A DIRECT ALGORITHM OF POSSIBILISTIC CLUSTERING WITH PARTIAL SUPERVISION

Received 18th May; accepted 26th June.

Dmitri A. Viattchenin

Abstract:

Fuzzy clustering plays an important role in intelligent systems design and the respective methods constitute a part of the areas of automation and robotics. This paper describes a modification of a direct algorithm of possibilistic clustering that takes into account the information coming from the labeled objects. The clustering method based on the concept of allotment among fuzzy clusters is the basis of the new algorithm. The paper provides the description of basic ideas of the method and the plan of the basic version of a direct possibilistic-clustering algorithm. A plan of modification of the direct possibilistic-clustering algorithm in the presence of information from labeled objects is proposed. An illustrative example of the method's application to the Sneath and Sokal's two-dimensional data in comparison with the Gaussian-clustering method is carried out. Preliminary conclusions are formulated.

Keywords: *clustering, fuzzy tolerance, fuzzy cluster, membership degree, allotment, typical point, labeled object, and partial supervision*

1. Introduction

Some remarks on fuzzy approach to clustering are considered in the first subsection. The second subsection includes a brief review of partially supervised fuzzy clustering methods and the aims of the paper.

1.1 Preliminary remarks

In general, cluster analysis refers to a spectrum of methods, which try to divide a set of objects $X=(x_1,\ldots x_n)$ into subsets, called clusters, which are pair wise disjoint, all non empty and reproduce X *via* union. Heuristic methods, hierarchical methods, optimization methods and approximation methods are used as approaches to the cluster analysis problem solving.

Clustering algorithms can also in general be divided into two types: hard versus fuzzy. Hard clustering assigns each object to exactly one cluster. In fuzzy clustering, founded upon fuzzy set theory [19], a given pattern does not necessarily belong to only one cluster, but can have varying degrees of memberships in several clusters. In heuristic methods of fuzzy clustering different researchers proposed hierarchical methods of fuzzy clustering and optimization methods of fuzzy clustering. These algorithms are described in [15].

The most widespread approach in fuzzy clustering is the optimization approach and the traditional optimization methods of fuzzy clustering are based on the concept of fuzzy partition. The initial set $X=(x_1,\ldots x_n)$ of n objects represented by the matrix of similarity coefficients, the matrix of dissimilarity coefficients or the matrix of object attributes, should be divided into c fuzzy clusters. Namely, the grade μ_{li}, $1 \le l \le c$, $1 \le i \le n$, to which an object x_i belongs to the fuzzy cluster A^l should be determined. For each object x_i, $i=1,\ldots n$ the grades of member-ship should satisfy the conditions of a fuzzy partition:

$$\sum_{l=1}^{c} \mu_{li} = 1, \;\; 1 \le i \le n; \;\; 0 \le \mu_{li} \le 1, \;\; 1 \le l \le c \qquad (1)$$

In other words, the family of fuzzy sets $P(X)=\{A^l \mid l=\overline{1,c}, c \le n\}$ is the fuzzy partition of the initial set of objects $X=(x_1,\ldots x_n)$ if condition (1) is met. Different authors proposed objective function-based fuzzy clustering algorithms, which are considered by Hoeppner, Klawonn, Kruse and Runkler [4].
If, on the other hand, condition

$$\sum_{l=1}^{c} \mu_{li} \ge 1, \;\; 1 \le i \le n; \;\; 0 \le \mu_{li} \le 1, \;\; 1 \le l \le c \qquad (2)$$

is met for each object x_i $1 \le i \le n$, then the corresponding family of fuzzy sets $C(X)=\{A^l \mid l=\overline{1,c}, c \le n\}$ 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.

A possibilistic approach to clustering was proposed by Krishnapuram and Keller [5]. The concept of possibilistic partition is the basis of possibilistic clustering methods and membership values μ_{li}, $1 \le l \le c$, $1 \le i \le n$ can be interpreted as a typicality degree. For each object x_i, $i=1,\ldots n$ the grades of membership should satisfy the conditions of a possibilistic partition:

$$\sum_{l=1}^{c} \mu_{li} > 0, \;\; 1 \le i \le n; \;\; 0 \le \mu_{li} \le 1, \;\; 1 \le l \le c \qquad (3)$$

So, the family of fuzzy sets $Y(X)=\{A^l \mid l=\overline{1,c}, c \le n\}$ is the possibilistic partition of the initial set of objects $X=(x_1,\ldots x_n)$ if condition (3) is met. The possibilistic approach to clustering was developed by Łęski [6], Zhang and Leung [20], Yang and Wu [18] and other researchers. This approach can be considered as a way in the optimization approach in fuzzy clustering because all methods of possibilistic clustering are objective function-based methods.

Heuristic algorithms of fuzzy clustering display high level of essential clarity and low level of a complexity. Some heuristic clustering algorithms are based on a definition of a cluster concept and the aim of these algorithms is cluster detection conform to a given definition. Mandel [8] notes that such algorithms are called algorithms of direct classification or direct clustering algorithms. Direct heuristic algorithms of fuzzy clustering are simple and very effective in many cases.

Fuzzy clustering can be considered as 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 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.

An outline for a new heuristic method of fuzzy clustering is presented by Viattchenin in [16], where concepts of fuzzy α-cluster and allotment among fuzzy α-clusters were introduced and a basic version of direct fuzzy clustering algorithm was described. The basic version of direct fuzzy clustering algorithm requires that the number c of fuzzy α-clusters be fixed. Some modifications of the basic version of the algorithm for different parameters of classification can be elaborated [17]. The version of the algorithm, which is presented in [16] is called the D-AFC(c)-algorithm. Note at this point that the name of AFC-algorithm was used for the fuzzy clustering algorithm, which was proposed by Dave in [3]. The allotment of elements of the set of classified objects among fuzzy clusters can be considered as a special case of possibilistic partition. That is why the D-AFC(c)-algorithm can be considered as a direct algorithm of possibilistic clustering.

