Запропоновано метод побудови нечіткої логічної системи типу 2, який дозволяє в процесі ідентифікації багатовимірних об'єктів, що слабо формалізуються, змінювати кількість входів без втрати системою властивості адекватного відображення предметної області. Для цього запропоновано критерій оптимізації, за яким збільшується кількість взаємної інформації, що відображається зі входів системи на її виходи

Ключові слова: нечітка логічна система, нечіткі множини типу 2, інтервальна функція належності, критерій оптимізації

Предложен метод построения нечеткой логической системы типа 2, позволяющий в процессе идентификации слабо формализованных многомерных объектов изменять количество входов без потери системой свойства адекватно отображать предметную область. Для этого предложено использование критерия оптимизации, с помощью которого максимизируется количество взаимной информации, отображаемой со входов системы на её выходы

Ключевые слова: нечеткая логическая система, нечеткие множества типа 2, интервальная функция принадлежности, критерий оптимизации

UDC 004.891.2
DOI: 10.15587/1729-4061.2017.101635

A METHOD OF BUILDING TYPE-2 FUZZY LOGIC SYSTEMS IN MULTIDIMENSIONAL OBJECTS IDENTIFICATION PROBLEMS

N. Kondratenko
PhD, Associate Professor*
E-mail: kondrn2014@gmail.com
O. Snihur
Postgraduate student*
E-mail: olha.snihur@gmail.com
*Department of information security
Vinnytsia National Technical University
Khmelnytske highway, 95, Vinnytsia, Ukraine, 21021

1. Introduction

The use of new information technologies requires developing decision support systems, the principal purpose of which are applications in identification problems, arising in science, medicine, engineering, industry etc. As stated in [1], multidimensional objects identification in many areas implies creating relation operators between input and output signals functions based on the information available to the developer. When modeling such objects, traditional method developers focus on enhancing the adequacy of modeled objects descriptions and accounting for more of the factors that affect decision-making processes. All of this results in creating bulky mathematical constructs, which are impossible to utilize in practice. Besides, there are a lot of objects, for which most of the input and output parameters cannot be described numerically. They are called partially formalized objects. Presently one of the more developed and common scientific directions for describing uncertainties for multidimensional partially formalized objects is fuzzy set theory and fuzzy modeling [2, 3].

Fuzzy models utilize the basic concepts of fuzzy set theory, fuzzy rules and fuzzy inference for their operation. Depending on the fuzziness degree of the sets considered while developing a fuzzy model, type-1, generalized type-2 and interval type-2 fuzzy models are distinguished. Type-1 fuzzy models are developed based on fuzzy sets with crisp membership grades values. Yet such fuzzy models provide only crisp (point) values at their outputs. A value of a generalized type-2 fuzzy set membership function is a type-1 fuzzy set. Applications of such models are extremely rare due to high computation cost. Instead interval fuzzy sets are used as a simplified presentation of generalized sets. Mathematical operations on interval type-2 fuzzy sets deal with marginal points of the membership function value interval only and disregard specific qualities of the distribution existing within this interval. Interval type-2 fuzzy models are developed based on fuzzy sets with interval membership grades values. Unlike type-1 fuzzy models, these models provide both point and interval values at their outputs and process different kinds of uncertainties efficiently enough [4]. Using fuzzy set theory techniques for knowledge formalization automatically makes a researcher face the problem of choosing the right fuzzy set type for building the fuzzy model.

Based on the above considerations, the question on analyzing fuzzy systems regarding their capabilities to reflect the subject area in multidimensional objects identification tasks is of current interest.

2. Literature review and problem statement

The modern approach to solving complicated multidimensional partially formalized objects identification tasks is associated with fuzzy logic systems (FLS) [5, 6]. For instance, such problems as diagnostics of a financial condition of a company [7] and transport logistics [8] are successfully solved in terms of fuzzy values.

Developing a type-1 fuzzy set based FLS is common in medical diagnostics [9, 10]. Such systems are usually multidimensional, and always include a membership functions parameters tuning stage in order to ensure adequate representation of the subject area by the model. Various random search based optimization methods are used for tuning, usually genetic algorithms [11–13].

Applications of genetic algorithms for complicated optimization tasks shows that these methods have advantages over traditional ones. In [14], a genetic algorithm for tuning a fuzzy model in medical diagnostics is considered. Its performance is being evaluated against a series of test functions and multidimensional tasks with a high number of local extremes. A genetic algorithm for training a type-2 fuzzy model with multiple inputs and outputs is introduced in [14]. The proposed algorithm uses four crossover schemes: single-point, multi-point, weighted mean and weighted average. Computer experiments in [14] have shown high performance of the genetic algorithm when solving complicated multi-parametric optimization problems, but it has a bulky structure, and the time required for its completion in complex optimization tasks nonlinearly depends on dimensionality.

In [15], the problem of identifying objects with multiple inputs and multiple outputs is being solved, but this model, like the ones considered previously, does not allow for the input data set to contain missing values.

