Construction of hybrid intellectual monitoring and decision-making systems

Z Sh Juraev, D T Muhamediyeva and D M Sotvoldiev

Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Amir Temur street, 108, 100200, Tashkent, Uzbekistan

Abstract. The development of algorithms for constructing hybrid intelligent monitoring and decision-making systems, optimization of problem solving with fuzzy initial information, analysis of the correspondence model of complex difficult to formalize processes is considered. The main difference between the proposed development and the traditional ones is the use of modern intelligent information technologies for the development of algorithmic tools for constructing hybrid intelligent monitoring and decision-making systems.

1. Introduction

The ability to represent the knowledge contained in the data in a way that is understandable to the user is one of the main advantages of fuzzy models in comparison with other non-linear systems, such as, for example, neural networks. The construction of a fuzzy classification model based on the available data set allows one to obtain a solution to the classification problem in the form of an interpreted set of fuzzy if-then-rules. However, the construction of a model with both high classification accuracy and good interpretability of the results is a rather complicated task. Initially, the main goal in constructing fuzzy classification models was to increase the classification accuracy. For the automatic training of model parameters, a large number of training algorithms from the theory of neural networks and genetic algorithms have been proposed [1-3]. In recent years, more attention has been paid to the construction of not only accurate, but also a fairly simple (interpreted) fuzzy monitoring and classification model [1-2]. The choice of the optimal correlation between the accuracy and interpretability of the model occurred in most cases due to the presentation of these characteristics as separate functions as part of a single scalar optimization function, which allowed the use of optimization methods to build a fuzzy model. An example of this approach is a fuzzy classification model constructed using a genetic algorithm [4]. In the present work, a multicriteria genetic algorithm is used to construct a fuzzy classification model. This algorithm allows you to build several sets of undominated fuzzy rules that are optimal with respect to three objective functions: classification accuracy, the number of fuzzy rules in the set, and the total length of fuzzy rules. The end user is given the opportunity among the obtained sets of rules to choose the most suitable from the point of view of correlation between classification accuracy and interpretability of the corresponding fuzzy classification model.

2. Fuzzy monitoring model.
The analysis of fuzzy decision-making models (DM) in the tasks of the decision-making support system (DSS) of monitoring and control [1-3] shows that their main functional module is the organizer of fuzzy inference. The core of the latter is the fuzzy production rule, in which the conditions and conclusions are formulated in the form of fuzzy linguistic statements (judgments) regarding the meanings of certain linguistic variables.

Therefore, a fuzzy conclusion and fuzzy linguistic utterances, as well as methods for their transformation, orderliness and organization, occupying a central place in intellectual DSS.

The process of fuzzy inference is a set of procedures implemented in a certain sequence in accordance with the methods used. This process combines all the basic concepts of the theory of fuzzy sets: membership functions, linguistic variables, fuzzy logical operations, methods of fuzzy implication and fuzzy composition. An appropriately organized set of considered procedures forms a fuzzy inference system (FIS) [1].

This section discusses the main models of the production system: fuzzy production rules, as well as the functional diagram, stages and procedures of strategic offensive arms.

Of the fuzzy decision-making models obtained in [5-6] in monitoring and control systems, the characteristic, which includes all the basic procedures of fuzzy inference, is a set of fuzzy mathematical expressions for assessing a fuzzy situation. Therefore, the analysis and generalization of the rules of fuzzy products will be carried out on the example of these fuzzy models, which have the following form.

For the convenience of subsequent calculations, we simplify these expressions by introducing the notation: \( A = (X \cup Y) \), \( B=S \) where \( A \) - the generalized (combined) set of input variables defined by the Cartesian product \( X \times Y \), \( B \) - the set of output variables [2].

Membership function \( A \) has the form

\[
\mu_A(a) = \mu_X(x) \cap \mu_Y(z).
\]

With this in mind, the generalized model of fuzzy inference for decision-making for assessing the fuzzy situation is written in the form [2]

R: if \( A \rightarrow B \),

\[
\mu_R(a, b) = \mu_{A \rightarrow B}(a, b),
\]

\[
\mu_B(b) = \max(\mu_1(a) \cdot \mu_2(a, b)),
\]

where the \( t \)-operation of the \( t \)-norm.

