Control and Cybernetics

vol. **33** (2004) No. 2

A new heuristic algorithm of fuzzy clustering

by

Dmitri A. Viattchenin

National Academy of Sciences of Belarus
GIS Research and Engineering Centre
Surganov St. 6, 220012 Minsk, Belarus
e-mail: viattchenin@mail.ru

Abstract: This paper deals with a new method of fuzzy clustering. The basic concepts of the method are introduced as resulting from the consideration of the fundamental fuzzy clustering problem. The paper provides the description of the general plan of the algorithm and an illustrative example. An analysis of the experimental results of the method's application to the Anderson's Iris data is carried out. Some preliminary conclusions and the ways of prospective investigations are given.

Keywords: fuzzy clustering, fuzzy tolerance, fuzzy cluster, membership degree, allotment, typical point.

1. Introduction

Some remarks on fuzzy approach to clustering are considered in the first subsection. The second subsection includes a short consideration of the modification to the fuzzy cluster analysis problem.

1.1. Preliminary remarks

To begin with, cluster analysis is a structural approach to solving the problem of object classification without training samples. Clustering methods aim at partitioning of a set of objects into subsets, called clusters, so that the objects belonging to the same cluster are as similar as possible and, vice versa, the objects belonging to different clusters are as dissimilar as possible. Clustering methods are called also automatic classification methods and numerical taxonomy methods. Heuristic, hierarchical, optimization and approximation methods are the main approaches to the cluster analysis problem solving.

Since the fundamental Zadeh's (1965) paper was published, fuzzy set theory has been applied to many areas such as learning, decision-making, control and classification. The idea of fuzzy approach application to clustering problems was proposed by Bellman, Kalaba and Zadeh (1966). Heuristic methods of fuzzy clustering, hierarchical methods of fuzzy clustering and optimization methods of fuzzy clustering were proposed by different researchers. Fuzzy clustering methods are considered at length, for instance, by Höppner, Klawonn and Kruse (1997).

The most common and widespread approach in fuzzy clustering is the optimization approach. It can be mentioned that very interesting methods were proposed by Bezdek (1981) and Pedrycz (1985). Moreover, interesting results were presented by Dunn (1974), Jajuga (1991) and others. However, heuristic algorithms are simple and very effective in many cases, because heuristic algorithms of fuzzy clustering display high level of essential clarity and a low level of complexity. Algorithms of Gitman and Levine (1970), Tamura, Higuchi and Tanaka (1971), Couturier and Fioleau (1997) are very good illustrations of this characterisation.

An outline for a new heuristic fuzzy clustering method was presented in Viattchenin (2002), where a man-machine approach to fuzzy classification was described. The main goal of the present paper is a detailed consideration of the automatic version of the method. For this purpose, a short consideration of the problem statement is presented and the fuzzy modification of the cluster analysis problem is formulated. Some theoretical premises of the allotment among fuzzy clusters (AFC) method are considered. The general plan of the AFC-algorithm is described. Illustrative examples are shown and conclusions are formulated.

1.2. A modification of the fuzzy cluster analysis problem

The traditional optimization methods of fuzzy clustering are based on the concept of fuzzy partition. For example, the approaches of Bezdek (1981) and Pedrycz (1985) assume that the input data form the objects by attributes matrix. Such information may not be available. Many classical techniques require that the available data consist only of coefficients of pairwise similarity or dissimilarity between objects. Very few algorithms have been developed to produce partition matrices from this type of input data. Algorithms of Ruspini (1970), Roubens (1978) and Windham (1985) are worth noting from this point of view. Thus, the set $X = \{x_1, \ldots, x_n\}$ of n objects represented by either the matrix of similarity coefficients, the matrix of dissimilarity coefficients or the matrix of objects by attributes, should be divided into c fuzzy clusters. Namely, the grade μ_{li}, $1 \leq l \leq c$, $1 \leq i \leq n$ in which an object x_i belongs to the fuzzy cluster A^l should be determined. For each object x_i the grades of membership should satisfy the conditions of a fuzzy partition:

$$\sum_{l=1}^{c} \mu_{li} = 1, \quad 1 \leq i \leq n; \quad 0 \leq \mu_{li} \leq 1, \quad 1 \leq l \leq c \tag{1}$$

In other words, the family of fuzzy sets $P(X) = \{A^l \mid l = \overline{1,c}, c \leq n\}$ is the fuzzy partition of the initial set of objects $X = \{x_1, \ldots, x_n\}$ if condition (1) is met. If, on the other hand, condition

$$\sum_{l=1}^{c} \mu_{li} \geq 1, \quad 1 \leq i \leq n; \quad 0 \leq \mu_{li} \leq 1, \quad 1 \leq l \leq c \tag{2}$$

is met for each object x_i, $1 \leq i \leq n$, then the corresponding family of fuzzy sets $C(X) = \{A^l \mid l = \overline{1,c}, c \leq n\}$ is the fuzzy coverage of the initial set of objects $X = \{x_1, \ldots, x_n\}$.

The concept of fuzzy coverage is used mainly in heuristic fuzzy clustering procedures. As for the concept of fuzzy partition Zadeh (1977) notes that the conditions (1) are very difficult. Moreover, conditions (1) and (2) consider any fuzzy set $A^l \in X$, $1 \leq l \leq c$ as a potential fuzzy cluster and this circumstance can stand in the way of the correct solving of the problem of classification.

Ruspini (1982) notes that fuzzy clustering is a technique of representation of the initial set of objects by fuzzy clusters. The structure of the set of objects can be described by some fuzzy tolerance, that is - a fuzzy binary symmetric reflexive intransitive relation. So, a fuzzy cluster can be understood as some fuzzy subset originated by fuzzy tolerance relation stipulating that the similarity degree of the fuzzy subset elements is not less than some threshold value. In other words, the value of a membership function of each element of the fuzzy cluster is the degree of similarity of the object to the center of fuzzy cluster. The concept of the fuzzy α-cluster satisfies these conditions. The concept is the basis of the method and is introduced in the paper.