Type-2 fuzzy models development methods that are based on experimental data, and are capable of solving problems in endocrine system diseases diagnostics area, are presented in [17, 18]. The proposed models and methods, unlike the previous ones, possess the capability to account for missing values in input data sets, thus allowing to introduce artificial input vector magnitude corrections. Still, the issue of such correction factor impacting the value of the fuzzy logic system's interval output value is not investigated in these studies. The issue of a model maintaining its capability to adequately reflect the subject area under conditions of input vector magnitude modifications is not investigated either.

Interval fuzzy set mathematics is used for modeling of complex systems under conditions of input data uncertainty in [19]. The results of the latter research are recommended for application in long-term natural processes study programs in order to shorten the time required for decision-making and to save cost during preliminary research in the hydrogeological exploration area.

The principal problem in most of the identification tasks is the unknown nature of the relation between input variables and the output variable. Besides, under conditions of multidimensional relations, a developer faces the problem of building an identification model that would not have to evaluate all inputs, but only the minimum required number of them. In this case, the input variable set is selected in such way that it ensures the location of the output variable value within bounds acceptable for making the decision. For this reason, within the given formulation of the problem, the requirements for estimating the degree of identity (adequacy) of a model, developed for an object, are increased. The classical identification theory considers relations based on mathematical statistics or information theory as the main characteristic of the identity of a model and a real object [20].

Consider the specifics of developing fuzzy logic systems regarding their capabilities to adequately reflect their subject areas in multidimensional objects identification problems. We shall consider all identification objects partially formalized; the term "adequacy" will be used to define the degree of a model's identity to the real object.

Developing a fuzzy logic system usually consists of the following stages:

1. Determining parameters and their ranges of variation.
2. Estimating qualitative and numerical parameters.
3. Creating a knowledge base and an inference tree.
4. Membership functions generation.
5. Fuzzy model tuning.

According to [13], a fuzzy logic system may be presented as

$$Y=\mathbf{F}(\mathbf{X}, \mathbf{W}, \mathbf{B}, \mathbf{C}),$$

where $\mathbf{X}=(x_1, x_2, ..., x_n)$ is the input variables vector; $\mathbf{W}=(w_1, w_2, ..., w_N)$ is the rules weights vector. Gaussian function is chosen for input variables membership functions:

$$\mu(x) = e^{-\left(\frac{x-b}{c}\right)^2},$$

where $\mathbf{B} = (b_1, b_2, ..., b_q)$, $\mathbf{C} = (c_1, c_2, ..., c_q)$ are membership functions parameters vectors; N is the number of rules, q is the determined number of terms.

Consider a learning data set consisting of M experimental data pairs:

$$(X^t, Y^t),\ t=1...M,$$

where $X^t = (x_1^t, x_2^t, ... x_n^t)$, Y^t are an input vector and a corresponding output variable value of the input-output tuple. The problem of developing an FLS that would be adequate to the subject area is solved by tuning the fuzzy logic system according to the least-squares method: find a vector $(\mathbf{W}, \mathbf{B}, \mathbf{C})$ that would ensure

$$\sum_{t=1}^{M}((F(X^t, W, B, C) - y^t)^2 \to \min.$$

Various optimization methods are used for tuning, the most common of them is a genetic algorithm. For its successful implementation all unknown parameters are organized in a single vector:

$$\mathbf{S}=(\mathbf{W}, \mathbf{B}, \mathbf{C})=(w_1, w_2, ..., w_N, b_1, c_1, b_2, c_2, ..., b_q, c_q).$$

The vector **S** unambiguously defines the FLS $y=\mathbf{F}(\mathbf{X}, \mathbf{W}, \mathbf{B}, \mathbf{C})$. A genetic algorithm has a number of specific procedures, including chromosome generation, crossover and mutation operations implementation, as well as the extremely important selection procedure, or fitness function estimation. As shown in [13], a genetic algorithm is an efficient enough fuzzy model tuning technique. Yet, the resulting fuzzy model loses its transparency, the initial membership functions undergo significant deformations and lose their linguistic meaningfulness. The obtained model is also extremely hard to modify. In order to correct the number of input parameters used for developing the system, changes need to be introduced into the initial parameter set. In case of both type-1 and interval fuzzy models, modifying the number of input parameters results in training the model from scratch. This may result in long training time, which is unacceptable in many practical applications.

Based on the studies and theoretical background considered above, the problem of developing interval type-2 fuzzy systems under the assumption of possible input data set magnitude corrections remains unsolved. The need for such corrections to be introduced by an expert often arises in applied systems, which have their knowledge bases generated from experimental data. At the same time, a requirement for the system to maintain its interval output within predefined margins exists.

3. The aim and objectives of the study

The aim of the research is expanding the capabilities for enhancing subject area reflection adequacy with type-2 fuzzy sets in partially formalized objects identification problems. We propose to develop a type-2 fuzzy logic system, allowing to modify the number of model inputs in the process of objects identification, without the fuzzy system losing its capability for adequate reflection of the subject area.

In order to reach this aim set for the present research, the following objectives were to be achieved:

– to propose a criterion that would enable developing a type-2 fuzzy logic system that would be able to adequately reflect the subject area by using interval type-2 fuzzy sets;

– to develop a method for fuzzy logic system creation from experimental data. It should be possible to modify the resulting system's number of input variables without losing the subject area reflection adequacy.

