Data Processing System for Non-Stationary Processes Based on The Synthesis of Soft Computing Components

Kholmonov Sunatillo Makhmudovich
PhD in Technical Sciences, Department of Information Technologies, Samarkand State University, Samarkand, Uzbekistan

Mirsanova Gulbahor Muhammadi qizi
Graduate student, Department of Information Technologies, Samarkand State University, Samarkand, Uzbekistan

ABSTRACT

The problem has been formulated and the scientific and methodological foundations of data processing systems for non-stationary objects based on neural networks, fuzzy set models, fuzzy inference algorithms, and neuro-fuzzy networks have been developed. Are proposed mechanisms for structural and parametric identification, finding a set of terms of linguistic variables, rules of inference, determining the coefficients of fuzzy rules, using the properties of self-adaptation, self-regulation, network organization, and the formation of databases and knowledge. The efficiency of the implemented algorithms was evaluated on the basis of test unimodal and multimodal output functions.

Keywords: Non-Stationary Process, Processing, Neural Network, Fuzzy Set, Fuzzy Inference, Neuro-Fuzzy Network.

Relevance of the topic. A promising direction for improving and developing models, algorithms for data processing systems of non-stationary objects is the use of neural networks (NN), fuzzy set models, fuzzy inference algorithms, neuro-fuzzy networks, the mechanisms of which are focused on identifying time series and non-linear "input-output" dependencies in conditions of incomplete information, uncertainty [1,2].

However, the use of NFN to identify non-stationary objects is associated with such difficulties as the large dimension of the fuzzy model parameter vector, the complexity of the mathematical description and representation of objects, the dependence of the approximation efficiency on the size of the formed training sample, establishing the adequacy of the new model, transferring the dynamic properties of time series to the NN [3,4].

The key point in the implementation of these approaches is the creation of mechanisms for using the unique properties of self-adaptation, self-regulation and network organization of the NN to identify and process data of non-stationary objects [5,6].

Fuzzy identification of non-stationary objects. The identification data processing model based on the NFN is represented with tuples:

$$A=\langle \{X_i, y_i\}, BR, DB, I, G(BR), L(DB), F\rangle$$

where $\{X_i y_i\}$, $(i=\overline{1,n})$ is the training sample; BR - base of fuzzy rules; DB - database; I – fuzzy inference mechanism; $G(BR)$ - generation and optimization of BP; $L(DB)$ – database generation and optimization; F is a function that estimates the minimum standard deviation between the reference characters of the output variable y_j and the values obtained on the basis of fuzzy identification in the form of a point estimate for $F(X_j, P)$, P is the vector of fuzzy logic parameters.

The problem of structural identification is solved and methods for finding a set of terms of linguistic variables, rules of inference, determining the coefficients of fuzzy rules, processing data of non-stationary objects are investigated [7].

As a basic model of fuzzy inference, it is proposed to use the Takagi-Sugeno model of the first order with rules of the form [8]:

Π_i: ЕСЛИ есть A_{i1} И...И x_j есть A_{ij} И...И x_m есть A_i

ТО

$$y = k_{i1}x_1 + \ldots + k_{ij}x_j + \ldots + k_{im}x_m + k_{i0}, \quad i = 1, \ldots, n,$$

where A_i are fuzzy sets and their corresponding membership functions (MF) built in the input space x_i;

$k_{ij}(j = 0, \ldots, m)$ are the coefficients of the function arguments.

The adopted model is characterized by the use of a gaussian MF

$$\mu(x) = \exp\left[-\left(\frac{x-c}{\sigma}\right)^2\right], \qquad (1)$$

where c is the center of the fuzzy set; σ is the steepness of the function.

To form a fuzzy knowledge base, a method is proposed, which is a procedure for sequential identification by a training sample [9,10].

The result of the fuzzy control rule based on the KB is represented by real numbers, and the indices of the fuzzy sets of input variables are tied to the coefficients of the fuzzy rules in the form

$$Ab = < IFS_1, IFS_2, \ldots, IFS_n, k_0, k_1, \ldots, k_n >,$$

where $IFS_1, IFS_2, \ldots, IFS_n$ are indices of fuzzy sets for n input variables; $k_0, k_1, \ldots, k_n$ – coefficients of fuzzy rules [11,12].