1.2 A problem of partially supervised fuzzy clustering

Partially supervised fuzzy clustering plays a unique role in discovering structure in data realized in the presence of labeled patterns. The circumstance is very useful in speech recognition systems and for elaboration of the robot vision systems. Some other problems related to robotics and automation can be successfully solved on the basis of partial-supervised clustering methods.

A priori knowledge about belonging of some objects can be very useful for classification in many cases. This fact was the basis of an approach to fuzzy clustering with partial supervision. Algorithms of fuzzy clustering with partial supervision were proposed by Pedrycz in [9]. Numerical experiments show that knowledge concerning membership of a small portion of the patterns significantly improve clustering results in such a sense that the partition matrix detects a real structure existing in the data set. Moreover, the speed of convergence of the scheme has been improved. These facts are demonstrated by Pedrycz in [10].

Different researchers developed the idea of partial supervision in fuzzy clustering. For example, Bensaid, Hall, Bezdek, and Clarke proposed an original semi-supervised modification of the FCM-algorithm [1]. The method is well suited to problems such as image segmentation. In particular, the procedure was effectively applied to magnetic resonance images segmentation [1]. Very interesting and important results in the area of fuzzy clustering with partial supervision are presented by Bouchachia and Pedrycz in [2].

The main goal of the present paper is consideration of the modification of the D-AFC(c)-algorithm in the case of the presence of labeled objects. For this purpose, an outline of the method of possibilistic clustering based on the concept of allotment of elements of the set of classified objects among fuzzy clusters is presented. A mechanism of partial supervision for the method is proposed and a modification of the algorithm is described. The illustrative examples of application of the proposed method to the Sneath and Sokal's two-dimensional data in comparison with the basic version of the algorithm and the Li and Mukaidono's GCM-algorithm are given. Concluding remarks are stated and perspectives of research work are considered.

2. Outline of the method

Basic concepts of the method and a plan of the basic version of the algorithm are considered in the first subsection. A mechanism of partial supervision for the method and a modification of the algorithm are proposed in the second subsection.

2.1 Basic concepts

Let us recall the basic concepts of the fuzzy clustering method based on the concept of allotment among fuzzy clusters, which was proposed in [16]. The concept of fuzzy tolerance is the basis for the concept of fuzzy α-cluster. That is why definition of fuzzy tolerance must be considered in the first place.

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 $X=(x_1,\ldots x_n)$ with $\mu_T=(x_i, x_j)\in[0,1], \forall x_i, x_j\in X$ being its membership function.

Definition 2.1. *Fuzzy tolerance is the fuzzy binary intransitive relation, which possesses the symmetricity property*

$$\mu_T(x_i,x_j)=\mu_T(x_j,x_i),\ \forall x_i,x_j\in X, \tag{4}$$

and the reflexivity property

$$\mu_T(x_i,x_j)=1,\ \forall x_i\in X. \tag{5}$$

The notions of powerful fuzzy tolerance, feeble fuzzy tolerance and strict feeble fuzzy tolerance were considered in [16], as well. In this context the classical fuzzy tolerance in the sense of Definition 2.1. was called usual fuzzy tolerance and this kind of fuzzy tolerance was denoted by T_2. So, the notions of powerful fuzzy tolerance, feeble fuzzy tolerance and strict feeble fuzzy tolerance must be considered too.

Definition 2.2. *The feeble fuzzy tolerance is the fuzzy binary intransitive relation, which possesses the symmetricity property (4) and the feeble reflexivity property*

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

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 (6):*

$$\mu_T(x_i, x_j) < \mu_T(x_i, x_i),\ \forall x_i, x_j \in X. \quad (7)$$

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 (4) and the powerful reflexivity property. The powerful reflexivity property is defined as the condition of reflexivity (5) together with the condition*

$$\mu_T(x_i, x_j) < 1,\ \forall x_i, x_j \in X, x_i \neq x_j. \quad (8)$$

This kind of fuzzy tolerance is denoted by T_3.

Fuzzy tolerances T_1 and T_0 are subnormal fuzzy relations if the condition $\mu_T(x_i, x_i) < 1,\ \forall x_i \in X.$ is met. The fact was demonstrated in [12]. The kind of the fuzzy tolerance imposed determines the nature of the implied of fuzzy clusters, as demonstrated in [13]. However, the essence of the method here considered 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 fuzzy cluster, the concept of the fuzzy cluster's typical point and the concept of the fuzzy allotment of objects.

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

Let $X=(x_1,...x_n)$ 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 the fuzzy tolerance matrix are fuzzy sets $\{A^1,..., A^n\}$. Let $\{A^1,..., A^n\}$ be fuzzy sets on X, which are generated by a fuzzy tolerance T.

Definition 2.5. *The -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, simply, fuzzy cluster.* *So* $A^l_{(\alpha)} \subseteq A^l,\ \alpha \in (0,1],\ A^l \in \{A^1,\ldots,A^n\}$ *and* μ_{lj} *is the membership degree of the element* $x_l \in X$ *for some fuzzy cluster* $A^l_{(\alpha)},\ \alpha \in (0,1],\ l \in [1,n]$. *Value of α is the toleran-ce threshold of fuzzy clusters elements.*

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

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

where a α-level $A^l_\alpha = \{x_i \in X \mid \mu_{A^l}(x_i) \geq \alpha\}, \alpha \in (0,1]$ of a fuzzy set A^l is the support of the fuzzy cluster $A^l_{(\alpha)}$. So, condition $A^l_\alpha = Supp(A^l_{(\alpha)})$ is met for each fuzzy cluster $A^l_{(\alpha)},\ \alpha \in (0,1],\ l \in [1,n]$. Membership degree can be interpreted as a degree of typicality of an element to a fuzzy cluster. The value of a membership function of each element of the fuzzy cluster in the sense of definition 2.5 is the degree of similarity of the object to some typical object of fuzzy cluster. So, fuzzy clusters in the definition 2.5 are different from fuzzy clusters in the sense (3) from the methodological positions.

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

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

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

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

Obviously, a typical point of a fuzzy cluster does not depend on the value of tolerance threshold. 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 $X=(x_1,...x_n)$. If condition*

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

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 $X=(x_1,...x_n)$ among fuzzy clusters $\{A^l_{(\alpha)}, l = \overline{1,c}, 2 \leq c \leq n\}$ for some value of the tolerance threshold $\alpha, \alpha \in (0,1]$.