4. Developing a type-2 fuzzy logic system adequate to the subject area

The present research proposes an information approach to justifying the choice of the fuzzy set type used for fuzzy logic system generation. We are introducing a mutual information function between model's inputs and outputs. In this case, the output value Y entropy can be represented as a sum of the amount of information about Y contained in the input value X, and the mean conditional entropy Y over X [20]:

$$H(Y)=I(X,Y)+H(Y/X). \quad (1)$$

As a fitness function, or target optimization function that changes until it reaches the maximum, we select mutual information between model's inputs and outputs, i. e. I(X,Y). According to the symmetry property I(Y,X)=I(X,Y), the equation (1) can be presented as

$$I(X,Y)=H(X)-H(X/Y). \quad (2)$$

Based on the assumption that, according to the information maximum (Infomax) principle, mutual information must strive for the maximum, i. e. $I(X,Y)\to\max,$ it can be seen from the equation (2) that reaching this goal is possible with unconditional entropy H(X) increasing, and conditional entropy H(X/Y) decreasing. For a fuzzy logic system, the entropy H(X) of its inputs would strive for maximum, when the input values are set using Gaussian interval membership functions, and with the FLS functioning for type-2 membership function values the uniform distribution law is applied (taken as a hypothesis). Fig. 1 features a fragment of applying an input value of x_0, and a uniform type-2 membership function values distribution $\mu(\mu(x_0))$.

Based on the above considerations, we will be solving the problem of developing an FLS with interval membership functions with an optimization criterion [18]. Let NM(MP, ~MP) be a fuzzy model defined by several parameters, MP the membership functions parameters set, ~MP – other parameters. Then the optimization criterion will look like

$$[\Psi(NM(MP,\sim MP))\to\max]\to\min, \quad (3)$$

where Ψ(NM(MP, ~MP)) is the optimization function, according to which the amount of mutual information between the model's inputs and outputs will increase. The first part of the criterion defines an FLS with interval membership functions that would, while maintaining the system's capability for adequate decision making, provide the maximum possible increase in the subject area reflection adequacy. The second part of the criterion states the increase in the system parameter set magnitude through optimization procedures. The latter impacts the increase in conditional entropy H(Y/X). This approach fits the definitions of structural and parametric identification stages.

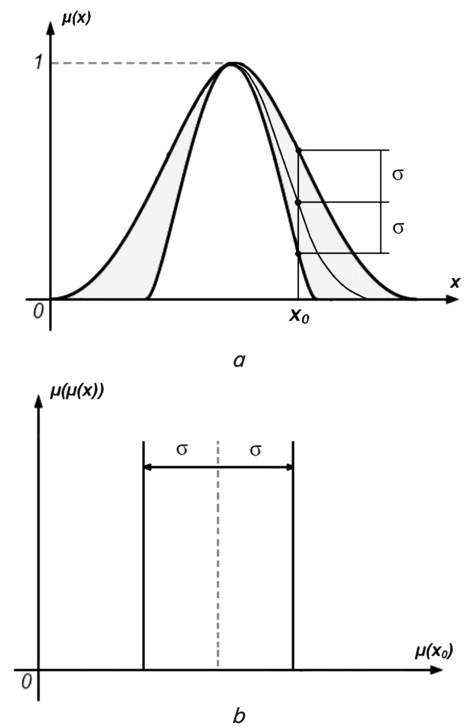

Fig. 1. Interval membership function: *a* – A fragment of applying input value x_0; *b* – Uniform type-2 membership function values distribution $\mu(\mu(x_0))$

We represent the mathematical model for solving the problem stated in this paper as an interval fuzzy identification model. The structure of a fuzzy model with interval membership functions for a "multiple inputs – single output" type multidimensional object is presented in Fig. 2 [16].

The model reflects crisp inputs (x_1,..., x_p) into interval and crisp outputs: $\tilde{Y}=[y_l;y_r].$ Interval type-2 fuzzy sets will be used for describing linguistic variables' fuzzy terms. The mathematical model will then be an interval type-2 model, including a rule base (fuzzy knowledge base), a fuzzification procedure, a fuzzy inference procedure, a type reduction procedure, and a defuzzification procedure (Fig. 2).

In a general case, for n≥1 outputs the fuzzy model rule base will be defined as follows:

R^1: IF $x_1=\tilde{F}_1^1$ AND ... AND $x_p=\tilde{F}_p^1$,
THEN $y_1=G_1^1,...,y_n=G_n^1$;

R^l: IF $x_1=\tilde{F}_1^l$ AND ... AND $x_p=\tilde{F}_p^l$,
THEN $y_1=G_1^l,\ldots,y_n=G_n^l$;

R^M: IF $x_1=\tilde{F}_1^M$ AND ... AND $x_p=\tilde{F}_p^M$,
THEN $y_1=G_1^M,\ldots,y_n=G_n^M$,

where $\tilde{F}_k^l$, k=1, ..., p, l=1, ... m is the type-2 interval fuzzy set of the k-th antecedent of the l-th rule; $\tilde{G}_k^l$, k=1,..., n, l=1,..., M is the type-1 interval set of the k-th consequent of the l-th rule, as defined by its left margin $y_{kl}^{G^l}$ and its right margin $y_{kr}^{G^l}$: $G_k^l=\left[y_{kl}^{G^l},y_{kr}^{G^l}\right]$; M is the number of rules.