The required set of fuzzy control rules is defined as:

IF x_1 *is* IFS_{11} *...AND...* x_n *is* IFS_{1n} *THEN* R_1:

$$y = k_{10} + k_{11}x_1 + \ldots + k_{1n}x_n;$$

IF x_1 *is* IFS_{21} *...AND...* x_n *is* IFS_{2n} *THEN* R_2:

$$y = k_{20} + k_{21}x_1 + \ldots + k_{2n}x_n;$$

...

IF x_1 *is* IFS_{q1} *...AND...* x_n *is* IFS_{qn} *THEN* R_q:

$$y = k_{q0} + k_{q1}x_1 + \ldots + k_{qn}x_n.$$

The obtained results of theoretical studies in the form of a set of fuzzy rules made it possible to develop the following structural identification algorithm based on the rules and regulations of genetic systems [13,14].

Algorithm for structural identification of NFN. The principle of structural identification of time series is to develop the following procedures: generation of the initial parameter Ab; loop execution for each variant of the time series Ag_i; generation of inference rules with coefficients taken from knowledge base; calculating the affinity of Ab to Ag_i.

The identification algorithm selects the parameters of the fuzzy model and the appropriate membership function (MF) and includes the following steps [15,16].

Step 1. Generation of the most appropriate control rules from the KB.

Step 2. Formation of a set of linguistic terms.

Step 3. Fuzzy identification of a non-stationary process, i.e. selection of informative features from set C, formation of a modified set C^*.

Step 4. Computing the affinity of the set C^* to the controlled Ag_i.

Step 5. Correction of the result of

identification and approximation and obtaining A, filling the memory M.

Step 6. Suppression of the parameters of the time series model leading to Ab.

Step 7. Selection of an informative training sample corresponding to Ab by eliminating the d worst samples.

Step 8. Checking the stop criteria. If not reached, then go to step 2, otherwise go to step 4.

Step 9. Finish.

Based on the results of the algorithm, a set of inference rules with calculated coefficients of fuzzy rules is established.

As a result of identifying the "input-output" dependencies for n input variables $x_1, x_2, \ldots, x_n$ and one output variable y, in accordance with the developed algorithm, a training sample of the form

$$\left(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}, y^{(k)}\right),$$

$$k = \overline{1, K},$$

where $x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}, y^{(k)}$ are the values of input variables $x_1, x_2, \ldots, x_n$ and output variable y in the k th example; K is the total number of examples in the training set.

The domain of definition of the input variable is divided into segments, on each of which a Gaussian MF is set, and the minimum and maximum values of the input variable are also set.

Next, bases of inference rules are formed, on the basis of which parametric identification is performed.

Algorithm for parametric fuzzy identification. The purpose of parametric identification is to find such a fuzzy NFN model with such a set of parameters for which the quality of identification will be the best. Parametric identification includes the procedure for optimizing the primary values of the MF parameters, which consists in changing the parameters of the terms of all linguistic variables. The identified object is represented by a fixed length string:

$$Ab = \langle c_{11}, \ldots, c_{1m},\ \ldots,\ c_{n1}, \ldots, c_{nm},\ \sigma_{11}, \ldots, \sigma_{m1},\ \ldots,\ \sigma_{n1}, \ldots, \sigma_{nm} \rangle,$$

where c_{ij}, σ_{ij}, $i = \overline{1, n}$; $j = \overline{1, m}$; are the parameters of gaussian MFs of the form (2) for input variables, each of which has terms; each of the MFs is specified in the range of a closed interval of real numbers $X = [x_{\min}, x_{\max}]$.

Object identification based on fuzzy logic rules is represented as

$$Ab = < c_{11}, \ldots, c_{1m}, \ldots, c_{n1}, \ldots, c_{nm}; \sigma_{11}, \ldots, \sigma_{1m}, \ldots, \sigma_{n1}, \ldots \sigma_{nm}; k_{10}, \ldots, k_{1n}, \ldots, k_{q0}, \ldots, k$$

,

where $k_{i0}, \ldots, k_{in}$, $i = \overline{1, q}$ are coefficients q of fuzzy rules

The developed parametric identification algorithm is presented in the form of the following steps.

Step 1. Initialize the initial Ab and run the loop for the Ab_i parameters.

Step 2. Selecting an appropriate MF for Ab_i and calculating the affinity for Ag.

Step 3. Choice of n most appropriate parameters.