Thus, the fuzzy problem formulation in cluster analysis can be defined in general as the problem of finding of the unique representation of the initial set of objects by fuzzy clusters. The concept of representation was used in Viattchenin (2002). However, the notion of representation has specific meaning in pattern recognition. That is why in this paper the term of allotment among fuzzy clusters will be used. The method introduced uses a special definition of the allotment concept and this fuzzy modification of the cluster analysis problem is more general than modifications based on fuzzy partition or fuzzy coverage concepts, because fuzzy partition and fuzzy coverage can be considered as special kinds of allotment. In the essence, an adequate allotment is the allotment, which corresponds to either most natural allocation of objects to fuzzy clusters or to the researcher's opinion about the aims of classification. Detection of a given number of compact and well-separated fuzzy clusters can be considered as the aim of classification.

2. Outline of the approach

The basic concepts of the method are introduced in the first subsection of the section. The basic concepts of the method are illustrated on the data of

Tamura's (1971) short example in the second subsection. The general plan of the AFC-algorithm is considered in the third subsection.

2.1. Basic concepts

Let us consider the conceptual and methodological bases of the method. The concept of fuzzy tolerance is the basis for the concept of fuzzy cluster. That is why definitions of fuzzy tolerance must be considered in the first place. The notions of powerful fuzzy tolerance, feeble fuzzy tolerance and strict feeble fuzzy tolerance were introduced by Viattchenin (1997, 1998). In this context the classical fuzzy tolerance is called usual fuzzy tolerance.

Let $X = \{x_1, \ldots, x_n\}$ be the initial set of elements and $T : X \times X \to [0, 1]$ some binary fuzzy relation on X with $\mu_T(x_i, x_j)$, $\forall x_i, x_j \in X$ being its membership function.

DEFINITION 2.1 *The usual fuzzy tolerance is the fuzzy binary intransitive relation which possesses the symmetricity property*

$$\mu_T(x_i, x_j) = \mu_T(x_j, x_i), \quad \forall x_i, x_j \in X \tag{3}$$

and the reflexivity property

$$\mu_T(x_i, x_i) = 1, \quad \forall x_i \in X. \tag{4}$$

This kind of fuzzy tolerance is denoted by T_2.

DEFINITION 2.2 *The feeble fuzzy tolerance is the fuzzy binary intransitive relation which possesses the symmetricity property (3) and the feeble reflexivity property*

$$\mu_T(x_i, x_j) \leq \mu_T(x_i, x_i), \quad \forall x_i, x_j \in X. \tag{5}$$

This kind of fuzzy tolerance is denoted by T_1.

DEFINITION 2.3 *The strict feeble fuzzy tolerance is the feeble fuzzy tolerance with strict inequality in (5):*

$$\mu_T(x_i, x_j) < \mu_T(x_i, x_i), \quad \forall x_i, x_j \in X. \tag{6}$$

This kind of fuzzy tolerance is denoted by T_0.

DEFINITION 2.4 *The powerful fuzzy tolerance is the fuzzy binary intransitive relation which possesses the symmetricity property (3) and the powerful reflexivity property. The powerful reflexivity property is defined as the condition of reflexivity (4) together with the condition*

$$\mu_T(x_i, x_j) < 1, \quad \forall x_i, x_j \in X, \quad x_i \neq x_j. \tag{7}$$

This kind of fuzzy tolerance is denoted by T_3.

Fuzzy tolerances T_1 and T_0 are subnormal fuzzy relations if the condition:

$$\mu_T(x_i, x_i) < 1, \quad \forall x_i \in X \tag{8}$$

is met.

The kind of the fuzzy tolerance imposed determines the nature of the implied of fuzzy clusters, as demonstrated by Viattchenin (1999). However, the essence of the method does not depend on the kind of fuzzy tolerance. That is why the method herein is described for any fuzzy tolerance T.

Let us consider the general definition of the fuzzy cluster concept, the concept of the fuzzy cluster's typical point and the concept of the fuzzy allotment of objects. The concept of level fuzzy set will be used in the definition of fuzzy cluster concept. The question of level fuzzy sets was considered by Radecki (1977).

The number c of fuzzy clusters can be equal the number of objects, n. This is taken into account in further considerations.

Let $X = \{x_1, \ldots, x_n\}$ be the initial set of objects. Let T be a fuzzy tolerance on X and α be α-level value of T, $\alpha \in (0, 1]$. Columns or lines of fuzzy tolerance matrix are fuzzy sets $\{A^1, \ldots, A^n\}$. Let $\{A^1, \ldots, A^n\}$ be fuzzy sets on X, which are generated by a fuzzy tolerance T.

Definition 2.5 *The α-level fuzzy set $A^l_{(\alpha)} = \{(x_i, \mu_{A^l}(x_i)) \mid \mu_{A^l}(x_i) \geq \alpha, x_i \in X, l \in [1, n]\}$ is fuzzy α-cluster or fuzzy cluster in simple words. So $A^l_{(\alpha)} \subseteq A^l$, $l = 1, \ldots, n$, $\alpha \in (0, 1]$ and μ_{li} is the membership degree of the element $x_i \in X$ for some fuzzy cluster $A^l_{(\alpha)}$, $l \in [1, n]$, $\alpha \in (0, 1]$. Value of α is the tolerance threshold of fuzzy clusters elements.*

The membership degree of the element $x_i \in X$ for some fuzzy cluster $A^l_{(\alpha)}$, $l \in [1, n]$, $\alpha \in (0, 1]$ can be defined as

$$\mu_{li} = \begin{cases} \mu_{A^l}(x_i), & x_i \in A^l_\alpha, \\ 0, & else. \end{cases} \tag{9}$$

where an α-level $A^l_\alpha = \{x_i \in X \mid \mu_{A^l}(x_i) \geq \alpha\}$ of a fuzzy set A^l is the support of the fuzzy cluster $A^l_{(\alpha)}$. So, a condition $A^l_\alpha = Supp\left(A^l_{(\alpha)}\right)$ is met for each fuzzy cluster $A^l_{(\alpha)}$, $l = \overline{1, n}$, $\alpha \in (0, 1]$. The membership degree can be interpreted as a degree of typicality of an element to a fuzzy cluster. The membership function of fuzzy clusters in the sense of definition 2.5 is denoted by μ_{li} and the notation is not changed from the notation of the membership function of fuzzy clusters in the sense (1) and (2). However, fuzzy clusters in the sense of definition 2.5 are different from fuzzy clusters in the sense (1) and (2) from the essential and methodological positions.