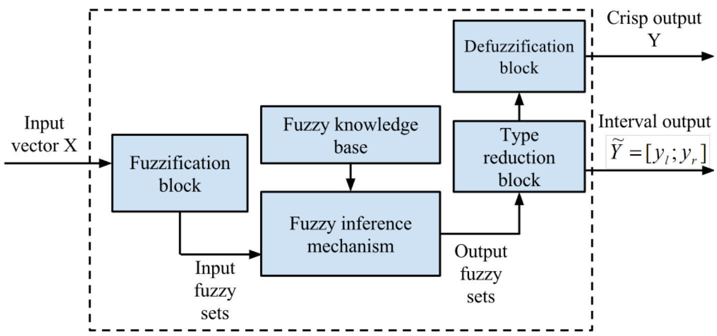

Fig. 2. Structure of an interval type-2 fuzzy model

In order to describe interval type-2 fuzzy sets of the linguistic variables terms, primary Gaussian membership functions with constant centers and uncertain deviations will be used.

A Gaussian primary membership function with a constant center and a variable deviation $\sigma\in[\sigma_l,\sigma_u]$ is defined as

$$\mu_A(x)=e^{-\frac{1}{2}\left(\frac{x-m}{[\sigma_l,\sigma_u]}\right)^2}. \quad (4)$$

A graph of a primary Gaussian membership function with a constant center and a variable deviation (4) is shown in Fig. 3.

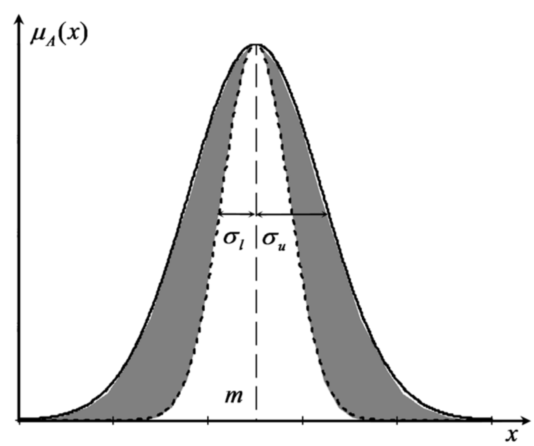

Fig. 3. A Gaussian primary membership function with a constant center and an uncertain deviation

For calculating the output rules fuzzy sets, the minimum t-norm will be used. The output set of a rule R^l will then be derived from the formula

$$\mu_{\tilde{B}^l}(y_k=\int_{b^l\in\left[\underline{f}^l*\mu_{G^l}(y_k),\bar{f}^l*\mu_{G^l}(y_k)\right]}1/b^l), \quad (5)$$

where $*$ is the t-norm operator; $\underline{f}^l$ and $\bar{f}^l$ are the lower and upper bounds of the activation interval $\left[\underline{f}^l,\bar{f}^l\right]$, derived from formulae:

$$\underline{f}^l=\mathop{T}_{k=1}^{p}\underline{\mu}_{\tilde{F}_k^l}(x_k), \quad (6)$$

$$\bar{f}^l=\mathop{T}_{k=1}^{p}\bar{\mu}_{\tilde{F}_k^l}(x_k), \quad (7)$$

where $\underline{\mu}_{\tilde{F}_k^l}(x_k)$ and $\bar{\mu}_{\tilde{F}_k^l}(x_k)$ are the lower and upper membership grades $\mu_{\tilde{F}_k^l}(x_k)$.

The output fuzzy sets of rules $\tilde{B}^l$ are forwarded to the type reduction block of the fuzzy model at once, without being merged into a single set.

Output interval type-2 fuzzy sets type reduction to interval type-1 sets is performed with the sets center method, expressed by the formula (8), where $Y_k(x)$, k=1, ..., n, is an interval set, defined by left and right bounds y_{kl} and y_{kr}.

$$Y_k(x)=\left[y_{kl},y_{kr}\right]=$$

$$=\int_{y_k^1\in\left[y_{kl}^{G^1},y_{kr}^{G^1}\right]}\cdots\int_{y_k^M\in\left[y_{kl}^{G^M},y_{kr}^{G^M}\right]}\int_{f^1\in\left[\underline{f}^1,\bar{f}^1\right]}\cdots\int_{f^M\in\left[\underline{f}^M,\bar{f}^M\right]}1\Bigg/\frac{\sum_{l=1}^{M}f^l y_k^l}{\sum_{l=1}^{M}f^l}. \quad (8)$$

The Karnik-Mendel algorithm is used for calculating the left and right bounds y_{kl} and y_{kr} [21].

5. A method of direct interval fuzzy model generation from experimental data

In order to develop an interval type-2 FLS according to the criterion (3), the method of direct fuzzy model generation from experimental data will be used. Its main stages are described below.

Stage 1. To satisfy the condition of adequate subject area reflection, the following relation must be applied:

$$H(X)\to\max.$$

A possible implementation of this stage, given that input data is available, consists of the following steps.