Here \( R \) - a fuzzy product in which the condition (antecedent) \( A \) and the conclusion (consequent) \( B \) are fuzzy linguistic statements of the following types:

1. \( \langle A \text{ is } \alpha \rangle \), where \( A \) - the name of the linguistic variable, \( \alpha \) - the value \( A' \) is a specific fuzzy subset, which corresponds to a separate linguistic term from the base term set \( T \) of linguistic variable \( A \).

2. \( \langle A \text{ is very } \alpha \rangle \), where \( \alpha \) - the modifier. Type “VERY”, “MORE OR LESS”, “MUCH MORE”, “MUCH LESS”, and others that change values through concentration operations \( (CON(A)) \) or \( (DIL(A)) \).

3. A complex statement formed from statements of types 1 and 2 and fuzzy logical operations in the form of connectives: AND, OR, IF – THEN “NOT”: as situations from the definition of statements of type 2 they are through operations \( (CON(A)) \) or \( (DIL(A)) \) are reduced to statements of type 1. Therefore, statements of type 3 composed of statements of type 1 are usually considered. If we take into account that the \( \text{AND} / \text{OR} \) operations, and the operations of fuzzy negation, are modifiers, then usually statements with compound \( \text{AND} / \text{OR} \) (respectively, operations unification \( (\cup) \) and intersection \( (\cap) \) of fuzzy sets.

One of the main procedures for generating fuzzy conclusions with specific values of conclusions is the operation of calculating the values of membership functions \( \mu_B(b) \) and \( \mu_R(a, b) = \mu_{A \rightarrow B}(a, b) \). The latter are determined by specific types of expressions for operations of the triangular norm and norm, the rules of fuzzy compilation [1-2].
As noted above, the central component, the core of the fuzzy inference system (FIS), is the production system, which consists of a set of production rules. These rules reflect the heuristic ideas of experts about the problem (situation, task) of the given variable in the field. Therefore, the totality of such rules is considered as a set of fuzzy models of the studied problems. Together with the fuzzy inference rules, these rules form FIS, which is a subsystem for decision-making in decision support systems (DDSS) [3].

The formalized presentation of standard methods, rules and mechanisms of fuzzy inference in the form of a systematic set of strategic offensive arms models is one of the main theoretical and methodological problems of constructing support systems for making poorly structured decisions (SSMPSD) [3].

In general terms FIS, by analogy with the generalized model of intellectual DDSS [2, 3], can be represented by the following set of declarative and procedural objects

\[ \text{FIS} = \langle A, L_0, L_H, P_0, p_H, p_D, p_D^0, R(p_D^0), R(P), R(P_D), R(F) \rangle, \]

where the \( A \)-finite alphabet set of identifiers of conditions and conclusions in the used linguistic statements \( (L) \) and production rules \( (P) \);

- \( L_0 \) – current (or initial) set \( L \);
- \( L_H \) – the set of admissible and possible \( L \) for a given problem area (task);
- \( p_0 \) – current (initial) set \( P \);
- \( p_H \) – set of admissible and possible \( P \);
- \( p_D^0 \) – current (initial) set of strategies (fuzzy inference mechanisms) for the search and formation of solutions \( D \);
- \( p_D^H \) – the set of acceptable and possible strategies for the search and formation of \( D \);
- \( R(P_D^0) \) – the rule of choosing the current search strategy \( D \) from the set \( p_D^0 \);
- \( R(P) \) – rules for modifying and replenishing the current set \( p_0 \) by adjusting and replenishing it with rules from the set \( p_H \);
- \( R(P_D) \) – rules for modifying and replenishing the current search strategies \( p_D^0 \) by adjusting and replenishing it with rules from the set \( p_D^H \);
- \( R(F) \) – the rules for modifying the FIS model by expanding the alphabet, adjusting and replenishing sets \( L_H, p_H, \) and \( p_D^H \).