In other words, if columns or lines of fuzzy tolerance T matrix are fuzzy sets $\{A^1, \ldots, A^n\}$ on X then fuzzy clusters $\{A^1_{(\alpha)}, \ldots, A^n_{(\alpha)}\}$ are fuzzy subsets of fuzzy sets $\{A^1, \ldots, A^n\}$ for some value α, $\alpha \in (0,1]$. The value zero for a fuzzy set membership function is eqnivalent to non-belonging of an element to a fuzzy set. That is why values of tolerance threshold α are considered in the interval $(0,1]$.

DEFINITION 2.6 *If T is a fuzzy tolerance on X, where X is the set of elements, and $\{A^1_{(\alpha)}, \ldots, A^n_{(\alpha)}\}$ is the family of fuzzy clusters for some $\alpha \in (0,1]$, then the point $\tau^l_e \in A^l_\alpha$, for which*

$$\tau^l_e = \arg\max_{x_i} \mu_{li}, \quad \forall x_i \in A^l_\alpha \tag{10}$$

is called a typical point of the fuzzy cluster $A^l_{(\alpha)}$, $l \in [1,n]$, $\alpha \in (0,1]$.

Obviously, a typical point of a fuzzy cluster does not depend on the value of tolerance threshold. A unique typical point τ^l of some fuzzy cluster $A^l_{(\alpha)}$ can be interpreted as a center of the fuzzy cluster. So, the membership degree can be interpreted as a degree of tolerance of each element to the center of the fuzzy cluster in this case. Moreover, a fuzzy cluster can have several typical points. That is why symbol e is the index of the typical point.

DEFINITION 2.7 *Let $R^\alpha_z(X) = \{A^l_{(\alpha)} \mid l = \overline{1,c}, 2 \leq c \leq n, \alpha \in (0,1]\}$ be a family of fuzzy clusters for some value of tolerance threshold α, $\alpha \in (0,1]$, which are generated by some fuzzy tolerance T on the initial set of elements $X = \{x_1, \ldots, x_n\}$. If a condition*

$$\sum_{l=1}^{c} \mu_{li} > 0, \quad \forall x_i \in X \tag{11}$$

is met for all $A^l_{(\alpha)}$, $l = \overline{1,c}$, $c \leq n$, then the family is the allotment of elements of the set $X = \{x_1, \ldots, x_n\}$ among fuzzy clusters $\{A^l_{(\alpha)}, l = \overline{1,c}, 2 \leq c \leq n\}$ for some value of tolerance threshold α, $\alpha \in (0,1]$.

It should be noted that several allotments $R^\alpha_z(X)$ can exist for some tolerance threshold α, $\alpha \in (0,1]$. That is why symbol z is the index of an allotment.

The condition (11) requires that every object of the set $X = \{x_1, \ldots, x_n\}$ must be classified. In other words, every object x_i, $i = \overline{1,n}$ must be assigned to at least one fuzzy cluster $A^l_{(\alpha)}$, $l = \overline{1,c}$, $c \leq n$ with the membership degree higher than zero. The condition $2 \leq c \leq n$ requires that the number of fuzzy clusters in $R^\alpha_z(X)$ must be more than two. Otherwise, the unique fuzzy cluster will contain all objects with different positive membership degrees.

The concept of allotment is the central point of the method. But the next concept introduced should be paid attention to, as well.

DEFINITION 2.8 *Allotment $R^\alpha_I(X) = \{A^l_{(\alpha)} \mid l = \overline{1,n}, \alpha \in (0,1]\}$ of the set of objects among n fuzzy clusters for some tolerance threshold α is the initial allotment of the set $X = \{x_1, \ldots, x_n\}$.*

In other words, if initial data are represented by a matrix of some fuzzy tolerance T then lines or columns of the matrix are fuzzy sets $A^l \subseteq X$, $l = \overline{1,n}$ and level fuzzy sets $A^l_{(\alpha)}$, $l = \overline{1,n}$, $\alpha \in (0,1]$ are fuzzy clusters. These fuzzy clusters constitute an initial allotment for some tolerance threshold and they can be considered as clustering components.

2.2. Illustrative examples

Let $X = \{x_1, x_2, x_3, x_4, x_5\}$ be the object set and the initial data matrix be as presented in Table 1. The matrix represents a fuzzy powerful tolerance T_3.

Table 1. Matrix of initial data

T_3	x_1	x_2	x_3	x_4	x_5
x_1	1	0.8	0	0.1	0.2
x_2	0.8	1	0.4	0	0.9
x_3	0	0.4	1	0	0
x_4	0.1	0	0	1	0.5
x_5	0.2	0.9	0	0.5	1

Let us consider examples of the basic concepts of the method. Columns or lines of the matrix of fuzzy tolerance T_3 are fuzzy sets

$$\begin{aligned}
A^1 &= \{(x_1,1),(x_2,0.8),(x_3,0),(x_4,0.1),(x_5,0.2)\},\\
A^2 &= \{(x_1,0.8),(x_2,1),(x_3,0.4),(x_4,0),(x_5,0.9)\},\\
A^3 &= \{(x_1,0),(x_2,0.4),(x_3,1),(x_4,0),(x_4,0)\},\\
A^4 &= \{(x_1,0.1),(x_2,0),(x_3,0),(x_4,1),(x_5,0.5)\},\\
A^5 &= \{(x_1,0.2),(x_2,0.9),(x_3,0),(x_4,0.5),(x_5,1)\}.
\end{aligned}$$