Let X be an experimental data set:

$$X=\{X_1, X_2, \ldots, X_n\},$$

where $X_i=(x_{1i}, x_{2i},\ldots, x_{ki}, y_i)$, i=1, ..., n; n is the number of experimental data vectors; k is the number of input variables; y is the output value.

The execution sequence is as follows.

Regular fuzzy model generation is performed based on the experimental data set X. To accomplish this, we shall use an approach to fuzzy model development, under which a fuzzy model is built based on the experimental data defining centers of rules' antecedents and consequents fuzzy sets. Thus, a model is obtained, for which all input parameters values are known. It should be stated that such a model has a redundancy associated with it. As a membership function for input variables, the Gaussian membership function is selected:

$$\mu(x) = e^{-\left(\frac{x-m}{c}\right)^2}.$$

For interval membership functions description, the modified Gaussian form will be used. The general form of an interval membership function is:

$$\mu(x) = e^{-\left(\frac{x-m}{[\min(c),\max(c)]}\right)^2},$$

where [min(c), max(c)] is the variation range of the Gaussian function's c parameter (contained within the interval $\sigma \in [\sigma_l, \sigma_u]$), Fig. 3.

In order to perform a transition from a regular to an interval membership function, while simultaneously maintaining the adequacy of decisions taken by the system, the algorithm [17] will be used. It is an algorithm for controlled stretching of one of the input variables membership functions parameter under the restriction of containing the system's interval output within predefined margins.

The algorithm thus allows to build interval type-2 membership functions for input variables and makes sure that the interval fuzzy model operates in a way the most adequate to the subject area.

Stage 2. In the second stage, corrections of the input vector and the experimental data set according to the developer's definition are implemented. As a result of the corrections, missing values in the experimental data and the input vector may appear. Data set transformation is performed with the algorithm [17], which takes experimental data incompleteness into account. According to this algorithm, the correction procedure occurs in combination with interval fuzzy model development; the interval membership functions centers are determined by the respective experimental data. Such approach to developing a model provides fast generation of membership functions, which corresponds to the reality from the conducted experiments' point of view.

The data set transformation algorithm consists of four steps:

1. Consider an initial experimental data set X:

$$X=\{X_1, X_2, ..., X_n\},$$

where $X_i=(x_{1i}, x_{2i},..., x_{si}, y_i)$, i=1,..., n; n is the number of experimental data vectors; s is the number of input variables; y is the output value.

Some of the variables in the experimental data set are unknown, since the input variables set is reduced.

2. Remove the columns with variables values, which are from an expert's point of view unsubstantial for solving the problem at hand, from the initial experimental data set. A correction of the experimental data set takes place, i. e. a transformed data set is created.

3. Follow the procedures from stage 1, i. e. a type-2 fuzzy model with interval membership functions is developed based on the corrected experimental data set.

4. Calculate the output value.

The final step of the resulting model is testing it on verified data, and its operation quality assessment.

Thus, the capabilities of the algorithm [17] to account for missing experimental data values are used for further investigation of the impact of input vector magnitude corrections based on expert's recommendations on the adequacy of the subject area reflection by the fuzzy logic system.

6. Examples of the proposed method's practical application

The present section describes computer modeling for multidimensional objects identification problems. Changes in an identification object's input parameter set magnitude will be considered in conjunction with the results of interval FLS generation.

The proposed method's operation will be shown applied to the problem of artesian well state evaluation. A well as a complex natural system is characterized by high redundancy of its parameters. There is a significant number of natural and anthropogenic factors impacting its state and potential operational characteristics; attempts at accounting for and mathematically describing their influence on the general condition of the system being researched cause the redundancy.

The stages of the algorithm's operation are considered in detail below.

Stage 1. As stated in [19], a hydrogeological system is described as follows:

– 84 input variables (Table 1);

– single output value – artesian well potential for further research and operation on the scale of 0 to 10;

– learning data set.

Table 1

Input variables (fragment)

Variable denotation	Parameter name	Range of values	Exploration stage number	Terms
x_1	Distance to inhabited areas, km	0–50	1	{L – low, M – medium, H – high}
x_2	Distance to highways of state significance, km	0–50	1	{L – low, M – medium, H – high}
x_3	Precipitation (region average), mm	400–800	1	{L – low, M – medium, H – high}
...				

The initial experimental data set X consists of 20 input vectors (Table 2).

Table 2

Learning data set of artesian wells data (fragment)

Variable	1	2	3	4	5	6	7	8	9	10	11	...
x_1	1,5	3	12	22	18	15	15	26	37	35	27	...
x_2	43	12	12,5	25	4	11	32	2	5	2	10	...
x_3	589	589	589	638	638	590	590	562	562	589	589	...
...												

Based on this learning data set, a rule base is built by assigning a fuzzy variable term from Table 1 to every value from Table 2. The obtained formalized knowledge base is presented in Table 3.

Membership functions parameters values of the type-1 fuzzy logic system's term sets are shown in Table 4.

By transforming membership functions according to the algorithm [17], a type-2 fuzzy logic system is obtained, with its membership functions presented in interval form, as shown in Table 5.

Table 3

Type-1 fuzzy logic system knowledge base (fragment)

	R^1	R^2	R^3	R^4	R^5	R^6	R^7	R^8	R^9	R^{10}	R^{11}	...
x_1	H	H	H	M	M	H	M	M	H	H	M	...
x_2	H	H	M	M	L	L	H	H	M	M	M	...
x_3	L	L	M	L	H	H	H	H	M	M	M	...
...												