The FIS presented in this way reflects the main property of intelligent systems - adaptation to changing situations in the studied dynamic problem domain by replenishing and updating the values and subject area, adjusting and modifying solution search strategies depending on the nature of the changing conditions and goals of the tasks and other procedures [1].

The models, rules and mechanisms of fuzzy inference systems (FIS) considered in the previous section occupy the main place in the knowledge database as well as models of the inference mechanism and simulation of the operational subsystem of DSS, oriented to support the adoption of poorly structured decisions in the face of uncertainty. For such intelligent DSS we will use the abbreviation DDSS [2-3].

In monitoring and control systems of complex dynamic processes, decisions are formed and made in real time, i.e. in conditions of its deficit. This requires maximum time reduction for the implementation of the entire set of procedures for processing the initial information, fuzzy inference mechanisms, and the formation of recommendations for making decisions.
This implies the relevance of the task of developing theoretical and methodological tools providing automation of the following processes in FIS and DDSS:

- designing and forming a set of production rules (PR) from a given set of permissible, with specific values of conditions and conclusions for describing a fuzzy model of the studied processes and tasks;
- checking the correctness of the constructed rules and bringing them into line with the criteria for the adequacy of fuzzy models;
- building, on the basis PR of the generated functional modules, FIS models and checking their correctness;
- replenishment, adjustment and modification of PR and FIS models;
- Constructing algorithms and programs by using the modules of typical algorithms and programs to solve FIS and making poorly structured decisions based on the constructed FIS.

This problem can be solved by various formalization methods based on the theory of algorithms, automata, fuzzy sets, fuzzy inference, fuzzy correspondences, fuzzy gradients and Petri nets and others.

One promising approach is based on the theoretical and methodological principles of the algorithmization of the study of complex processes and solving related problems [2], as well as methods of algorithmization of the generation and synthesis of algorithms and software systems [3, 4]. Algorithmic systems of this class [2] make it possible to formalize and automate the stages of scientific research in a chain: experience — laws — tasks — mathematical models — algorithms — programs — a computational experiment — analysis of the results and return of the feedback to the corresponding steps for corrections.

This section discusses the main provisions of the proposed conceptualization of algorithmic FIS, as well as the structure and interaction scheme of its functional modules, called algorithmic banks, in the process of solving FIS and SSMPSD [3].

3. The conceptual basis of the algorithmic system of fuzzy inference.

The main provisions and principles of the proposed concept of constructing an algorithmic system of fuzzy inference (AS-FIS) are as follows:

1) a systematic description, in the chosen language of formalism, and an ordered presentation of basic concepts, paradigms; models, methods, and laws describing the poorly formalizable processes under study and related tasks;

2) a formalized representation of the typical shells of FIS in the form of rule constructions and fuzzy inference mechanisms;

3) an orderly presentation of the rules for constructing, modifying, updating, replenishing and checking the correctness of the PR and FIS models;

4) a systematic presentation of typical algorithms and software modules that implement standard procedures for the formation of PR and FIS models, as well as solving FIS and SSMPSD problems;

5) the modular organization of the structure of AS-FIS in the form of a functionally complete set of interacting algorithmic banks (AB), providing automation of the relevant FIS and SSMPSD procedures;

6) the interactive interaction of AS-FIS with the DM at the stages of setting goals and the conditions of the tasks to be solved, as well as analyzing current results and correcting them;

7) automatic interaction (AB) of the system at the stages of solving the problem (computational experiment).

In the course of solving the optimization problem, a set of combinations of linguistic attribute values \( S \subset \Omega \) is searched, which is the optimal solution for the optimization function \( F(S) \) that determines the accuracy of the classification of the data set:

\[
F(S) \rightarrow \max, \text{ where } S \subset \Omega.
\]

The paper considers the process of constructing a compact and interpretable set of fuzzy rules with a high degree of monitoring accuracy, which is an example of solving the multicriteria optimization
problem. A multicriteria genetic algorithm is used to search for optimal solutions for several optimized functions (criteria).