It should be noted that several allotments $R^\alpha_z(X)$ could 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 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 $2 \leq l \leq c$ 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 possibly with different positive membership degrees.

Obviously, the definition of the allotment among fuzzy clusters (11) is similar to the definition of the possibilistic partition (3). Moreover, each column $R_{(l)}$, $l=1,...,c$ of the allotment matrix $R_{c\times n}=[\mu_{li}]$, $i=1,...,n$ $l=1,...,c$ can be considered as a possibility distribution on X. So, the allotment among fuzzy clusters can be considered as the possibilistic partition and fuzzy clusters in the sense of definition 2.5 are elements of the possibilistic partition. However, the concept of allotment will be used in further considerations.

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* $R_I^{\alpha}(X)=\{A_{(\alpha)}^l \mid l=\overline{1,n},\alpha\in(0,1]\}$ *of the set of objects among* n *fuzzy clusters for some tolerance threshold* α, $\alpha\in(0,1]$ *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 T then lines or columns of the matrix are fuzzy sets $A^l\subseteq X$, $l=\overline{1,n}$ and level fuzzy sets $A_{(\alpha)}^l$, $l=\overline{1,n}$, $\alpha\in(0,1]$ are fuzzy clusters. These fuzzy clusters constitute an initial allotment for some tolerance threshold and they can be considered as clustering components.

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 can be fixed.

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

$$\bigcup_{l=1}^{c} A_{\alpha}^l = X, \tag{12}$$

and condition

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

are met for all fuzzy clusters $A_{(\alpha)}^l$, $l=\overline{1,c}$ of some allotment $R_z^{\alpha}(X)=\{A_{(\alpha)}^l \mid l=\overline{1,c},c\le n,\alpha\in(0,1]\}$ then the allotment is the allotment among fully separate fuzzy clusters.

However, fuzzy clusters in the sense of definition 2.5 can have an intersection area. This fact was demonstrated in [17]. If the intersection area of any pair of different fuzzy cluster is an empty set, then condition (13) is met and 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.

So, the conditions (12) and (13) can be generalized for a case of particularly separate fuzzy clusters. Condition

$$\sum_{l=1}^{c} card(A_{\alpha}^l)\ge card(X),\quad \forall A_{(\alpha)}^l\in R_z^{\alpha}(X), \tag{14}$$

$$\alpha\in(0,1],\quad card(R_z^{\alpha}(X))=c,$$

and condition

$$card(A_{\alpha}^l\cap A_{\alpha}^m)\le w,\quad \forall A_{(\alpha)}^l, A_{(\alpha)}^m,\quad l\ne m,\quad \alpha\in(0,1], \tag{15}$$

are generalizations of conditions (12) and (13). Obviously, if $w=0$ in conditions (14) and (15) then conditions (12) and (13) are met.

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 $R_z^{\alpha}(X)$ for the value of α and the family of fuzzy clusters should satisfy either the conditions (6) and (7) or the conditions (14) and (15). So, the construction of adequate allotments $R_z^{\alpha}(X)=\{A_{(\alpha)}^l \mid l=\overline{1,c},c\le n,\alpha\in(0,1]\}$ for every α, $\alpha\in(0,1]$ is a trivial problem of combinatorics.

Several adequate allotments can exist. Thus the problem consists in the selection of the unique adequate allotment $R^*(X)$ from the set B of adequate allotments, $B=\{R_z^{\alpha}(X)\}$ which is the class of possible solutions of the concrete classification problem and depends on the parameters the classification problem. The selection of the unique adequate allotment $R^*(X)$ from the set $B=\{R_z^{\alpha}(X)\}$ of adequate allotments must be made on the basis of evaluation of allotments. The criterion

$$F_1(R_z^{\alpha}(X),\alpha)=\sum_{l=1}^{c}\frac{1}{n_l}\sum_{i=1}^{n_l}\mu_{li}-\alpha\cdot c, \tag{16}$$

where c is the number of fuzzy clusters in the allotment $R_z^{\alpha}(X)$ and $n_l=card(A_{\alpha}^l)$, $A_{(\alpha)}^l\in R_z^{\alpha}(X)$ is the number of elements in the support of the fuzzy cluster $A_{(\alpha)}^l$, can be used for evaluation of allotments. The criterion

$$F_2(R_z^{\alpha}(X),\alpha)=\sum_{l=1}^{c}\sum_{i=1}^{n_l}(\mu_{li}-\alpha), \tag{17}$$

can also be used for evaluation of allotments. Both criteria were proposed in [14].

Maximum of criterion (10) or criterion (11) 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

$$R^*(X)=\arg\max_{R_z^{\alpha}(X)\in B} F(R_z^{\alpha}(X),\alpha), \tag{18}$$

where $B=\{R_z^{\alpha}(X)\}$ is the set of adequate allotments corresponding to the formulation of a concrete classification problem and criteria (16) and (17) are denoted by $F(R_z^{\alpha}(X),\alpha)$.

The criterion (16) can be considered as the average total membership of objects in fuzzy clusters of the allotment $R_z^{\alpha}(X)$ minus $\alpha\cdot c$. The quantity $\alpha\cdot c$ regularizes with respect to the number of clusters c in the allotment $R_z^{\alpha}(X)$. The criterion (17) can be considered as the total membership of objects in fuzzy clusters of the allotment $R_z^{\alpha}(X)$ with an appreciation through the value α of tolerance threshold. The condition (18) must be met for the some unique allotment $R_z^{\alpha}(X)\in B(c)$. Otherwise, the number c of fuzzy clusters in the allotment sought $R^*(X)$ is suboptimal.