Table 4

Membership functions of the type-1 fuzzy logic system (fragment)

Variable	Variable name	Terms					
x_1	Distance to inhabited areas, km	L		M		H	
		b	c	b	c	b	c
		0	15.4	25	15.4	50	15.4
x_2	Distance to highways of state significance, km	L		M		H	
		b	c	b	c	b	c
		0	15.4	25	15.4	50	15.4
x_3	Precipitation (region average), mm	L		M		H	
		b	c	b	c	b	c
		400	43.9	600	43.9	800	43.9
...							

Table 5

Membership functions of the interval type-2 fuzzy logic system (fragment)

Variable	Variable name	Terms					
x_1	Distance to inhabited areas, km	L		M		H	
		b	c	b	c	b	c
		0	[9.03; 84.93]	25	[6.57; 61.78]	120	[23.61; 221.86]
x_2	Distance to highways of state significance, km	L		M		H	
		b	c	b	c	b	c
		–0.01	[11.29; 106.11]	25	[6.57; 61.78]	50	[6.57; 61.78]
x_3	Precipitation (region average), mm	L		M		H	
		b	c	b	c	b	c
		301.25	[11.81; 110.97]	600	[18.74; 176.12]	480.4	[18.74; 176.12]
...							

Stage 2.

1. The initial experimental data set X is identical to the data set from Table 1 of the first stage. Missing values will appear in this data set later on, as a result of artificial feature removal.

2. Introducing corrections into the experimental data set based on expert's recommendations. Receiving any meaningful output over the entire set of 84 parameters is close to impossible, considering the difficulties related to input data redundancy. Thus, an initial interference of an expert in order to optimize the input feature space is a necessity resulting from problem dimensionality. One of the possible outcomes of such optimization is shown below; the resulting input feature space X' consists of 38 features:

$X'=\{x_3, x_4, x_6, x_7, x_{12}, x_{13}, x_{14}, x_{15}, x_{18}, x_{19}, x_{24}, x_{26}, x_{32}, x_{33}, x_{34}, x_{35}, x_{36}, x_{37}, x_{40}, x_{41}, x_{43}, x_{44}, x_{45}, x_{47}, x_{50}, x_{54}, x_{60}, x_{62}, x_{64}, x_{67}, x_{74}, x_{75}, x_{76}, x_{79}, x_{80}, x_{82}, x_{83}, x_{84}\}$.

3. Similarly to the procedures of the first stage, another type-2 fuzzy model is built, this time for the input feature set suggested by the expert. Table 6 shows membership functions of the parameters remaining after the expert had removed all non-informative and redundant features.

Table 6

Membership functions of the type-2 interval fuzzy logic system on the reduced input feature set (fragment)

Variable	Variable name	Terms					
x_3	Precipitation (region average), mm	L		M		H	
		b	c	b	c	b	c
		301.25	[24.53; 30.75]	600	[38.93; 48.81]	480.4	[38.93; 48.81]
x_4	Existing contamination sources	L		M		H	
		b	c	b	c	b	c
		0	[1.92; 2.41]	3.33	[1.92; 2.41]	7.13	[1.92; 2.41]
x_6	Plumbum concentration (average over existing wells in the area), mg/dm^3	L		M		H	
		b	c	b	c	b	c
		0	[0.02; 0.03]	0	[0.01; 0.02]	0.144	[0.01; 0.02]
...							

4. Table 7 shows an example of determining the output value on verified data of the test data set under conditions of missing values in the input data. The effect of missing values appearing in the input data is created due to the requirement to obtain a result long before the process of collecting experimental data is completed. The results of the system's operation with the different number of missing values in the input data are shown in Table 7.

Table 7

Model operation on uncertain hydrogeological data (fragment)

№	HGE stage	Known features/total	Model output	Expert's estimation – well potential
1	1	9/38	[0.5; 9.19]	high
	2	16/38	[5.88; 8.45]	
	3	32/38	[7.7; 8.36]	
2	1	9/38	[5.39; 9.74]	high
	2	16/38	[5.45; 9.58]	
	3	32/38	[8.23; 8.47]	
...				

The number of missing values of the input variables decreases from the first hydrogeological exploration (HGE) stage to the third, when all of the input variables values are known.

7. Interval type-2 FLS operation results discussion

Analysis of the results received from the interval fuzzy logic system for artesian well state evaluation shows a general trend of the resulting output variable interval to expand with the increase of the number of unknown input variables. At the same time, the system maintains its capability to provide an interpretable interval at its output. For instance, on stages 2 and 1 of the example 2, Table 7, the result may generally be considered acceptable, since the model maintains its capability to reflect the subject area in an adequate manner, although with a significant degree of uncertainty associated with the decision taken. Further removal of input variables causes the uncertainty zone to fill the entire output parameter domain of definition, or its major part, as can be seen on stage 1 of the example 1, Table 7.

It is confirmed by the experiments results that the use of the proposed model generation method based on experimental data not only enables estimations of the output parameter value, but also drawing a conclusion as to whether extending the input vector is advisable.