So, α-level fuzzy sets

$$\begin{aligned}
A^1_{(0.1)} &= \{(x_1,1),(x_2,0.8),(x_4,0.1),(x_5,0.2)\},\\
A^2_{(0.1)} &= \{(x_1,0.8),(x_2,1),(x_3,0.4),(x_5,0.9)\},\\
A^3_{(0.1)} &= \{(x_2,0.4),(x_3,1)\},\\
A^4_{(0.1)} &= \{(x_1,0.1),(x_4,1),(x_5,0.5)\},\\
A^5_{(0.1)} &= \{(x_1,0.2),(x_2,0.9),(x_4,0.5),(x_5,1)\}
\end{aligned}$$

are fuzzy clusters for the value of tolerance threshold $\alpha = 0.1$. These fuzzy clusters constitute the initial allotment $R^{0.1}_I(X) = \{A^1_{(0.1)}, A^2_{(0.1)}, A^3_{(0.1)}, A^4_{(0.1)}, A^5_{(0.1)}\}$ for the tolerance threshold $\alpha = 0.1$. So, an allotment $R^{0.1}_z(X)$ for the value of tolerance threshold $\alpha = 0.1$ is any family of fuzzy clusters which are elements of the initial allotment $R_I^{0.1}(X)$ for the tolerance threshold $\alpha = 0.1$ and the family of fuzzy clusters should satisfy the conditions of Definition 2.7. For example, the family of fuzzy clusters $R_1^{0.1}(X) = \{A^3_{(0.1)}, A^4_{(0.1)}\}$ can be considered as the allotment of the object set X among two fuzzy clusters. The object x_3 is the center of the fuzzy cluster $A^3_{(0.1)} = \{(x_2, 0.4), (x_3, 1)\}$ because the object is the unique typical point τ^3 of the fuzzy cluster $A^3_{(0.1)}$. The object x_4 is the center of the fuzzy cluster $A^4_{(0.1)} = \{(x_1, 0.1), (x_4, 1), (x_5, 0.5)\}$ because $x_4 = \tau^4$.

2.3. General plan of the AFC-algorithm

Thus, the problem of fuzzy cluster analysis can be defined in general as the problem of discovering the unique allotment $R^*(X)$, resulting from the classification process, which corresponds to either most natural allocation of objects among fuzzy clusters or to the researcher's opinion about classification. In the first case, the number of fuzzy clusters c is not fixed. In the second case, the researcher's opinion determines the kind of the allotment sought and the number of fuzzy clusters c is fixed. The second case is the subject of consideration.

If some allotment $R_z^\alpha(X) = \{A^l_{(\alpha)} \mid l = \overline{1,c}, c \le n, \alpha \in (0,1]\}$ corresponds to the formulation of a concrete problem, then this allotment is an adequate allotment. In particular, if conditions

$$\bigcup_{l=1}^{c} A^l_\alpha = X, \quad A^l_{(\alpha)}, l = 1, \ldots, c, \quad \alpha \in (0,1], \tag{12}$$

and

$$card(A^l_\alpha \cap A^m_\alpha) = 0, \quad \forall A^l_{(\alpha)}, A^m_{(\alpha)}, \quad l \ne m, \quad \alpha \in (0,1] \tag{13}$$

are met for all fuzzy clusters $A^l_{(\alpha)}$, $l = \overline{1,c}$ of some allotment $R_z^\alpha(X) = \{A^l_{(\alpha)} \mid l = \overline{1,c}, c \le n, \alpha \in (0,1]\}$ then the allotment is the adequate allotment. The adequate allotment $R_z^\alpha(X)$ for some value of tolerance threshold α, $\alpha \in (0,1]$ is a family of fuzzy clusters which are elements of the initial allotment $R_I^\alpha(X)$ for the value of α and the family of fuzzy clusters should satisfy the conditions (12) and (13). In other words, the family of supports $\{A^l_\alpha \mid l = \overline{1,c}, c \le n, \alpha \in (0,1]\}$ of fuzzy clusters $A^l_{(\alpha)}$, $l = \overline{1,c}$ of the adequate allotment $R_z^\alpha(X) = \{A^l_{(\alpha)} \mid l = \overline{1,c}, c \le n, \alpha \in (0,1]\}$ is the hard partition of the initial set of objects $X = \{x_1, \ldots, x_n\}$. So, the construction of adequate allotments $R_z^\alpha(X) = \{A^l_{(\alpha)} \mid l = \overline{1,c}, c \le n\}$ for every α, $\alpha \in (0,1]$ is a trivial problem of combinatorics.

The number c of fuzzy clusters in the allotment sought $R^*(X)$ must be fixed. Thus, the problem consists in the selection of the unique adequate allotment $R^*(X)$ from the set $B(c)$ of adequate allotments, $B(c) = \{R_z^\alpha(X)\}$. So, the condition $c = card(R_z^\alpha(X))$, $\forall R_z^\alpha(X) \in B(c)$ must be met. In other words, the set $B(c)$ of adequate allotments is the class of possible solutions of the classification problem.

The selection of the unique adequate allotment $R^*(X)$ from the set $B(c) = \{R^\alpha_z(X)\}$ of adequate allotments must be made on the basis of evaluation of allotments. The criterion

$$F(R^\alpha_z(X), \alpha) = \sum_{l=1}^{c} \frac{1}{n_l} \sum_{i=1}^{n_l} \mu_{li} - \alpha \cdot c, \tag{14}$$

where c is the number of fuzzy clusters and $n_l = card(A^l_\alpha)$, $A^l_{(\alpha)} \in R^\alpha_z(X)$ is the number of elements in the support of the fuzzy cluster $A^l_{(\alpha)}$, can be used for evaluation of allotments.

The maximum of criterion (14) corresponds to the best allotment of objects among c fuzzy clusters. So, the classification problem can be characterized formally as determination of the solution $R^*(X)$ satisfying

$$R^*(X) = \arg \max_{R^\alpha_z(X) \in B(c)} F(R^\alpha_z(X), \alpha) \tag{15}$$

where $B(c) = \{R^\alpha_z(X)\}$ is the set of adequate allotments for the given number c of fuzzy clusters.

The criterion (14) can be considered as the average total membership of objects in fuzzy clusters of the allotment minus $\alpha \cdot c$. The quantity $\alpha \cdot c$ regularizes with respect to the number of clusters c in the allotment $R^\alpha_z(X)$.