Step 4. Formation of the set of linguistic terms of fuzzy rules C.

Step 5. Selection of informative features from the set , formation of the modified set C^*.

Step 6. Calculation of the affinity of the set C^* to the set *Ag.*

Step 7. Correction of the result of identification and approximation and obtaining Ab, filling the memory M.

Step 8. Selection of an informative training sample corresponding to Ab by eliminating d worse samples.

Step 9: Checking the Stop Criteria If not achieved, go to step 2, otherwise, go to step 4. The result is a word form with better affinity, containing the parameters of the MF.

The solution of problems of identification and approximation of data in the second direction was studied in detail in [4] by the authors. Below, we will consider the construction of a network for identifying and approximating time series and "input-output" dependencies based on fuzzy inference models and a neural network.

Identification of non-stationary objects based on NFN. As an example, we consider a network with two input variables, each of which is represented by two MFs and four fuzzy control rules. The outputs of the neurons of the first layer are the values of the Gaussian MF ($\mu_{Ai}(x_1)$ and $\mu_{Bi}(x_2)$, $i = 1,2$),calculated in accordance with (1).

The weights of the neurons of the second layer are obtained by multiplying the outputs of the first layer and represent the weight of the rules:

$$w_i = \mu_{Ai}(x_1) \times \mu_{Bi}(x_2),$$
$$i = \overline{1,4}.$$

In the third layer, the weight of the i-th neuron is calculated as the ratio of the weight of the i-th rule to the sum of the weights of all rules: $\overline{w}_i = \dfrac{w_i}{\sum w_i}$, $i = \overline{1,4}$.

The activation function of neurons of the fourth layer is represented as

$$\overline{w}_i f_i = p_i x_1 + q_i x_2 + r_i,$$
$$i = \overline{1,4},$$

where $\{p_i, q_i, r_i\}$ are the fuzzy inference parameters according to the i-th rule, in the case of the Takagi-Sugeno model of the 1st order.

In the neuron of the fifth layer, the output is calculated as the sum of the outputs of the previous layer:

$$F = \sum \overline{w}_i f_i, \qquad i = \overline{1,4}.$$

To evaluate the effectiveness of the implemented algorithms, experiments are organized on the basis of the test output functions given in Table. 1, where function 1 is unimodal and function 2 is multimodal.

Table 1

	Test function	Function scope
1	$F_1(x_1, x_2) = x_1^2 + x_2^2$	$x_1,\ x_2 \in [-3,3]$
2	$F_2(x_1, x_2) = 3(1 - x_1)^2 \exp(-x_1^2 - (x_2 + 1)^2) - -10\left(\frac{x_1}{5} - x_1^3 - x_2^5\right)\exp(-x_1^2 - x_2^2) - \frac{1}{3}\exp(-(x_1 + 1)^2 - x_2^2)$	$x_1,\ x_2 \in [-3,3]$

The formed training sample consisted of 121 examples. Each test function is calculated from all data in the training set.

The initial values of the MF parameters are set in such a way that the MFs are evenly distributed over the domain of definition of the output functions.

It has been established that identification algorithms based on fuzzy inference models require significant computational time, which is associated with the length of the string of parameters presented in the antibody. When 3 terms are used for linguistic evaluation of input variables, the length of the antibody string is 39 parameters, in the case of 5 terms - 95 parameters, in the case of 7 terms - 175 parameters.

Identification algorithms based on NFN change not only the coefficients of fuzzy rules, but also the number of fuzzy inference rules. If in algorithms of fuzzy inference models the number of rules for two input variables with five MFs was 25 for both test functions, then for algorithms based on NFN the number of rules is reduced to 19 for test function 1 and to 23 for test function 2.

Conclusion. Thus, methods, algorithms and a software package for processing data of non-stationary objects have been developed based on the use of models of fuzzy sets, fuzzy logic, neural networks, neuro-fuzzy networks, as well as their synthesized models. The obtained mechanisms make it possible to carry out identification, approximation of time series, dynamic smoothing, using the methods of parallel information processing, intellectual analysis tools, unique properties of self-adaptation, self-regulation, network organization. It is determined that when synthesizing NFN algorithms, the best results of information processing are achieved due to the combined adjustment of both fuzzy rules parameters and MF parameters, and the accuracy of time series approximation is also increased. NFN makes it possible to eliminate the dependence of systems on the ergonomic factor, reduces the role of a human expert in making decisions in real conditions, and allows for operational control and correction of distortions of transmitted data under time constraints in an interactive mode.