The proposed method may be used for developing fuzzy logic advisors in diagnostic tasks, as well as for developing system state evaluation models in a broader sense. Natural system state evaluation is a partial case of this method's application. This paper provides justification for using type-2 fuzzy sets in mathematical models dealing with uncertain input data. The justification is performed theoretically, based on information theory considerations, and confirmed experimentally.

The conducted research is a continuation of investigating FLS adequacy in [18, 19].

8. Conclusions

1. An optimization criterion is proposed, measuring the increase in mutual information reflecting from a fuzzy logic system's inputs to its outputs. Maximization and minimization of respective components of the criterion allows to develop a type-2 fuzzy logic system, which reflects the subject area in an adequate manner.

2. A method of building a type-2 fuzzy logic system is proposed, which allows, while identifying partially formalized objects, to modify the number of model inputs without the fuzzy system losing its capability to adequately reflect the subject area. Experimental data is assumed as the primary source for generating fuzzy knowledge base rules. A sequence of steps required for developing a type-2 fuzzy logic system, optimal according to the considered criterion, is formulated.

References

1. Minaev, Yu. N. Metody i algoritmy resheniya zadach identifikacii i prognozirovaniya v usloviyah neopredelennosti v neyrosetevom logicheskom bazise [Text] / Yu. N. Minaev. – Moscow: Goryachaya liniya-Telekom, 2003. – 205 p.
2. Zadeh, L. A. Fuzzy sets as a basis for a theory of possibility [Text] / L. A. Zadeh // Fuzzy Sets and Systems. – 1999. – Vol. 100. – P. 9–34. doi: 10.1016/s0165-0114(99)80004-9
3. Liang, Q. Interval type-2 fuzzy logic systems: theory and design [Text] / Q. Liang, J. M. Mendel // IEEE Transactions on Fuzzy Systems. – 2000. – Vol. 8, Issue 5. – P. 535–550. doi: 10.1109/91.873577
4. Zeng, J. Type-2 Fuzzy Sets for Pattern Classification: A Review [Text] / J. Zeng, Z.-Q. Liu // 2007 IEEE Symposium on Foundations of Computational Intelligence. – 2007. doi: 10.1109/foci.2007.372168
5. Venkata Subba Reddy, P. Fuzzy Medical Expert Systems for Clinical Medicine Learning Through the Fuzzy Neural Network [Text] / P. Venkata Subba Reddy, A. Sadana // International Journal of Clinical Medicine Research. – 2015. – Vol. 2, Issue 5. – P. 54–60.
6. Dahiya, S. Analysis of groundwater quality using fuzzy synthetic evaluation [Text] / S. Dahiya, B. Singh, S. Gaur, V. K. Garg, H. S. Kushwaha // Journal of Hazardous Materials. – 2007. – Vol. 147, Issue 3. – P. 938–946.
7. Raskin, L. G. Informacionnoe obespechenie nechetkih ehkspertnyh sistem [Text] / L. G. Raskin, O. V. Seraya, T. I. Katkova, V. A. Golovko // Systemy obrobky informatsiyi. – 2013. – Issue 6 (113). – P. 13–16.
8. Raskin, L. G. Fuzzy model of nonlinear multi-index transportation problem [Text] / L. G. Raskin, O. V. Seraya, O. I. Dunaevskaya // Eastern-European Journal of Enterprise Technologies. – 2012. – Vol. 6, Issue 4 (60). – P. 15–17. – Available at: http://journals.uran.ua/eejet/article/view/5676/5106
9. Samuel, O. W. A web based decision support system driven by fuzzy logic for the diagnosis of typhoid fever [Text] / O. W. Samuel, M. O. Omisore, B. A. Ojokoh // Expert Systems with Applications. – 2013. – Vol. 40, Issue 10. – P. 4164–4171. doi: 10.1016/j.eswa.2013.01.030
10. Agrawal, P. Fuzzy rule-based medical expert system to identify the disorders of eyes, ENT and liver [Text] / P. Agrawal, V. Madaan, V. Kumar // International Journal of Advanced Intelligence Paradigms. – 2015. – Vol. 7, Issue 3/4. – P. 352. doi: 10.1504/ijaip.2015.073714
11. Sabzi, H. Z. Optimization of adaptive fuzzy logic controller using novel combined evolutionary algorithms, and its application in Diez Lagos flood controlling system, Southern New Mexico [Text] / H. Z. Sabzi, D. Humberson, S. Abudu, J. P. King // Expert Systems with Applications. – 2016. – Vol. 43. – P. 154–164. doi: 10.1016/j.eswa.2015.08.043
12. Sun, Z. Type-1/type-2 fuzzy logic systems optimization with RNA genetic algorithm for double inverted pendulum [Text] / Z. Sun, N. Wang, Y. Bi // Applied Mathematical Modelling. – 2015. – Vol. 39, Issue 1. – P. 70–85. doi: 10.1016/j.apm.2014.04.035
13. Rotshteyn, A. P. Intellektual'nye tekhnologii identifikacii: nechetkie mnozhestva, geneticheskie algoritmy, neyronnye seti [Text] / A. P Rotshteyn. – Vinnica: «UNIVERSUM–Vinnicya», 1999. – 320 p.
14. Bilan, S. M. Doslidzhennya mozhlyvostey henetychnoho alhorytmu v zadachakh pro komivoyazhera [Text] / S. M. Bilan, N. R. Kondratenko, O. A. Tkachuk // Reyestratsiya, zberihannya i obrobka danykh. – 2004. – Vol. 6, Issue 3. – P. 50–57.