There is a five-step procedure of classification:

1 Calculation of α-level values of the fuzzy tolerance T and construction of the initial allotment $R^\alpha_I(X) = \{A^l_\alpha \mid l = \overline{1, n}\}$ for every α-level, $\alpha \in (0, 1]$;

2 Construction of adequate allotments $R^\alpha_z(X) = \{A^l_{(\alpha)} \mid l = \overline{1, c}, c \leq n\}$ for every α, $\alpha \in (0, 1]$;

3 Construction of the set of adequate allotments $B(c) = \{R^\alpha_z(X)\}$ for the given number of fuzzy clusters c and different values of the tolerance threshold α, $\alpha \in (0, 1]$;

4 Calculation of the $F(R^\alpha_z(X), \alpha)$ for every allotment $R^\alpha_z(X)$ from the set $B(c)$;

5 If for some unique allotment $R^*(X)$ from the set $B(c)$ the value of $F(R^\alpha_z(X), \alpha)$ is maximal for a received value of α, $\alpha \in (0, 1]$, then the allotment $R^*(X)$ must be selected from the set $B(c) = \{R^\alpha_z(X)\}$.

The allotment $R^*(X)$ is the main result of the classification process. The value of tolerance threshold α, $\alpha \in (0, 1]$, which corresponds to the allotment $R^*(X)$, is the additional result of classification. The value of α is very important for the interpretation of results from the epistemological position.

3. Experimental results

The process of the AFC-algorithm execution is presented on the data of Tamura's (1971) short example in the first subsection of the section. Experimental results of the AFC-algorithm's application to the Anderson's Iris data are presented in the second subsection.

3.1. The Tamura's example

Let us consider the execution of the AFC-algorithm for the data of Tamura's (1971) short example and for the fixed number of fuzzy cluster $c = 3$. Table 2 presents a fragment of the execution of the AFC-algorithm process plan. The cardinality of adequate allotment of the object set X among fuzzy clusters is two for $\alpha = 0.1$. The adequate allotment of the object set X among three fuzzy clusters does not exist for $\alpha = 0.2$ and $\alpha = 0.4$. The cardinality of adequate allotments is more than three for $\alpha > 0.8$. That is why the execution of the AFC-algorithm process plan is presented for $\alpha = 0.5$ and $\alpha = 0.8$.

Table 2. Fragment of the process of the AFC-algorithm execution

α	Initial allotment, $R_I^{\alpha}(X) = \{A^1_{(\alpha)}, A^2_{(\alpha)}, A^3_{(\alpha)}, A^4_{(\alpha)}, A^5_{(\alpha)}\}$	The adequate allotment, $R_z^{\alpha}(X)$	The value of $F(R_z^{\alpha}(X), \alpha)$
0.5	$A^1_{(0.5)} = \{(x_1, 1), (x_2, 0.8)\}$ $A^2_{(0.5)} = \{(x_1, 0.8), (x_2, 1), (x_5, 0.9)\}$ $A^3_{(0.5)} = \{(x_3, 1)\}$ $A^4_{(0.5)} = \{(x_4, 1), (x_5, 0.5)\}$ $A^5_{(0.5)} = \{(x_2, 0.9), (x_4, 0.5), (x_5, 1)\}$	$R_1^{0.5}(X) = \{A^1_{(0.5)}, A^3_{(0.5)}, A^4_{(0.5)}\}$	1.150
0.8	$A^1_{(0.8)} = \{(x_1, 1), (x_2, 0.8)\}$ $A^2_{(0.8)} = \{(x_1, 0.8), (x_2, 1), (x_5, 0.9)\}$ $A^3_{(0.8)} = \{(x_3, 1)\}$ $A^4_{(0.8)} = \{(x_4, 1)\}$ $A^5_{(0.8)} = \{(x_2, 0.9), (x_5, 1)\}$	$R_1^{0.8}(X) = \{A^2_{(0.8)}, A^3_{(0.8)}, A^4_{(0.8)}\}$	0.500

The first column and the second column of Table 2 correspond to the first step of the AFC-algorithm. The third column of Table 2 corresponds to the second step of the AFC-algorithm. So, the set $B(c = 3) = \{R_1^{0.5}(X), R_1^{0.8}(X)\}$. This is the third step of the AFC-algorithm. The fourth column of Table 2 corresponds to the fourth step of the AFC-algorithm. The value of $F(R_z^{\alpha}(X), \alpha)$ is maximal for the $R_1^{0.5}(X) = \{A^1_{(0.5)}, A^3_{(0.5)}, A^4_{(0.5)}\}$ allotment. So, the allotment $R_1^{0.5}(X) = \{A^1_{(0.5)}, A^3_{(0.5)}, A^4_{(0.5)}\}$ is the result $R^*(X)$ of the classification. The matrix of object assignments is presented in Table 3.

The fuzzy cluster $A^1_{(0.5)} = \{(x_1, 1), (x_2, 0.8)\}$ corresponds to the first class and $x_1 = \tau^1$ (x_1 is the center of the first class). The fuzzy cluster $A^3_{(0.5)} = \{(x_3, 1)\}$ corresponds to the second class. The object x_3 is the center of the second class because $x_3 = \tau^3$. The fuzzy cluster $A^4_{(0.5)} = \{(x_4, 1), (x_5, 0.5)\}$ corresponds to the third class and $x_4 = \tau^4$. That is why x_4 is the center of the third class. These results can be presented by the diagram of Fig. 1.

Analysis of the example by no means provides an adequate test of the AFC-algorithm's performance. A more critical assessment can be made on the basis of its use in the next example.

Table 3. Matrix of object assignments

Object	Membership degree		
	Class 1	Class 2	Class 3
x_1	1.0	0.0	0.0
x_2	0.8	0.0	0.0
x_3	0.0	1.0	0.0
x_4	0.0	0.0	1.0
x_5	0.0	0.0	0.5

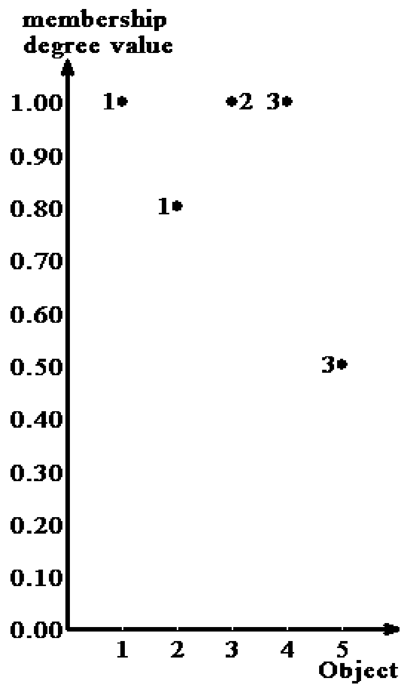

Figure 1. Diagram of object assignment (numbers accompanying points denote clusters)

3.2. The Anderson's Iris data

The Anderson's (1934) Iris data set consists of the sepal length, sepal width, petal length and petal width for 150 irises. The problem is to classify the plants into three subspecies on the basis of this information. The real assignments to the three classes are presented in Table 4.