Detection of fixed c number of fuzzy clusters can be considered as the aim of classification. So, the adequate allotment $R_z^{\alpha}(X)$ is any allotment among c fuzzy clusters in the case. There is the D-AFC(c)-algorithm:

1. Calculate α-level values of the fuzzy tolerance T and construct the sequence $0<\alpha_0<\alpha_1<\ldots<\alpha_l<\ldots<\alpha_Z\leq 1$ of α-levels;
2. Construct the initial allotment $R_I^{\alpha}(X)=\{A_{(\alpha)}^l \mid l=\overline{1,n}\}$, $\alpha=\alpha_l$ for every value α_l from the sequence $0<\alpha_0<\alpha_1<\ldots<\alpha_l<\ldots<\alpha_Z\leq 1$;
3. Let $w:=0$;
4. Construct allotments $R_z^{\alpha}(X)=\{A_{(\alpha)}^l \mid l=\overline{1,c}, c\leq n\}$, $\alpha=\alpha_l$ which satisfy conditions (14) and (15) for every value α_l from the sequence $0<\alpha_0<\alpha_1<\ldots<\alpha_l<\ldots<\alpha_Z\leq 1$;
5. Construct the class of possible solutions of the classification problem $B(c)=\{R_z^{\alpha}(X)\}, \alpha\in\{\alpha_1,\ldots,\alpha_Z\}$ for the given number of fuzzy clusters c and different values of the tolerance threshold α, $\alpha\in\{\alpha_1,\ldots,\alpha_Z\}$ as follows:
   if for some allotment $R_z^{\alpha}(X), \alpha\in\{\alpha_1,\ldots,\alpha_Z\}$ the condition $card(R_z^{\alpha}(X))=c$ is met,
   then $R_z^{\alpha}(X)\in B(c)$
   else let $w:=w+1$ and go to step 4.
6. Calculate the value of some criterion $F(R_z^{\alpha}(X),\alpha)$ for every allotment $R_z^{\alpha}(X)\in B(c)$;
7. The result $R^*(X)$ of classification is formed as follows:
   if for some unique allotment $R_z^{\alpha}(X)$ from the set $B(c)$ the condition (18) is met,
   then the allotment is the result of classification
   else the number c of classes is suboptimal.

The allotment $R_z^{\alpha}(X)=\{A_{(\alpha)}^l \mid l=\overline{1,c}, \alpha\in(0,1]\}$ among the given number of fuzzy clusters and the corresponding value of tolerance threshold α, $\alpha\in(0,1]$ are the results of classification.

2.2. A mechanism of partial supervision

Let us consider a subset of labeled objects $X_L=\{x_{L(1)},\ldots,x_{L(c)}\}$ and $X_L\subset X$. A condition $card(X_L)=c$ must be met for the subset. Let the membership grades $y_{l(j)}$, $l=1,\ldots,c$, $j=1,\ldots,c$ correspond to each labeled object $x_{L(j)}\in X_L$, $j=1,\ldots,c$ as follows: if $x_i\in X_L$ and $x_i=X_{L(j)}$, the values of $y_{l(j)}$ are given by researcher. So, detection of fixed c number of fuzzy clusters can be considered as the aim of classification and each labeled object must be assigned to a unique fuzzy cluster. Moreover, for each labeled object $x_i=X_{L(j)}$ its membership value μ_{li}, $l=1,\ldots,c$, $i=1,\ldots,n$ in the sought allotment $R^*(X)$ must be greater than a priori determined membership grade $y_{l(j)}\in(0,1]$. There is a seven-step procedure of classification:

1. Calculate α-level values of the fuzzy tolerance T and construct the sequence $0<\alpha_0<\alpha_1<\ldots<\alpha_l<\ldots<\alpha_Z\leq 1$ of α-levels;
2. Construct the initial allotment $R_I^{\alpha}(X)=\{A_{(\alpha)}^l \mid l=\overline{1,n}\}$, $\alpha=\alpha_l$ for every value α_l from the sequence $0<\alpha_0<\alpha_1<\ldots<\alpha_l<\ldots<\alpha_Z\leq 1$;
3. Let $w:=0$;
4. Construct allotments $R_z^{\alpha}(X)=\{A_{(\alpha)}^l \mid l=\overline{1,c}, c\leq n\}$, $\alpha=\alpha_l$, which satisfy conditions (14) and (15) for every value α_l from the sequence $0<\alpha_0<\alpha_1<\ldots<\alpha_l<\ldots<\alpha_Z\leq 1$;
5. Construct the class of possible solutions of the classification problem $B(c)=\{R_z^{\alpha}(X)\}, \alpha\in\{\alpha_1,\ldots,\alpha_Z\}$ for the given number of fuzzy clusters c and different values of the tolerance threshold α, $\alpha\in\{\alpha_1,\ldots,\alpha_Z\}$ as follows:
   if for some allotment $R_z^{\alpha}(X), \alpha\in\{\alpha_1,\ldots,\alpha_Z\}$ the condition $card(R_z^{\alpha}(X))=c$ is met
   and for every labeled object $x_{L(j)}=x_i$, $j=\overline{1,c}, i\in\{1,\ldots,n\}$ the condition $\mu_{li}\geq y_{lj}, A_{(\alpha)}^l\in R_z^{\alpha}(X)$, $l=1,\ldots,c$ is met,
   then $R_z^{\alpha}(X)\in B(c)$
   else let $w:=w+1$ and go to step 4.
6. Calculate the value of criterion $F(R_z^{\alpha}(X),\alpha)$ for every allotment $R_z^{\alpha}(X)\in B(c)$;
7. The result $R^*(X)$ of classification is formed as follows:
   if for some unique allotment $R_z^{\alpha}(X)$ from the set $B(c)$ the condition (18) is met,
   then the allotment is the result of classification
   else the number c of classes is suboptimal.

The proposed modification of the D-AFC(c)-algorithm can be called the D-AFC-PS(c)-algorithm. Obviously, the modification does not differ significantly from the basic version of the clustering procedure.

3. An illustrative example

The Sneath and Sokal's two-dimensional data and results of their processing by the GCM-algorithm are considered in the first subsection of the section. Results of three numerical experiments with the proposed procedure are presented in the second subsection.

3.1. The Sneath and Sokal's data

In this section an artificial data set is used for testing of the proposed clustering procedure. These data originally appear as Table 1 in [11] and are shown here in Fig. 1. Li and Mukaidono applied their GCM-algorithm [7] to this data set for the number of classes $c=2$. The results of the GCM application are presented also in Table 1.