The task of constructing an interpreted fuzzy monitoring model can be formulated as an optimization problem with three objective functions [2]:

\[
\begin{align*}
\text{max } f_1(S), \quad \text{min } f_2(S) \quad \text{and } \quad \text{min } f_3(S),
\end{align*}
\]

(1)

where \( f_1(S) \) – the number of correctly classified objects using the ruleset \( S \), \( f_2(S) \) - the number of fuzzy rules in \( S \), \( f_3(S) \) - the total number of prerequisite elements in \( S \).

The task of constructing an interpreted fuzzy classification model can be solved using the standard genetic algorithm, in this case, three objective functions are combined into one scalar function with the corresponding weight coefficients:

\[
F(S) = w_1 \cdot f_1(S) - w_2 \cdot f_2(S) - w_3 \cdot f_3(S) \rightarrow \text{max},
\]

(2)

where \( w_1, w_2, w_3 \) – real numbers from the interval [0,1] and \( \sum_{i=1}^{3} w_i = 1 \). In this case, the weighting coefficients should be initially determined by the user according to his preferences within the framework of the classification problem to be solved.

The multicriteria genetic algorithm uses the scalar optimization function (2), for which weights \( w_1, w_2, w_3 \) are randomly generated from the interval [0,1] when selecting each new pair of parental individuals. In addition, during the operation of the genetic algorithm, non-dominant rule sets are stored separately from the current population and are updated after each stage of generating a new generation of individuals. Below is a brief description of the stages of the genetic algorithm:

Initialization. An initial population of \( N \) sets of fuzzy rules is generated, which are individual individuals of the population, where \( N \) is the size of the population.

Evaluation. The values of three optimized functions for each set of rules in the current population are calculated. A separately stored population consisting of undominated sets of fuzzy rules is updated.

Selection. Repeating the following steps to select \( N \) pairs of rule sets (parental individuals):

1) The weights \( w_1, w_2, w_3 \) are randomly generated.

2) The utility function is calculated for each set of rules according to formula (2). A pair of rule sets is selected using the roulette wheel method, where the probability of selection of each of the sets is calculated as

\[
P(S) = \frac{F(S) - F_{\text{min}}}{\sum_{S} (F(S) - F_{\text{min}})},
\]

where \( F_{\text{min}} \) – the minimum value of the utility function in the current population.

Crossing and mutation. Generate a new set of rules for each pair of selected parent sets using cross and mutation operations with a predefined probability.

The strategy of elitism. A random selection of \( N_{\text{num}} \) non-dominated rule sets from a separately stored and updated population is randomly selected. Thus, a new population of \( N \) individuals is formed.

Steps 2 through 5 are repeated until the algorithm stops (for example, the number of generations).

The choice of cross and mutation operators depends on the encoding of the rule sets.

Each individual individual of the multicriteria genetic algorithm in this paper is an encoded set \( S \) of fuzzy rules, which is a fuzzy monitoring model. Due to the fact that the consequences and weights of individual fuzzy rules can be determined heuristically using a set of training data, only the composition and number of fuzzy classification rules are encoded. Thus, only elements of the premises of fuzzy rules are part of each individual. The fuzzy rule is represented in an individual as a combination of fuzzy sets of premises encoded in numerical form. For example, when each of the
characteristics $i$ is represented using six fuzzy sets with the corresponding linguistic values and each of the fuzzy sets is assigned a number (Table 2), then the fuzzy rule for a data set with five signs can be represented as individuals in the following form (Table 1).
Table 1. Coded representation of a fuzzy rule of type

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ |
|-------|-------|-------|-------|-------|
| 0     | -1    | -1    | -1    | 4     |

Table 2. Coding Linguistic Values of Fuzzy Sets

| Linguistic Value | Code |
|------------------|------|
| Small            | 0    |
| medium small     | 1    |
| Average          | 2    |
| medium small     | 3    |
| Big              | 4    |
| Empty            | -1   |

A set of fuzzy classification rules $S$ is encoded as a string of numbers, where each substring of length $n$ represents a separate fuzzy rule in $S$.

$n$ – the number of signs or elements of the rule background.

