 4.

3. The essence of this step coincides with step 3 of the algorithm for solving the problem in direct statement. Proceed to step 2.

4. Decrease the complexity of model $\mathbf{\Psi}^{(t)}$ for the inclusion in the area of permissible solutions. For models, in which $u_I^{(t)}=1$ and $v_J^{(t)}=1$, decrease the number of terms in the following way:

$$\mathbf{\Psi}_I'=(u_I^{(t)}-1,\mathbf{V}^{(t)},\mathbf{\Psi}_r^I);\quad \mathbf{\Psi}_J''=(\mathbf{U}^{(t)},v_J^{(t)}-1,\mathbf{\Psi}_r^J).$$

For the inputs and outputs, find such sets of terms $Q_x^{(t)}$ and $Q_y^{(t)}$, for which the conditions are satisfied:

$$Z(\mathbf{\Psi}_I')>\overline{Z}; \tag{11}$$

$$Z(\mathbf{\Psi}_J'')>\overline{Z}. \tag{12}$$

If at least one of conditions (11) or (12) is not met, then choose among permissible solutions a model that provides a lower inference error, otherwise proceed to step 5.

5. For terms $C_I\in Q_x^{(t)}$ and $d_J\in Q_y^{(t)}$, which satisfy conditions (11) and (12), find the magnitude, by which the inference error ΔE increases. Find the term, for which

$$\tilde{\Delta}E=\min\{\Delta E_x^L,\Delta E_y^M\},$$

where

$$\Delta E_x^L(u_L^{(t)})=\min_I\{\Delta E(u_I^{(t)}-1)\};$$

$$\Delta E_y^M(v_M^{(t)})=\min_J\{\Delta E(v_J^{(t)}-1)\}.$$

t:=t+1. For vector $\mathbf{\Psi}^{(t)}$, assign:

$$u_L^{(t)}=0,\ \ \mathbf{\Psi}^{(t)}:=\mathbf{\Psi}_L',\ \text{if}\ \tilde{\Delta}E=\Delta E_x^L;$$

$$v_J^{(t)}=0,\ \ \mathbf{\Psi}^{(t)}:=\mathbf{\Psi}_M'',\ \text{if}\ \tilde{\Delta}E=\Delta E_y^M.$$

Proceed to step 4.

5. Results of computer experiment

For the model-standard [15, 16], the number of terms is limited as follows:

$$\overline{k}_1=9,\ \ \overline{k}_2=7,\ \ \overline{m}=7.$$

The task implied the transformation of the expert zero-option of a knowledge base to the variant, which provides: $Z\to\min$ and $\overline{E}\le 0.5$ in the direct statement; $E\to\min$ and $\overline{Z}\le 30$ in the dual statement.

Results of the calculation of optimization problems are listed in Table 1, where each iteration represents the results of designing model $\mathbf{\Psi}^{(t)}$ for the current number of terms $k_i^{(t)}$ and $m^{(t)}$ with further arrangement of terms in vectors $\mathbf{U}^{(t)}$ and $\mathbf{V}^{(t)}$.

The first acceptable solution of the direct problem is obtained at step 4 by successive adding of terms with the highest gradients:

– term c_{24} at step 2 since:

$$\gamma_y(v_3)=\frac{0.6712-0.6104}{16-12}=0.0152,$$

$$\gamma_x^1(u_{12},u_{18})=\frac{0.6712-0.6380}{14-12}=0.0166,$$

$$\gamma_x^2(u_{24})=\frac{0.6712-0.5968}{15-12}=0.0248;$$

– terms c_{12} and c_{18} at step 3 since:

$$\gamma_y(v_3)=\frac{0.5968-0.5575}{19-15}=0.0098,$$

$$\gamma_x^1(u_{12},u_{18})=\frac{0.5968-0.5310}{17-15}=0.0329,$$

$$\gamma_x^2(u_{23})=\frac{0.5968-0.5632}{18-15}=0.0112;$$

– class d_3 at step 4 since:

$$\gamma_y(v_3)=\frac{0.5310-0.4873}{21-17}=0.0109,$$

$$\gamma_x^1(u_{13},u_{17})=\frac{0.5310-0.5250}{21-17}=0.0015,$$

$$\gamma_x^2(u_{23})=\frac{0.5310-0.5101}{20-17}=0.0070.$$

Table 1

Calculation of optimization problems

t	k_1	k_2	m	$u_{11},\ldots,u_{19}$	$u_{21},\ldots,u_{27}$	$v_1,\ldots,v_7$	Z	E
1	5	4	5	100111001	1100110	1101011	12	0.6712
2	5	5	5	100111001	1101110	1101011	15	0.5968
3	7	5	5	110111011	1101110	1101011	17	0.5310
4	7	5	6	110111011	1101110	1111011	21	0.4873
5	5	5	6	100111001	1101110	1111011	19	0.5575
6	7	6	6	110111011	1111110	1111011	22	0.4625
7	7	6	7	110111011	1111110	1111111	24	0.4318
8	9	6	7	111111111	1111110	1111111	28	0.3514
9	9	7	7	111111111	1111111	1111111	31	0.3007
10	7	7	7	110111011	1111111	1111111	27	0.3819

Model $\mathbf{\Psi}^{(4)}$ remains the solution of the direct problem. Decreasing the complexity leads to model $\mathbf{\Psi}^{(5)}$ leaving the region of permissible solutions. Further increase in the number of terms in model $\mathbf{\Psi}^{(6)}$ provides decreasing the inference error by $\Delta E=0.0248$ with increasing the number of rules by $\Delta Z=1$.

Solution of the dual problem was continued by adding terms with the highest gradients:

– term c_{23} at step 6 since:

$$\gamma_y(v_5)=\frac{0.4873-0.4509}{23-21}=0.0182,$$

$$\gamma_x^1(u_{13},u_{17})=\frac{0.4873-0.4716}{23-21}=0.0079,$$

$$\gamma_x^2(u_{23})=\frac{0.4873-0.4625}{22-21}=0.0248;$$

– class d_5 at step 7 since:

$$\gamma_y(v_5)=\frac{0.4625-0.4318}{24-22}=0.0153,$$

$$\gamma_x^1(u_{13},u_{17})=\frac{0.4625-0.4437}{26-22}=0.0047,$$

$$\gamma_x^2(u_{27})=\frac{0.4625-0.4560}{25-22}=0.0022;$$

– terms c_{13} and c_{17} at step 8 since:

$$\gamma_x^1(u_{13},u_{17})=\frac{0.4318-0.3514}{28-24}=0.0201,$$

$$\gamma_x^2(u_{27})=\frac{0.4318-0.3819}{27-24}=0.0166.$$

Model $\mathbf{\Psi}^{(8)}$ remains the solution of the dual problem. Further increase in the number of terms leads to model $\mathbf{\Psi}^{(9)}$ leaving the region of permissible solutions, and decreasing the complexity of model $\mathbf{\Psi}^{(10)}$ – to increasing the inference error by ΔE=0.0305 at decreasing the number of rules by ΔZ=1.

Matrices of fuzzy relations in the solutions of direct and dual tasks (Tables 2, 3) are associated with fuzzy rule bases, presented in Tables 4, 5. Results of structural and parametric tuning of models $\mathbf{\Psi}^{(4)}$ and $\mathbf{\Psi}^{(8)}$ are shown in Fig. 1, 2.