References

1. Jumanov, I. I., Djumanov, O. I., & Safarov, R. A. (2021, November). Mechanisms for optimizing the error control of micro-object images based on hybrid neural network models. In AIP Conference Proceedings (Vol. 2402, No. 1, p. 030018). AIP Publishing LLC.
2. Джуманов, О. И., & Холмонов, С. М. (2016). Методы идентификации микрообъектов на основе регуляции параметров глобальных и локальных характеристик изображений. In Проблемы оптимизации сложных систем (pp. 171-179).
3. Джуманов, О. И., & Холмонов, С. М. (2016). Методы оптимизации идентификации нестационарных объектов на основе нейро-нечеткой сети с настройкой параметров вычислительных схем. Проблемы вычислительной и прикладной математики, (3), 20-27.
4. Жуманов, И. И., & Холмонов, С. М. (2017). Исследование эффективности программного комплекса идентификации случайных временных рядов на основе генетической настройки параметров. Проблемы современной науки и образования, (19 (101)), 22-25.
5. Djumanov, O., & Kholmonov, S. (2012). Methods and algorithms of selection the informative attributes in systems of adaptive data processing for analysis and forecasting. Applied Technologies & Innovations, 8(3).
6. Djumanov, O. I., & Kholmonov, S. M. (2016). The modified model of training of neural networks in computer industrial systems with modules for nonstationary objects images processing. Control and Management. South Korea, Seoul-Uzbekistan. Tashkent, (5), 54-58.
7. Jumanov, I. I., & Xolmonov, S. M. (2021, February). Optimization of identification of non-stationary objects due to information properties and features of models. In IOP Conference Series: Materials Science and Engineering (Vol. 1047, No. 1, p. 012064). IOP Publishing.
8. Ibragimovich, J. I., Isroilovich, D. O., & Maxmudovich, X. S. (2020, November). Effective recognition of pollen grains based on parametric adaptation of the image identification model. In 2020 International Conference on Information Science and Communications Technologies (ICISCT) (pp. 1-5). IEEE.
9. Холмонов, С. М., & Жумаёзов, У. З. Оптимизация отбора информативных признаков нестационарных объектов на основе модифицированных вычислительных схем сверточной нейронной сети. ILMIY AXBOROTNOMA, 64.
10. Холмонов, С. М., & Сафаров, Р. (2013). Оценка состава полезных минералов в горной массе на основе восстановления прямых в спектре интенсивностей изображений. SCIENCE AND WORLD, 117.
11. Холмонов, С. М., & Хамроева, Б. Ш. (2016). Методы и алгоритмы компьютерного зрения изображений непрерывных объектов на основе синтеза методов статистического и нечеткого моделирования. Научная дискуссия: вопросы технических наук, (4), 65-70.
12. Makhmudovich, K. S., & Sobirjon o'g'li, S. V. (2022). Intellectualization of Data Processing of Non-Stationary Objects in a Complex Problem Environment. EUROPEAN JOURNAL OF INNOVATION IN NONFORMAL EDUCATION, 2(2), 364-368.
13. Makhmudovich, K. S. ., & Ismoilovich, N. A. . (2022). Increasing the Reliability of Texts of Electronic Documents Based on Soft Calculations under Parametric Uncertainty. Middle European Scientific Bulletin, 21, 126-132. Retrieved from

https://cejsr.academicjournal.io/index.php/journal/article/view/1072

14. Makhmudovich, K. S., & qizi, M. G. M. (2022). Optimization of Data Processing of Non-Stationary Processes Based on Setting the Parameters of Fuzzy Models. International Journal of Human Computing Studies, 4(2), 36-42. Retrieved from https://journals.researchparks.org/index.php/IJHCS/article/view/2723
15. Israilovich, D. O., Makhmudovich, K. S., & Uktomovich, Y. F. (2021). Increasing The Credibility Of Forecasting Random Time Series Based On Fuzzy Inference Algorithms. International Journal of Progressive Sciences and Technologies, 26(1), 12-15.
16. Djumanov, O. I., Kholmonov, S. M., & Ganiev, J. M. (2021). Optimization of data processing based on neural networks and properties of non-stationary objects. Theoretical & Applied Science, (4), 165-168.