Some modifications of the multicriteria algorithm proposed in [4,6] are proposed, namely:

1. An initial population is generated based on available training data.
2. After each generation, all fuzzy rules that do not participate in the classification of the training data set are removed from each individual. This operation does not impair the classification accuracy using a set of fuzzy rules encoded in individuals, and at the same time helps to reduce the number of rules and the total length of premises, i.e. minimizes the objective functions $f_2(S)$ and $f_3(S)$ (expression (1)).

To generate $N$ rules, $N$ objects of training data are randomly selected. For each object $x_i = (x_i^1, x_i^2, ..., x_i^n)$, $i = 1, ..., N$, one fuzzy rule is generated according to the following algorithm:

1) for each input feature $j$, $j = 1, ..., n$ of the data object, we find the membership function $\mu_{k_j}(x_i^j)$ (not taking into account the membership function with the linguistic value "empty") such that $\mu_{k_j}(x_i^j) = \max_{k=1,...,p_j} \{\mu_k(x_i^j)\}$;

2) create a rule with a premise $A = (\mu_{k_1}, ..., \mu_{k_n})$;

3) with a probability of $P=0.5$, we replace each of the background elements with an element with the linguistic meaning "empty";

4) we determine the corollary $C_i$ (class label) and the weight $V_i$ of the rule.

After enumerating all $N$ objects, we get $N$ rules, which are encoded by a string of numbers and define one individual. Thus, the generation of an initial population based on available data allows us to speed up the process of finding optimal solutions.

The advantage of the multicriteria genetic algorithm under consideration is the variable length of the string, which represents one individual. Initially, several individuals (rule sets) are generated that are of equal length. During the operation of the algorithm, the length of individual individuals can vary, which allows us to search for undominated sets of fuzzy rules in the multidimensional space of rule sets of different lengths. Line length is adjusted using a single point crossover operation with different break points for each parent. The crossover operation is applied with a certain probability $P_{cros}$ to each pair of parents selected for reproduction and allows generating new individuals with lengths different from the length of the parent ones.

The mutation operation makes it possible to slightly modify each of the individuals and is used with a certain probability $P_{mut}$. Suppose that the set of rules $S$ is modified. The mutation operation is to generate $\alpha \cdot |S|$ new fuzzy rules, where $|S|$ - the number of fuzzy rules in $S$, $0.1 \leq \alpha \leq 0.5$.

Moreover, one half of the new rules is generated on the basis of training data that were incorrectly
classified using a set of fuzzy classification rules \( S \). The second half of the new rules is generated using genetic operations from the rules available in \( S \). The main steps of the algorithm for modifying the set of rules \( S \) are as follows:

1. The value of the utility function of each fuzzy rule from the set \( S \) is estimated. The value of the utility function of the rule is determined by the number of training objects that are correctly classified by this rule.

2. \( \alpha \cdot |S| \) new rules are generated. Half or newer rules are generated using genetic operations. First, using the “roulette wheel” method, the required number of pairs of parental fuzzy rules is selected. Using the uniform crossover operation, descendants are generated. For each of the fuzzy descendant rules, a mutation operation is applied, which with some probability allows you to modify fuzzy sets of premises. The second part of the new fuzzy rules is generated based on incorrectly classified data objects. All new rules are joined to the ruleset \( S \), forming an extended set \( S^* \).

3. The utility function of each rule in the extended set \( S^* \) is calculated. After that, the fuzzy rules with the lowest value of the utility function are removed from the extended set of \( S^* \) rules.

After performing the above operations, the composition of the fuzzy rules of the original set \( S \) is modified, while their total number does not change. The use of the Michigan algorithm as a mutation operation during the operation of a multicriteria genetic algorithm allows you to speed up the process of searching for individual fuzzy rules that have higher classification accuracy.

4. The results of the simulation.

This section demonstrates the use of the multicriteria genetic algorithm to extract non-dominant sets of fuzzy classification rules and construct the resulting fuzzy monitoring model for two iris and wine datasets from a machine learning data archive.

In both cases, several fuzzy sets were specified for each of the attributes of the studied data sets. Thus, we uniformly partitioned the range of values of each of the attributes with fuzzy sets and additionally considered a fuzzy set corresponding to the linguistic meaning “empty”.