Table 4. Real objects assignmen in the Iris data set

Class		Numbers of objects
Number	Name	
1	SETOSA	1, 6, 10, 18, 26, 31, 36, 37, 40, 42, 44, 47, 50, 51, 53, 54, 55, 58, 59, 60, 63, 64, 67, 68, 71, 72, 78, 79, 87, 88, 91, 95, 96, 100, 101, 106, 107, 112, 115, 124, 125, 134, 135, 136, 138, 139, 143, 144, 145, 149
2	VERSICOLOR	3, 8, 9, 11, 12, 14, 19, 22, 28, 29, 30, 33, 38, 43, 48, 61, 65, 66, 69, 70, 76, 84, 85, 86, 92, 93, 94, 97, 98, 99, 103, 105, 109, 113, 114, 116, 117, 118, 119, 120, 121, 128, 129, 130, 133, 140, 141, 142, 147, 150
3	VIRGINICA	2, 4, 5, 7, 13, 15, 16, 17, 20, 21, 23, 24, 25, 27, 32, 34, 35, 39, 41, 45, 46, 49, 52, 56, 57, 62, 73, 74, 75, 77, 80, 81, 82, 83, 89, 90, 102, 104, 108, 110, 111, 122, 123, 126, 127, 131, 132, 137, 146, 148

The matrix of attributes is the matrix $X_{m\times n} = [x_i^t]$, $i = 1,\dots,n$, $t = 1,\dots,m$, where $n = 150$, $m = 4$. So, the value x_i^t is the value of the t-th attribute for i-th object. The data can be normalized as follows:

$$\mu_{x_i}(x^t) = \frac{x_i^t}{\max\limits_{x_i} x_i^t}, \quad i = 1,\dots,n, \tag{16}$$

for all attributes x^t, $t = 1,\dots,m$. So, each object can be considered as a fuzzy set x_i, $i = 1,\dots,n$ and $\mu_{x_i}(x^t) \in [0,1]$, $i = 1,\dots,n$, $t = 1,\dots,m$ are their membership functions. After application of the normalized Euclidean distance

$$d_E(x_i,x_j) = \sqrt{\frac{1}{m}\sum_{t=1}^{m}\left(\mu_{x_i}(x^t) - \mu_{x_j}(x^t)\right)^2}, \quad i,j = \overline{1,n}, \tag{17}$$

to the matrix of normalized data $X'_{m\times n} = [\mu_{x_i}(x^t)]$, $i = 1,\dots,n$, $t = 1,\dots,m$ a matrix of a fuzzy intolerance $I = [\mu_I(x_i,x_j)]$, $i,j = 1,\dots,n$ is obtained. The matrix of fuzzy intolerance relation is the matrix of dissimilarity coefficients. The matrix of fuzzy tolerance $T = [\mu_T(x_i,x_j)]$, $i,j = 1,\dots,n$ is obtained after application of complement operation

$$\mu_T(x_i,x_j) = 1 - \mu_I(x_i,x_j), \quad \forall i,j = 1,\dots,n \tag{18}$$

to the matrix of fuzzy intolerance $I = [\mu_I(x_i, x_j)]$, $i, j = 1, \dots, n$.

The object assignments resulting from the AFC-algorithm's application to the Anderson's Iris data are presented in Table 5.

Table 5. The results of AFC-algorithm application: the object assignment

Class		Numbers of objects
Number	Name	
1	SETOSA	1, 6, 10, 18, 26, 31, 36, 37, 40, 42, 44, 47, 50, 51, 53, 54, 55, 58, 59, 60, 63, 64, 67, 68, 71, 72, 78, 79, 87, 88, 91, 95, 96, 100, 101, 106, 107, 112, 115,124, 125, 134, 135, 136, 138, 139, 143, 144, 145, 149
2	VERSICOLOR	3, **5**, 8, 11, 12, 14, 19, 22, **25**, 28, 29, 30, 33, 38, 43, 48, **56**, 61, 65, 66, 69, 70, 76, 84, 85, 86, **90**, 92, 93, 94, 97, 98, 99, 103, 105, 109, 113, 114, 116, 117, 118, 119, 120, 121, 128, 129, 130, 133, 140, 141, 142, 150
3	VIRGINICA	2, 4, 7, **9**, 13, 15, 16, 17, 20, 21, 23, 24, 27, 32, 34, 35, 39,41, 45, 46,49, 52, 57, 62, 73, 74, 75, 77, 80, 81, 82, 83, 89,102,104,108, 110, 111, 122, 123, 126, 127, 131, 132, 137, 146, **147**, 148

By executing the AFC-algorithm for three classes $(1, 2, 3)$ we obtain the following: the first class is formed by 50 elements all being Iris Setosa; the second class by 52 elements, 48 of them being Iris Versicolor and 4 Iris Virginica; the third class by 48 elements, 46 of them being Iris Virginica and 2 Iris Versicolor. In other words, the first class corresponds to the Setosa subspecies, the second class corresponds to the Versicolor subspecies and the third class corresponds to the Virginica subspecies. So, there are six mistakes of classification. Misclassified objects are distinguished in Table 5. Membership values of the Setosa class are presented in Fig. 2. Membership values of the Versicolor class are presented in Fig. 3 and membership values of the Virginica class are presented in Fig. 4. The allotment $R^*(X)$, which corresponds to the result, was obtained for the tolerance threshold $\alpha = 0.811$.

The value of the membership function of the fuzzy cluster which corresponds to the first class is maximal for the seventy-second object and is equal one. So, the seventy-second object is the typical point of the fuzzy cluster which corresponds to the first class. The seventy-second object can be considered as the center of the Setosa class.