Table 2

Matrix of fuzzy relations for a direct problem

IF		THEN y					
		d_1	d_2	d_3	d_4	d_6	d_7
x_1	c_{11}	0	0.65	0.72	0.92	0	0
	c_{12}	0.74	0	0.65	0	0	0
	c_{14}	0.50	0.84	0	0.69	0.57	0
	c_{15}	0.72	0.96	0	0	0.92	1.00
	c_{16}	0.53	0.82	0	0.70	0.57	0
	c_{18}	0.77	0	0.63	0	0	0
	c_{19}	0	0.67	0.71	0.95	0	0
x_2	c_{21}	0	0.82	0.71	0.91	0	0
	c_{22}	0.74	0	0.73	0	0	0
	c_{24}	0.75	0.63	0	0.63	0.94	0
	c_{25}	0.64	0.52	0	0	0.87	1.00
	c_{26}	0.50	0.61	0	0.61	0.92	0

Table 3

Matrix of fuzzy relations for a dual problem

IF		THEN y						
		d_1	d_2	d_3	d_4	d_5	d_6	d_7
x_1	c_{11}	0	0.76	0.82	0.95	0	0	0
	c_{12}	0.82	0	0.65	0	0	0	0
	c_{13}	0.51	0.70	0	0.70	0.70	0	0
	c_{14}	0.52	0.64	0	0	0.83	0.92	0
	c_{15}	0.64	0.64	0	0	0.71	0.86	1.00
	c_{16}	0.50	0.63	0	0	0.85	0.93	0
	c_{17}	0.52	0.70	0	0.70	0.70	0	0
	c_{18}	0.81	0	0.64	0	0	0	0
	c_{19}	0	0.75	0.83	0.94	0	0	0
x_2	c_{21}	0	0.82	0.71	0.93	0	0	0
	c_{22}	0.78	0	0.72	0	0	0	0
	c_{23}	0.77	0.80	0	0.70	0.86	0	0
	c_{24}	0.73	0.64	0	0	0.83	0.76	0
	c_{25}	0.61	0.52	0	0	0.77	0.89	1.00
	c_{26}	0.50	0.80	0	0.70	0.70	0.70	0

Table 4

IF-THEN rules for a direct problem

Rule	IF		THEN
	x_1	x_2	y
1, 2 3	[0.61, 2.03] or [3.90, 5.45] [1.75, 4.12]	[0.60, 4.00] [0.60, 1.97]	d_1, [−0.25, 0.39]
4, 5 6	[0, 0.86] or [5.17, 6.00] [1.75, 4.12]	[1.68, 4.00] [0, 0.87]	d_2, [0.32, 0.80]
7, 8 9, 10	[0, 0.86] or [5.17, 6.00] [0.61, 2.03] or [3.90, 5.45]	[0.60, 1.97] [0, 0.87]	d_3, [0.73, 1.25]
11, 12 13, 14 15, 16	[0, 0.86] or [5.17, 6.00] [1.75, 2.72] or [3.30, 4.12] [1.75, 2.72] or [3.30, 4.12]	[0, 0.87] [1.68, 2.70] [3.32, 4.00]	d_4, [1.10, 1.97]
17, 18 19 20	[1.75, 2.72] or [3.30, 4.12] [2.54, 3.45] [2.54, 3.45]	[2.56, 3.50] [1.68, 2.70] [3.32, 4.00]	d_6, [1.85, 2.59]
21	[2.54, 3.45]	[2.56, 3.50]	d_7, [2.41, 3.20]

Table 5

IF-THEN rules for a dual problem

Rule	IF		THEN
	x_1	x_2	y
1, 2 3	[0.52, 1.46] or [4.68, 5.54] [1.32, 4.82]	[0.56, 4.00] [0.56, 1.42]	d_1, [−0.18, 0.34]
4, 5 6	[0, 0.64] or [5.42, 6.00] [1.32, 4.82]	[1.30, 4.00] [0, 0.68]	d_2, [0.26, 0.71]
7, 8 9, 10	[0, 0.64] or [5.42, 6.00] [0.52, 1.46] or [4.68, 5.54]	[0.56, 1.42] [0, 0.68]	d_3, [0.64, 1.00]
11, 12 13, 14 15, 16	[0, 0.64] or [5.42, 6.00] [1.32, 2.24] or [3.85, 4.82] [1.32, 2.24] or [3.85, 4.82]	[0, 0.68] [1.30, 2.18] [3.37, 4.00]	d_4, [0.92, 1.63]
17, 18 19, 20 21, 22 23	[1.32, 2.24] or [3.85, 4.82] [2.12, 2.76] or [3.34, 3.97] [2.12, 2.76] or [3.34, 3.97] [2.12, 3.97]	[2.06, 3.45] [2.06, 2.72] [3.37, 4.00] [1.30, 2.18]	d_5, [1.50, 2.25]
24, 25 26 27	[2.12, 2.76] or [3.34, 3.97] [2.64, 3.48] [2.64, 3.48]	[2.61, 3.45] [2.06, 2.72] [3.37, 4.00]	d_6, [2.11, 2.78]
28	[2.64, 3.48]	[2.61, 3.45]	d_7, [2.65, 3.37]