At the beginning of the genetic algorithm, each individual contained a set of 20 fuzzy rules. During the operation of the algorithm, the length of individuals or the number of rules varied. After each generation, rules that did not participate in the classification of the training data set were deleted from the sets of fuzzy rules. The entire process of the genetic algorithm was repeated 10 times, using various initial populations. The resulting non-dominated sets of fuzzy rules after each of the 10 starts of the genetic algorithm were compared with each other. Then, among the obtained ten groups of sets, non-dominated sets of fuzzy rules were again selected. That is, when one of the non-dominated sets in one of the groups is dominated by the set from the other group, it is removed from the finite set of resulting non-dominated sets of fuzzy classification rules. Table 3 and Table 4 show several non-dominant sets of fuzzy classification rules obtained after 10 runs of the multicriteria genetic algorithm.

| Table 3. Non-dominant classification rule sets for a dataset – Iris |
|---------------------------------------------------------------|
| Number of rules | The number of correctly classified objects | Correctly Classified Data Objects (%) |
|-----------------|-------------------------------------------|---------------------------------------|
| 2               | 98                                        | 65                                    |
| 3               | 141                                       | 94                                    |
| 5               | 144                                       | 96                                    |
| 7               | 147                                       | 98                                    |
Table 4. Non-dominant classification rule sets for a dataset - Wines

| Number of rules | The number of correctly classified objects | Correctly Classified Data Objects (%) |
|-----------------|-------------------------------------------|---------------------------------------|
| 3               | 155                                       | 87.07                                 |
| 4               | 167                                       | 93.82                                 |
| 7               | 173                                       | 97.19                                 |
| 10              | 177                                       | 99.43                                 |

Due to the fact that the data sets considered above correspond to three classes, in order to obtain high classification accuracy, it is necessary to consider only sets consisting of three or more rules. The selection of a finite set of fuzzy rules from a set of undominated sets depends on the preferences of the end user.

5. Conclusion.
The paper presents the process of constructing a fuzzy monitoring model using a multicriteria genetic algorithm. The stages of work, the method of coding solutions and the genetic operations of a multicriteria genetic algorithm are described. Modifications of the multicriteria algorithm are proposed, which make it possible to accelerate the search for optimal solutions.

The main advantage of using a multi-criteria genetic algorithm to build a fuzzy monitoring model is the ability to generate a large number of undominated sets of fuzzy rules.

Thus, the problem of finding a compromise between the degree of accuracy and interpretability of a fuzzy monitoring model can be solved by the end user by selecting the most suitable set of fuzzy monitoring rules for a particular task.

6. References

[1] Rutkovskaya D., Pliinsky M., Rutkovsky L. 2004 Neural networks, genetic algorithms and fuzzy systems. 452
[2] Aliev R.A., Aliev R. 2001 Theory of Intelligent Systems and Applications. 720
[3] Zaychenko Yu. 2006 The Fuzzy Group Method of Data Handling and Its Application for Economical Processes forecasting. 7 83
[4] Gladkov L.A., Kureichik V.V., Kureichik V.M. 2006 Genetic Algorithms. 320
[5] Kondrashina E.Yu., Litvintseva L.V., Pospelov D.A. 1989 Representation of knowledge about time and space in intelligent systems. 6 328
[6] Darrel Whitley. 1993 A Genetic Algorithm Tutorial. 93
[7] Muxamediyeva D.K. 2019 The property of the problem of reaction diffusion with double nonlinearity at the given initial conditions. 9 1095
[8] Muxamediyeva D.K. 2019 Methods for solving the problem of the biological population in the two-case. 1210
[9] Muhamediyeva D.T. and Niyozmatova N.A. 2019 Problems of constructing models of intellectual analysis of states of weakly formalizable processes. 1210
[10] Muhamediyeva D.K. 2019 Diffusion model solutions with double nonlinearity. 1260
[11] Muhamediyeva D.T. and Niyozmatova N.A. 2019 Approaches to solving the problem of fuzzy parametric programming in weakly structured objects. 1260
[12] Muhamediyeva D.T. and Sayfiyev J. 2019 Approaches to the construction of nonlinear models in fuzzy environment. 1260