Table 1. The Sneath and Sokal's data set and the results of its processing by the GCM-algorithm.

Numbers of objects, i	The data, $x_i=(\hat{x}^1,\hat{x}^2)$		Membership grades obtained from the GCM-algorithm	
	$\hat{x}^1$	$\hat{x}^2$	μ_{1i}	μ_{2i}
1	0	4	1.00	0.00
2	0	3	1.00	0.00
3	1	5	1.00	0.00
4	2	4	0.97	0.03
5	3	3	0.84	0.16
6	2	2	0.98	0.02
7	2	1	0.99	0.01
8	1	0	1.00	0.00
9	5	5	0.05	0.95
10	6	5	0.01	0.99
11	7	6	0.00	1.00
12	5	3	0.07	0.93
13	7	3	0.00	1.00
14	6	2	0.01	0.99
15	6	1	0.01	0.99
16	8	1	0.00	1.00

Obviously, two well-separated classes can be distinguished. In particular, the Gaussian membership function is sharper than the membership function obtained from the FCM-algorithm. The fact was shown by Li and Mukaidono in [7].

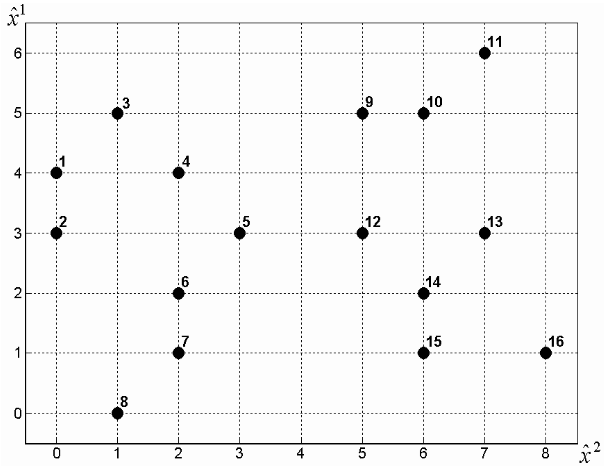

Fig. 1. The Sneath and Sokal's two-dimensional data set.

A diagram can illustrate the matrix of the fuzzy partition. Membership functions of two classes are presented in Fig. 2.

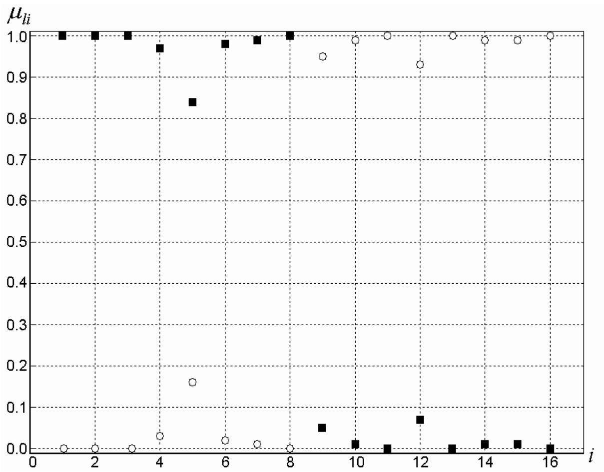

Fig. 2. Membership functions obtained from the GCM-algorithm.

Membership values of the first class are represented in Fig. 2 by ■ and membership values of the second class are represented by ○. The results, obtained from the GCM-algorithm will be very useful for the following considerations.

3.2. Experimental results

Let us consider results of application of the proposed D-AFC-PS(c)-algorithm to the Sneath and Sokal's data set. The matrix of attributes is the matrix $X_{m\times n} = [x_i^t]$, $i=1,\ldots,n, t=1,\ldots,m$, where $n=16$ and $m=2$. So, the value x_i^t is the value of the t-th attribute for the i-th object. The data can be normalized as follows:

$$x_i^t = \frac{x_i^t}{\max\limits_i x_i^t}, i = 1,\ldots,n, \qquad (19)$$

$i = 1,\ldots,n, t = 1,\ldots,m$

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

$$e(x_i, x_j) = \sqrt{\frac{1}{m}\sum_{t=1}^{m}\left(\mu_{x_i}(x^t) - \mu_{x_j}(x^t)\right)^2}, \quad i, j = \overline{1,n}, \qquad (20)$$

to the matrix of normalized data $\hat{X}_{n\times m} = [\mu_{x_i}(x^t)]$, $i = 1,\ldots,n, t = 1,\ldots,m$ the matrix of a fuzzy intolerance $I=[\mu_I(x_i, x_j)]$, $i,j=1,\ldots,n$ is obtained. The matrix $I=[\mu_I(x_i, x_j)]$, $i,j=1,\ldots,n$ is the matrix of pair wise dissimilarity coefficients. The matrix of fuzzy tolerance $T=[\mu_T(x_i, x_j)]$, $i,j=1,\ldots,n$ is obtained after application of the complement operation

$$\mu_T(x_i,x_j)=1-\mu_I(x_i,x_j), \forall i,j=1,\ldots n \qquad (21)$$

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

The results of the data set processing by the D-AFC(c)-algorithm must be considered in the first place. The matrix of the allotment between two fuzzy clusters is presented in Table 2.

Table 2. Results of the Sneath and Sokal's data set classification obtained from the D-AFC(c)-algorithm for $c=2$.

Numbers of objects, i	Membership grades	
	μ_{1i}	μ_{2i}
1	0.88214887	0.00000000
2	1.00000000	0.00000000
3	0.74826988	0.00000000
4	0.78754085	0.00000000
5	0.73483496	0.64644661
6	0.78754085	0.00000000
7	0.70537217	0.00000000
8	0.63556551	0.00000000
9	0.00000000	0.70537217
10	0.00000000	0.74826988
11	0.00000000	0.64644661
12	0.00000000	0.82322330
13	0.00000000	1.00000000
14	0.00000000	0.85268609
15	0.00000000	0.74826988
16	0.00000000	0.74826988

By executing the D-AFC(c)-algorithm for two classes we obtain the following: the first class is formed by 8 elements and the second class is composed of 9 elements. The fifth element belongs to both classes. The allotment $R^*(X)$, which corresponds to the result, was obtained for the tolerance threshold $\alpha = 0.57508171$.