For the solutions of a direct and a dual problem, the compromise "inference accuracy – complexity" is achieved by adding/removing output class d_5 and input terms c_{13}, c_{17} and c_{23}.

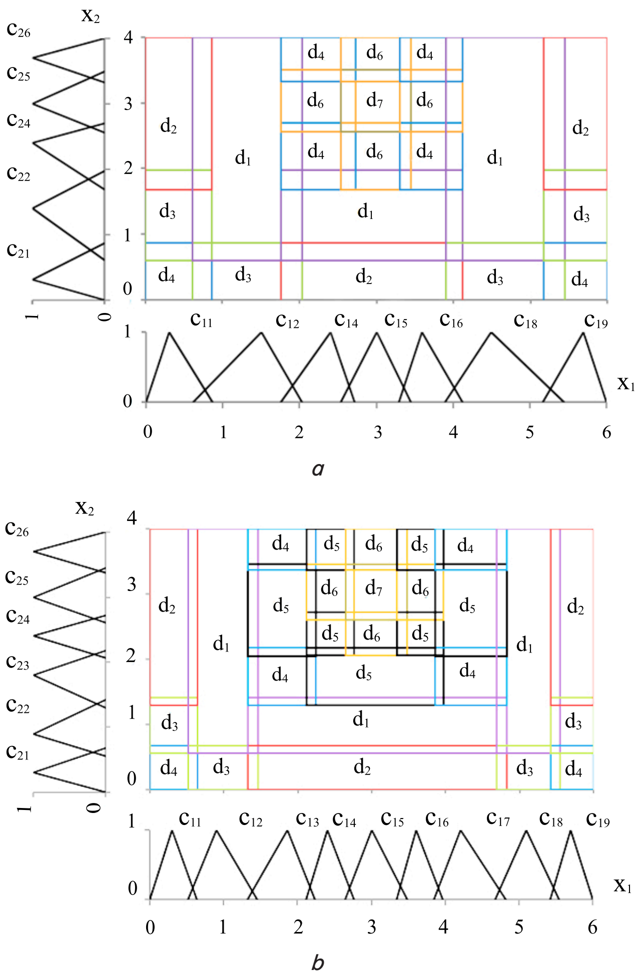

Fig. 1. Results of the structural tuning for solving: *a* — direct problem; *b* — dual problem; ____ — d_1; ____ — d_2; ____ — d_3; ____ — d_4; ____ — d_5; ____ — d_6; ____ — d_7

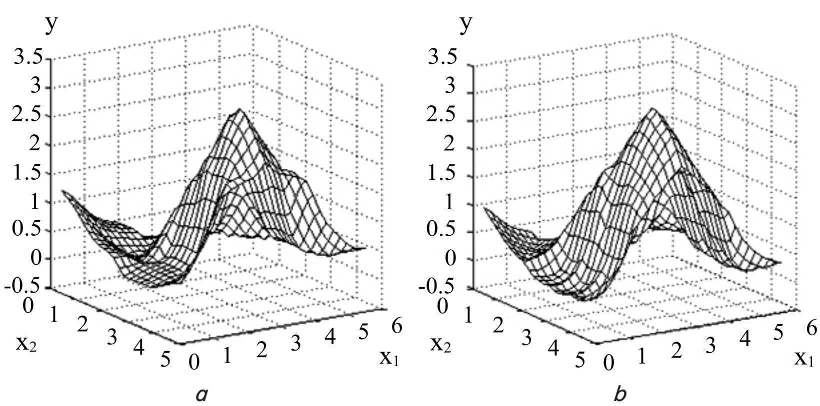

Fig. 2. Results of parametric tuning for solving: *a* — direct problem; *b* — dual problem