15. Kondratenko, N. R. Interval'ni nechitki modeli typu-2 v zadachakh identyfikatsiyi ob'yektiv z bahat'ma vkhodamy ta vykhodamy [Text] / N. R. Kondratenko, O. V. Cheboraka, O. A. Tkachuk // Systemy obrobky informatsiyi. – 2011. – Issue 3 (93). – P. 48–52.
16. Noshadi, A. Optimal PID-type fuzzy logic controller for a multi-input multi-output active magnetic bearing system [Text] / A. Noshadi, J. Shi, W. S. Lee, P. Shi, A. Kalam // Neural Computing and Applications. – 2015. – Vol. 27, Issue 7. – P. 2031–2046. doi: 10.1007/s00521-015-1996-7
17. Kondratenko, N. R. Nechitki lohichni systemy z vrakhuvannyam propuskiv v eksperymental'nykh danykh [Text] / N. R. Kondratenko, N. B. Zelins'ka, S. M. Kuzemko // Naukovi visti natsional'noho tekhnichnoho universytetu Ukrayiny "Kyyivs'kyy politekhnichnyy instytut". – 2004. – Issue 5. – P. 37–41.
18. Kondratenko, N. R. Pidvyshchennya adekvatnosti nechitkykh modeley za rakhunok vykorystannya nechitkykh mnozhyn typu 2 [Text] / N. R. Kondratenko // Naukovi visti natsional'noho tekhnichnoho universytetu Ukrayiny "Kyyivs'kyy politekhnichnyy instytut". – 2014. – Issue 6. – P. 56–61.
19. Kondratenko, N. Interval fuzzy modeling of complex systems under conditions of input data uncertainty [Text] / N. Kondratenko, O. Snihur // Eastern-European Journal of Enterprise Technologies. – 2016. – Vol. 4, Issue 4 (82). – P. 20–28. doi: 10.15587/1729-4061.2016.75679
20. Teoriya informacii i ee prilozheniya (sbornik perevodov) [Text] / A. A. Harkevich (Ed.). – Moscow: Fizmatgiz, 1959. – 328 p.
21. Karnik, N. N. Type-2 fuzzy logic systems [Text] / N. N. Karnik, J. M. Mendel, Q. Liang // IEEE Transactions on Fuzzy Systems. – 1999. – Vol. 7, Issue 6. – P. 643–658. doi: 10.1109/91.811231

Розглядається система масового обслуговування вигляду G/M/1/K з розподілом Вейбула. Для самоподібного трафіку розроблена імітаційна модель за допомогою програмного пакету Simulink в середовищі Matlab. За допомогою сплайн-функцій (лінійних та кубічних сплайнів) отримано відновлення самоподібного трафіку за його значеннями в вузлах інтерполяції.

Отримані результати дозволять передбачити необхідний обсяг буферних пристроїв, тим самим запобігти перенавантаження мережі та перевищення нормативних значень характеристик QoS

Ключові слова: самоподібний трафік, розподіл Вейбула, система масового обслуговування, відновлення, сплайн-функції

Рассматривается система массового обслуживания вида G/M/1/K с распределения с распределением Вейбулла. Для самоподобного трафика разработана имитационная модель с помощью программного пакета Simulink в среде Matlab. С помощью сплайн-функций (линейных и кубических сплайнов) получено восстановление самоподобного трафика по его значениям в узлах интерполяции.

Полученные результаты позволят предусмотреть требуемый объем буферных устройств, тем самым избежать перегрузок в сети и превышений нормативных значений характеристик QoS

Ключевые слова: самоподобный трафик, распределение Вейбулла, система массового обслуживания, восстановление, сплайн-функции

UDC 621.391
DOI: 10.15587/1729-4061.2017.102999

SPLINE-APPROXIMATION-BASED RESTORATION FOR SELF-SIMILAR TRAFFIC

I. Strelkovskaya
Doctor of Technical Sciences, Professor, Director of Educational and Research Institute of Infocommunications and Software Engineering
Department of Higher Mathematics*
E-mail: i.strelkovskaya@onat.edu.ua

I. Solovskaya
PhD, Associate Professor
Department of Switching Systems*
E-mail: i.solovskaya@onat.edu.ua

N. Severin
Senior Lecturer
Department of Information Technologies*
E-mail: n_severin@ukr.net

S. Paskalenko
Postgraduate student
Department of Higher Mathematics*
E-mail: s.paskalenko@onat.edu.ua

*O. S. Popov Odessa National Academy of Telecommunications
Kuznechna str., 1, Odessa, Ukraine, 65029

1. Introduction

The development of modern telecommunications is connected with the active implementation of Next Generation Networks (NGN), which are multi-service, multiprotocol and invariant to switching technologies. The concept of NGN implies active implementation of high-speed multiservice access technologies (xDSL, FTTx/PON, BWA), packet