The membership value of the ninety-eighth object is equal one for the fuzzy cluster which corresponds to the second class. Thus, the ninety-eighth object is

1

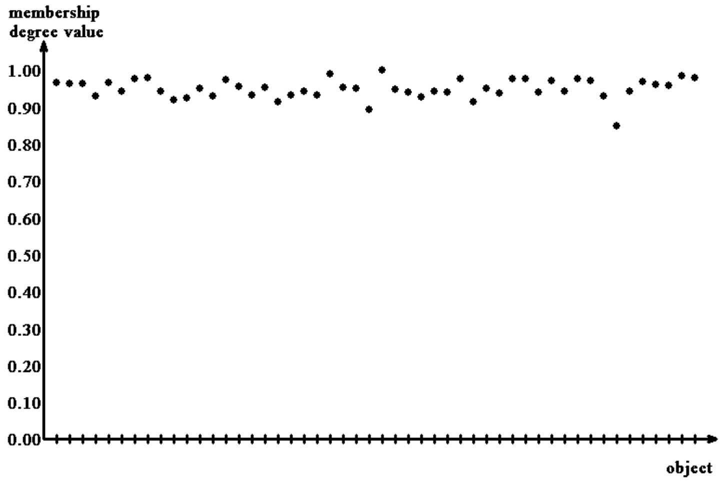

Figure 2. Membership values of the SETOSA class

2

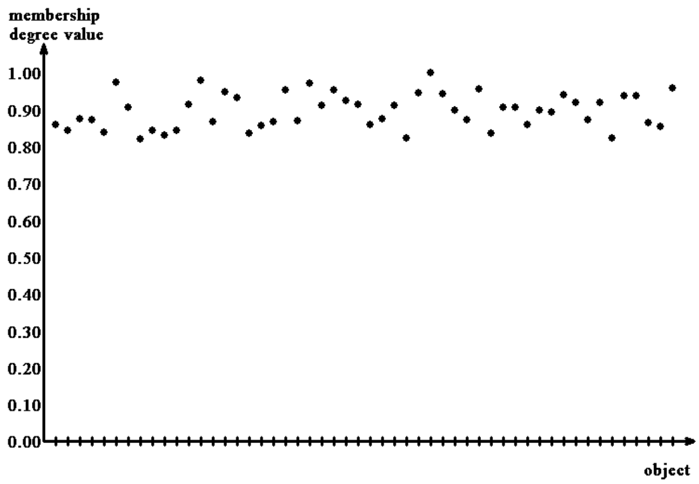

Figure 3. Membership values of the VERSICOLOR class

3

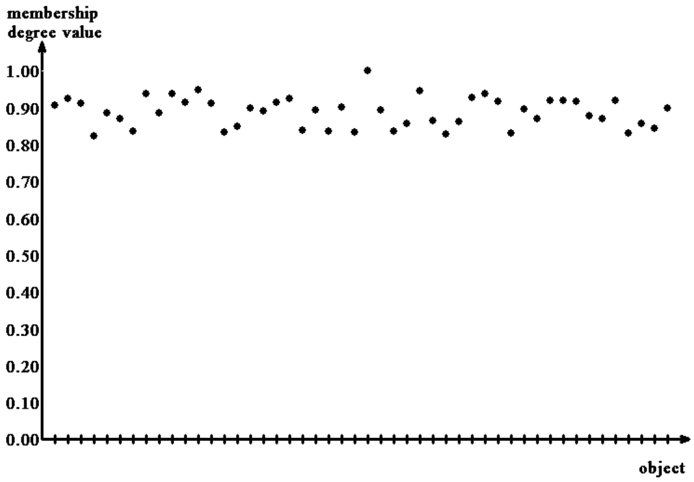

Figure 4. Membership values of the VIRGINICA class

the typical point of the fuzzy cluster which corresponds to the second class. So, the ninety-eighth object can be considered as the center of the Versicolor class.

The membership value of the seventy-third object is equal one for the fuzzy cluster which corresponds to the third class. That is why the seventy-third object is the typical point of the fuzzy cluster which corresponds to the third class and this object can be considered as the center of the Virginica class.

4. Concluding remarks

Preliminary conclusions are discussed in the first subsection of the section. The second subsection deals with the perspectives on future investigations.

4.1. Discussion

In conclusion it should be said that fuzzy clustering methods are very effective in any problem of data analysis. However, Zadeh (1980) notes that uniform theory of pattern classification based on fuzzy sets does not exist and creation of the uniform theory of pattern classification based on fuzzy sets will be a very long process, because conceptual basis must be revised for this purpose. The fuzzy cluster concept and the allotment concept have an epistemological motivation. Thus, the allotment method of fuzzy clustering has a sound justification and can be considered as an attempt of creation of the new conceptual basis and the uniform approach to fuzzy clustering.

Of course, the AFC-algorithm is based on the strong assumption that the number of clusters is fixed. However, the results of the AFC-algorithm application can be very well interpreted. The AFC-algorithm is very simple from the heuristic point of view. Moreover, the objective function-based fuzzy clustering algorithms are sensitive to initialization. Very often, the algorithms are initialized randomly many times, in the hope that some of the initializations lead to good clustering results. The AFC-algorithm is a heuristic fuzzy clustering procedure depending on the set $B(c) = \{R_z^\alpha(X)\}$ of adequate allotments only. That is why the AFC-algorithm clustering results are stable.

The AFC-algorithm can be applied directly to the data given as the matrix of tolerance coefficients. This means that it can be used with the objects by attributes data, by choosing a suitable metric to measure similarity or it can be used in situations where only objects by objects proximity data is available. The results of application of the AFC-algorithm to Anderson's Iris data shows that the AFC-algorithm is a precise and effective numerical procedure for solving classification problems.

4.2. Perspectives

The properties of the criterion (14) are the subject of special theoretical consideration. Some other criteria can be proposed and investigated also.

Other parameters of a clustering procedure can be considered. Firstly, the tolerance threshold can be determined a priori, so that the initial allotments $R_I^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1,n}\}$ and allotments $R_z^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1,c}, c \leq n\}$ are constructed for every α, $\alpha \in [\alpha^*, 1]$, where α^* is the tolerance threshold, defined by the researcher.