6. Discussion of results of assessing the complexity of tuning algorithms for a fuzzy classification knowledge base

The proposed method, as well as methods [14, 15], represents the formalization of improving transformations for an expert fuzzy knowledge base. At the same time, controlling variables are set, which are the number of input terms, output classes and rules. Improving transformations make it possible to formalize the process of generation of fuzzy knowledge base variants with a subsequent selection by the criteria of accuracy and costs or by the complexity of the tuning process.

Assume the number of rules (classes) is limited, and the number of input terms is unknown. Then the number of tuning parameters for the fuzzy classification knowledge base is 2nZ+2m for two-parameter membership functions [2] or upper and lower boundaries of interval rules [8–13]. Assume that in addition to the number of output classes and rules, the number of input terms is also limited. Then relations matrices are implanted into the antecedents of fuzzy rules, and the number of tuning parameters of a relational fuzzy knowledge base is $ZNm+2N+2m$ [4, 5].

If the number of rules (classes) is subjected to minimization, we limit the number of terms of input N_T and output M_T whose linguistic modification provides the required inference accuracy [14–16]. The number of tuning parameters of the rules generator based on the matrix of fuzzy relations is $N_T M_T+2N_T+2M_T$. An inverse inference for m output terms requires the solution of Z optimization problems with 2n variables for the upper and lower boundaries of the intervals [16].

Compared with [2, 4, 5, 8–13, 14–16], the proposed method allows us to decrease the number of tuning parameters to $Nm+2N+2m$ for partition matrices and the upper and lower boundaries of triangular membership functions. The shortcoming of the method is the necessity of obtaining linguistic IF-THEN rules, which are associated with a fuzzy partition matrix.

7. Conclusions

1. The models and methods were developed for the optimization design of fuzzy classification knowledge bases by the criteria "inference accuracy – complexity". A fuzzy relational model, which corresponds to a fuzzy classification knowledge base, was proposed. The problem on the optimization of a fuzzy knowledge base is reduced to the problem on the min-max clustering and comes down to selecting such partition matrices "inputs – output" that provide the required or extreme levels of accuracy and the number of rules.

2. The selection of output classes and input terms is reduced to the problem on discrete optimization of the algorithm reliability indicators, for the solution of which we employed the gradient method. This resolves a general problem in the methods of min-max clustering related to the selection of the number of output classes and minimization of the number of input terms without losing inference accuracy. The number and location of hyperboxes are determined by the relation matrix "input terms – output classes", and the sizes of hyperboxes are defined as a result of tuning of the triangular membership functions. Selection of the number of input and output terms in partition matrices may be performed both under the offline mode and by adaptive adding/removing of terms. A transition to the relational fuzzy model provides the simplification of the process of knowledge bases tuning both for the assigned and unknown output classes.