The value of the membership function of the fuzzy cluster, which corresponds, to the first class is maximal for the second object and is equal one. So, the second object is the typical point of the first fuzzy cluster. The membership value of the thirteenth object is equal one for the second fuzzy cluster. Thus, the thirteenth object is the typical point of the second fuzzy cluster. Membership functions of two classes of the allotment are presented in Fig. 3 and values, which equal zero, are not shown in the figure.

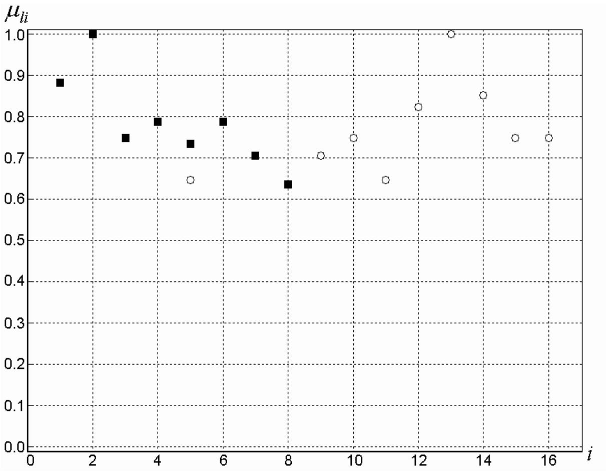

Fig. 3. Membership function obtained from the D-AFC(c)-algorithm for two classes.

Note, that the Gaussian membership function is sharper than the membership function, which is obtained from the D-AFC(c)-algorithm, but the essential interpretation of the results which obtained are from the D-AFC(c)-algorithm is better than in the case of GCM-algorithm.

The results are illustrated also in Fig 4. Supports of two fuzzy clusters are distinguished in Fig. 4 and typical points are denoted by **O**.

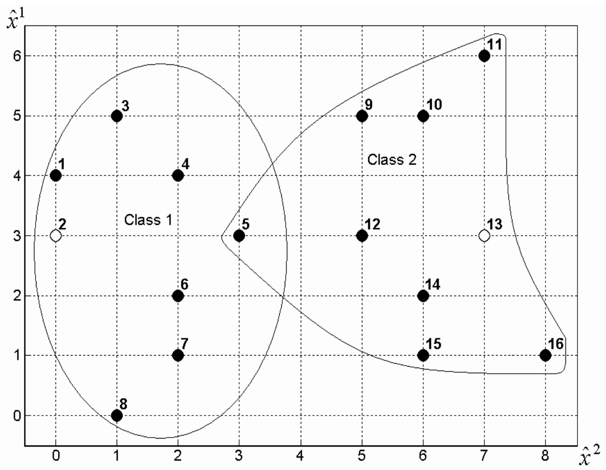

Fig. 4. Supports and typical points of fuzzy clusters obtained from the D-AFC(c)-algorithm for two classes.

Let us consider the results of experiments with the proposed D-AFC-PS(c)-algorithm. The first experiment was made for the set of labeled objects $X_L = \{x_5 = x_{L(1)}, x_9 = x_{L(2)}\}$ with their membership functions $y_{1(5)}=0.6$ and $y_{2(9)}=0.6$. Results of the first experiment are presented in Table 3.

Table 3. Results of the Sneath and Sokal's data set classification obtained from the D-AFC-PS(c)-algorithm in the first experiment.

Numbers of objects, i	Membership grades	
	μ_{1i}	μ_{2i}
1	0.60471529	0.00000000
2	0.70537217	0.00000000
3	0.52038065	0.00000000
4	0.64644661	0.49913270
5	0.74826988	0.50000000
6	0.88214887	0.00000000
7	1.00000000	0.00000000
8	0.85268609	0.00000000
9	0.00000000	0.78754085
10	0.00000000	0.85268609
11	0.00000000	1.00000000
12	0.64522112	0.60471529
13	0.49913270	0.64644661
14	0.62732200	0.52038065
15	0.64644661	0.00000000
16	0.46966991	0.00000000

By executing the D-AFC-PS(c)-algorithm for the set of labeled objects $X_L = \{x_5 = x_{L(1)}, x_9 = x_{L(2)}\}$ with their membership functions $y_{1(5)}=0.6$ and $y_{2(9)}=0.6$ we obtain the following: the first class is formed by 13 elements and the second class is composed of 8 elements. Five objects are elements of both classes. The allotment $R^*(X)$, which corresponds to the result, was obtained for the tolerance threshold $\alpha = 0.46966991$. Membership functions of two classes of the allotment are presented in Fig. 5.

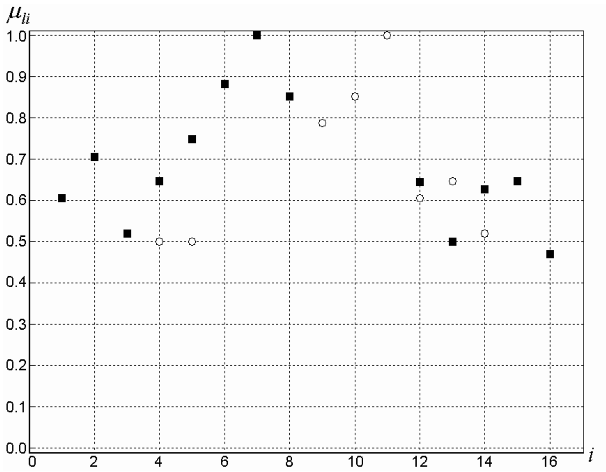

Fig. 5. Membership function obtained from the D-AFC-PS(c)-algorithm in the first experiment.

The value of the membership function of the first fuzzy cluster is equal one for the seventh object. The seventh object is the typical point of the first fuzzy cluster. The membership value of the eleven object is equal one for the second fuzzy cluster. Thus, the eleventh object is the typical point of the fuzzy cluster, which corresponds to the second class. Fig. 6 illustrates supports of fuzzy clusters and their typical points.