Secondly, an intersection of different fuzzy clusters is a natural feature of fuzzy clusters in the sense of conditions (1) and (2). However, fuzzy clusters in the sense of Definition (2.5) can have an intersection area. This fact was demonstrated by Viattchenin (2001). If an intersection area of different fuzzy clusters is an empty set, then fuzzy clusters are called fully separate fuzzy clusters. Otherwise, fuzzy clusters are called particularly separate fuzzy clusters and $w = \{0, \ldots, n\}$ is the maximum number of elements in the intersection area of different fuzzy clusters. Obviously, for $w = 0$ fuzzy clusters are fully separate fuzzy clusters. The maximal number of elements in the intersection area of fuzzy clusters w can be considered as a parameter of the algorithm. So, the conditions (12) and (13) can be generalized and the modification of the AFC-algorithm for particularly separate fuzzy clusters case can be elaborated.

Thirdly, a researcher can determine the minimal number of elements in a fuzzy cluster, too. In other words, if u is a minimal number of elements in a fuzzy cluster, then $card(A_\alpha^l) \geq u$, $\forall l = \overline{1,c}$, where $A_\alpha^l = Supp\left(A_{(\alpha)}^l\right)$ for each fuzzy cluster $A_{(\alpha)}^l$, $l = \overline{1,c}$, $\alpha \in (0,1]$. If the parameter u is not determined, then $u = 1$.

The fuzzy allotment method can be considered as a basis for elaboration of hierarchical fuzzy clustering algorithms. Algorithms of classification on graphs can be elaborated also, because the matrix of fuzzy tolerance can be presented as a weighting graph. These algorithms can be very effective in some particular cases.

These perspectives for investigations are of great interest both from the theoretical point of view and from the practical one as well.

Acknowledgements

I would like to thank the anonymous referees and Dr. Jan W. Owsinski, Editor, for their constructive reviews and useful remarks, Prof. Janusz Kacprzyk, Prof. Eulalia Szmidt, Prof. Yuri S. Kharin and Prof. Eugene E. Zhuk for their interest in my investigations and fruitful discussions during the paper preparation. I also thank the Director of the Systems Research Institute of the Polish Academy of Sciences for the possibility of conducting the investigations and the Mianowski Fund for financial support. I am grateful to Mr. D. Novikau and Mr. D. Dobyshau for elaborating the experimental software.

References

Anderson, E. (1934) The irises of the Gaspe peninsula. *Bulletin of the American Iris Society* **59**, 2–5.

Bellman, R., Kalaba, R. and Zadeh, L.A. (1966) Abstraction and pattern classification. *Journal of Mathematical Analysis and Applications* **13**, 1–7.

Bezdek, J.C. (1981) *Pattern Recognition with Fuzzy Objective Function Algorithms.* New York, Plenum Press.

Couturier, A. *and* Fioleau, B. (1997) Recognizing stable corporate groups: a fuzzy classification method. *Fuzzy Economic Review* **II**, 35–45.

Dunn, J.C. (1974) A fuzzy relative of the ISODATA process and its use in detecting compact, well-separated clusters. *Journal of Cybernetics* **3**, 32–57.

Gitman, I. and Levine, M.D. (1970) An algorithm for detecting unimodal fuzzy sets and its application as a clustering technique. *IEEE Transactions on Computers* **C-19**, 583–593.

Höppner, F., Klawonn, F. and Kruse, R. (1997) *Fuzzy-clusteranalyse verfahren für die bilderkennung, classification und datenanalyse.* Wiesbaden, Vieweg Verlag.

Jajuga, K. (1993) L_1-norm based fuzzy clustering. *Fuzzy Sets and Systems* **39**, 43–50.

Pedrycz, W. (1985) Algorithms of fuzzy clustering with partial supervision. *Pattern Recognition Letters* **3**, 13–20.

RADECKI, T. (1977) Level fuzzy sets. *Journal of Cybernetics* **7**, 189–198.

ROUBENS, M. (1978) Pattern classification problems and fuzzy sets. *Fuzzy Sets and Systems* **1**, 239–253.

RUSPINI, E.H. (1970) Numerical methods for fuzzy clustering. *Information Sciences* **2**, 319–350.

RUSPINI, E.H. (1982) Recent Developments in Fuzzy Clustering. In: R.R.Yager, ed., *Fuzzy Set and Possibility Theory: Recent Developments.* New York, Pergamon Press.

TAMURA, S., HIGUCHI, S. and TANAKA, K. (1971) Pattern Classification Based on Fuzzy Relations. *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-1**, 61-66.

VIATTCHENIN, D.A. (1997) Some remarks to concept of fuzzy similarity relation for fuzzy cluster analysis. In: *Proceedings of the Fourth International Conference PRIP'97* **1**. Szczecin, Wydawnictwo Uczelniane Politechniki Szczecinskiej.

VIATTCHENIN, D.A. (1998) On projections of fuzzy similarity relations. In: *Proceedings of the Fifth International Conference CDAM'98* **2**. Minsk, Belarusian State University.

VIATTCHENIN, D.A. (1999) Remarks on kinds of fuzzy clusters. In: *Proceedings of the International Conference "Quality of Life: Statistical Data Analysis"*. Wroclaw, Wroclaw University of Economics.

VIATTCHENIN, D.A. (2001) On the intersection of fuzzy clusters. In: *Proceedings of International Scientific and Practical Workshop "Integrated Models and Soft Computing in Artificial Intelligence"*. Moscow, Science (in Russian).

VIATTCHENIN, D.A. (2002) Human-computer approach to fuzzy classification problem based on the concept of representation. In: *Proceedings of the Second International Conference "Quality of Life"*. Wroclaw, Wroclaw University of Economics.

WINDHAM, M.R. (1985) Numerical classification of proximity data with assignment measures. *Journal of Classification* **2**, 157–172.

ZADEH, L.A. (1965) Fuzzy Sets. *Information and Control* **8**, 338–353.

ZADEH, L.A. (1977) Fuzzy Sets and Their Application to Pattern Classification and Cluster Analysis. In: J.Van Ryzin, ed., *Classification and Clustering.* New York, Academic Press.