References

1. Yager, R. Essentials of fuzzy modeling and control [Text] / R. Yager, D. Filev. – New York: John Willey & Sons, 1994. – 408 p.
2. Rotshtein, A. P. Fuzzy Evidence in Identification, Forecasting and Diagnosis [Text] / A. P. Rotshtein, H. B. Rakytyanska. – Heidelberg: Springer, 2012. – 314 p. doi: 10.1007/978-3-642-25786-5
3. Mandal, S. SISO fuzzy relational inference systems based on fuzzy implications are universal approximators [Text] / S. Mandal, B. Jayaram // Fuzzy Sets and Systems. – 2015. – Vol. 277. – P. 1–21. doi: 10.1016/j.fss.2014.10.003
4. Scherer, R. Relational Modular Fuzzy Systems [Text] / R. Scherer // Studies in Fuzziness and Soft Computing. – Springer Berlin Heidelberg, 2012. – P. 39–50. doi: 10.1007/978-3-642-30604-4_4
5. Gonzalez, A. An Efficient Inductive Genetic Learning Algorithm for Fuzzy Relational Rules [Text] / A. Gonzalez, R. Perez, Y. Caises, E. Leyva // International Journal of Computational Intelligence Systems. – 2012. – Vol. 5, Issue 2. – P. 212–230. doi: 10.1080/18756891.2012.685265
6. Graves, D. Clustering with proximity knowledge and relational knowledge [Text] / D. Graves, J. Noppen, W. Pedrycz // Pattern Recognition. – 2012. – Vol. 45, Issue 7. – P. 2633–2644. doi: 10.1016/j.patcog.2011.12.019
7. De Carvalho, F. de A. T. Relational partitioning fuzzy clustering algorithms based on multiple dissimilarity matrices [Text] / F. de A. T. de Carvalho, Y. Lechevallier, F. M. de Melo // Fuzzy Sets and Systems. – 2013. – Vol. 215. – P. 1–28. doi: 10.1016/j.fss.2012.09.011
8. Bargiela, A. Optimised Information Abstraction in Granular Min/Max Clustering [Text] / A. Bargiela, W. Pedrycz // Smart Innovation, Systems and Technologies. – 2013. – P. 31–48. doi: 10.1007/978-3-642-28699-5_3
9. Gaspar, P. Parameter Influence in Genetic Algorithm Optimization of Support Vector Machines [Text] / P. Gaspar, J. Carbonell, J. L. Oliveira // Advances in Intelligent and Soft Computing. – 2012. – P. 43–51. doi: 10.1007/978-3-642-28839-5_5
10. Wu, Z. A fuzzy support vector machine algorithm for classification based on a novel PIM fuzzy clustering method [Text] / Z. Wu, H. Zhang, J. Liu // Neurocomputing. – 2014. – Vol. 125. – P. 119–124. doi: 10.1016/j.neucom.2012.07.049
11. Mohammed, M. F. Improving the Fuzzy Min-Max neural network with a K-nearest hyperbox expansion rule for pattern classification [Text] / M. F. Mohammed, C. P. Lim // Applied Soft Computing. – 2017. – Vol. 52. – P. 135–145. doi: 10.1016/j.asoc.2016.12.001
12. Seera, M. A modified fuzzy min-max neural network for data clustering and its application to power quality monitoring [Text] / M. Seera, C. P. Lim, C. K. Loo, H. Singh // Applied Soft Computing. – 2015. – Vol. 28. – P. 19–29. doi: 10.1016/j.asoc.2014.09.050
13. Reyes-Galaviz, O. F. Granular fuzzy modeling with evolving hyperboxes in multi-dimensional space of numerical data [Text] / O. F. Reyes-Galaviz, W. Pedrycz // Neurocomputing. – 2015. – Vol. 168. – P. 240–253. doi: 10.1016/j.neucom.2015.05.102
14. Rotshtein, A. P. Optimal design of rule-based systems by solving fuzzy relational equations [Text] / A. P. Rotshtein, H. B. Rakytyanska // Studies in Computational Intelligence. – 2014. – P. 167–178. doi: 10.1007/978-3-319-06883-1_14
15. Rakytyanska, H. Optimization of composite fuzzy knowledge bases on rules and relations [Text] / H. Rakytyanska // Inf. Technol. Comput. Eng. – 2015. – Issue 1. – P. 17–26.
16. Rakytyanska, H. Fuzzy classification knowledge base construction based on trend rules and inverse inference [Text] / H. Rakytyanska // Eastern-European Journal of Enterprise Technologies. – 2015. – Vol. 1, Issue 3 (73). – P. 25–32. doi: 10.15587/1729-4061.2015.36934
17. Rotshtein, A. P. Fuzzy Genetic Object Identification: Multiple Inputs/Multiple Outputs Case [Text] / A. P. Rotshtein, H. B. Rakytyanska // Advances in Intelligent and Soft Computing. – 2012. – P. 375–394. doi: 10.1007/978-3-642-23172-8_25