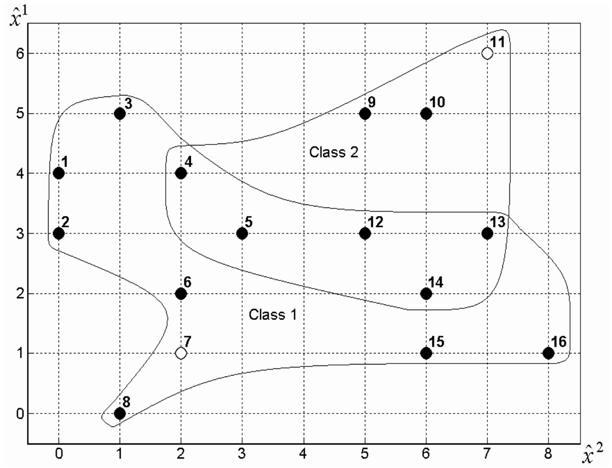

Fig. 6. Supports and typical points of fuzzy clusters obtained from the D-AFC-PS(c)-algorithm in the first experiment.

The second experiment was made for the set of labeled objects $X_L = \{x_7 = x_{L(1)}, x_5 = x_{L(2)}\}$ with their membership functions $y_{1(7)}=0.8$, $y_{2(5)}=0.8$. The results are presented in Table 4.

Table 4. Results of the Sneath and Sokal's data set classification obtained from the D-AFC-PS(c)-algorithm in the second experiment.

Numbers of objects, i	Membership grades	
	μ_{1i}	μ_{2i}
1	0.00000000	0.54261462
2	0.63556551	0.55805826
3	0.00000000	0.57508171
4	0.00000000	0.70982524
5	0.60471529	0.82322330
6	0.74826988	0.70982524
7	0.85268609	0.64522112
8	1.00000000	0.00000000
9	0.00000000	0.76429774
10	0.00000000	0.74826988
11	0.00000000	0.60471529
12	0.00000000	1.00000000
13	0.00000000	0.82322330
14	0.00000000	0.85268609
15	0.54261462	0.74826988
16	0.00000000	0.64522112

By executing the D-AFC-PS(c)-algorithm for the set of labeled objects $X_L = \{x_7 = x_{L(1)}, x_5 = x_{L(2)}\}$ with their membership functions $y_{1(7)}=0.8$ and $y_{2(5)}=0.8$ we obtain the following: the first class is formed by 6 elements and the second class is composed of 15 elements. Five objects are elements of both classes. The allotment $R^*(X)$ was obtained for the tolerance threshold $\alpha = 0.52859548$.

The value of the membership function of the first fuzzy cluster is maximal for the eighth object. So, the eighth object is the typical point of the fuzzy cluster, which corresponds to the first class. The membership value of the twelfth object is equal one for the second fuzzy cluster. The twelfth object is the typical point of the second fuzzy cluster. Membership functions of two classes of the allotment are presented in Fig. 7.

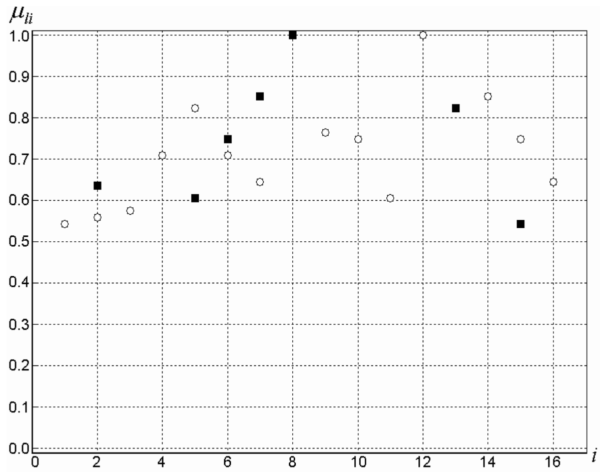

Fig. 7. Membership function obtained from the D-AFC-PS(c)-algorithm in the second experiment.

Fig. 8 illustrates supports of fuzzy clusters and their typical points.

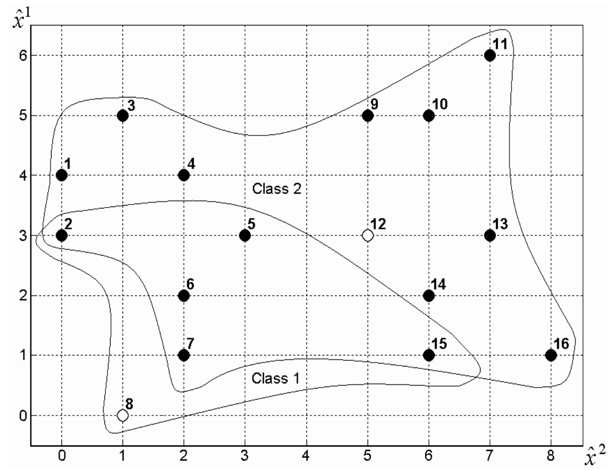

Fig. 8. Supports and typical points of fuzzy clusters obtained from the D-AFC-PS(c)-algorithm in the second experiment.

The third experiment was performed for the set of labeled objects $X_L = \{x_5 = x_{L(1)}, x_{16} = x_{L(2)}\}$ with their membership functions $y_{1(5)}=1.0$ and $y_{2(16)}=1.0$. The results of the experiment are presented in Table 5. Obviously, the labeled objects will be typical points in the sought allotment in this case.

Table 5. Results of the Sneath and Sokal's data set classification obtained from the D-AFC-PS(c)-algorithm in the third experiment.

Numbers of objects, i	Membership grades	
	μ_{1i}	μ_{2i}
1	0.70982524	0.00000000
2	0.73483496	0.00000000
3	0.70537217	0.00000000
4	0.85268609	0.00000000
5	1.00000000	0.00000000
6	0.85268609	0.00000000

Numbers of objects, i	Membership grades	
	μ_{1i}	μ_{2i}
7	0.74826988	0.00000000
8	0.60471529	0.00000000
9	0.70537217	0.00000000
10	0.64522112	0.00000000
11	0.50000000	0.00000000
12	0.82322330	0.64522112
13	0.64644661	0.74826988
14	0.70982524	0.78754085
15	0.64522112	0.82322330
16	0.00000000	1.00000000

By executing the D-AFC-PS(c)-algorithm for the set of labeled objects $X_L = \{x_5 = x_{L(1)}, x_{16} = x_{L(2)}\}$ with their membership functions $y_{1(5)}=1.0$, $y_{2(16)}=1.0$ we obtain that the first class is formed by 15 elements and the second class is composed of 5 elements. Four objects are elements of both classes. Membership functions of two classes are presented in Fig. 9. The allotment $R^*(X)$ was obtained for the tolerance threshold $\alpha = 0.500$.

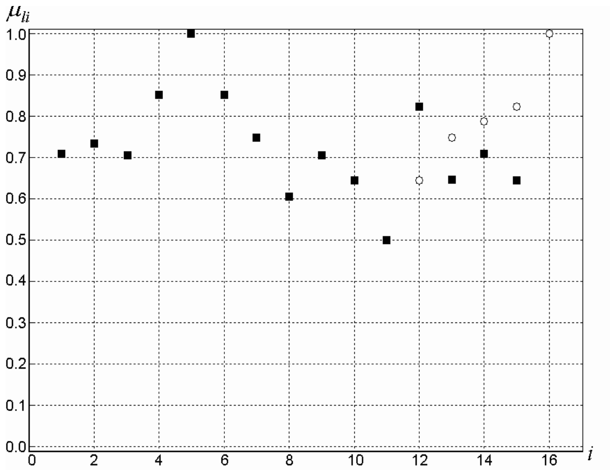

Fig. 9. Membership function obtained from the D-AFC-PS(c)-algorithm in the third experiment.

Fig. 9 illustrates supports of fuzzy clusters and their typical points.

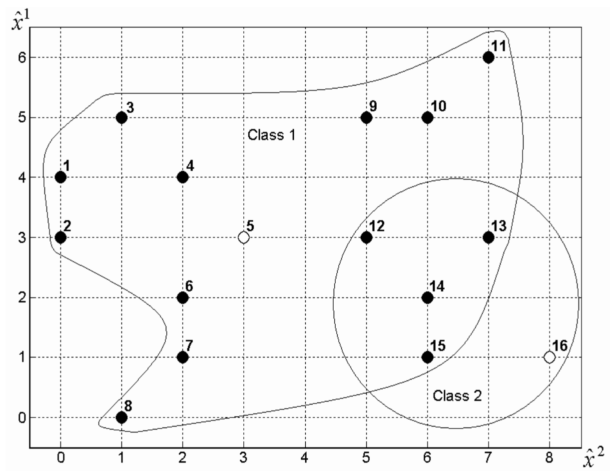

Fig. 10. Supports and typical points of fuzzy clusters obtained from the D-AFC-PS(c)-algorithm in the third experiment.

Membership values of the first class are represented in all figures by ■ and membership values of the second class are represented in all figures by **O**. Results of experiments show that the results of classification depend on the set of labeled objects and their a priori membership functions.

4. Concluding remarks

Some remarks to the results of numerical experiments are made in the first subsection. Perspectives on future investigations are outlined in the second subsection.

4.1. Discussion of the experimental results

In conclusion it should be said that the concept of fuzzy cluster and allotment have an epistemological motivation. That is why the results of application of the fuzzy clustering method based on the allotment concept can be very well interpreted. Moreover, the fuzzy clustering method based on the allotment concept depends on the set of adequate allotments only. That is why the clustering results are stable.

The D-AFC-PS(c)-algorithm of possibilistic clustering is proposed in the paper. The algorithm is based on the mechanism of partial supervision. Numerical experiments show that a result of the D-AFC-PS(c)-algorithm application to the data set depends on the choice of the labeled objects and on their a priori membership functions.

The D-AFC-PS(c)-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 objects by objects proximity data is available. The results of application of the D-AFC-PS(c)-algorithm to the Sneath and Sokal's data set show that the D-AFC-PS(c)-algorithm is a precise and effective numerical procedure for solving classification problem in the case of the presence of labeled objects.

4.2. Perspectives

Given membership functions can be different for different labeled objects. A problem of choosing of the membership function values for the labeled objects must be investigated. Moreover, the method can be extended for the case of presence of a few labeled objects for every class in the sought allotment. These perspectives for investigations are of great interest both from the theoretical point of view and from practical one as well.

ACKNOWLEDGEMENTS

I am grateful to Dr. Jan W. Owsinski and Prof. Janusz Kacprzyk for their useful remarks during the paper preparation. I would like to 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 and Siemens for financial support. I also thank Mr. Aliaksandr Damaratski for elaborating experimental software.

AUTHOR
Dmitri A. Viattchenin - United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Surganov St. 6, 220012 Minsk, Belarus, e-mail: viattchenin@mail.ru.

References

[1] Bensaid A.M., Hall L.O., Bezdek J.C., and Clarke L.P., "Partially supervised clustering for image segmentation", *Pattern Recognition*, vol. 29, no. 5, 1996, pp. 859-871.
[2] Bouchachia A. and Pedrycz W., "Enhancement of fuzzy clustering by mechanisms of partial supervision", *Fuzzy Sets and Systems*, vol. 157, issue 13, 2006, pp. 1733-1759.
[3] Davé R.N., "Use of the adaptive fuzzy clustering algorithm to detect lines in digital images", *Intelligent Robots and Computer Vision*, 1989, vol. 1192, pp. 600-611.
[4] Höppner F., Klawonn F., Kruse R. and Runkler T., *Fuzzy Cluster Analysis: Methods for classification, data analysis and image recognition*, Chichester: Wiley Intersciences, 1998.
[5] Krishnapuram R. and Keller J.M., "A Possibilistic Approach to Clustering", *IEEE Transactions on Fuzzy Systems*, vol. 1, 1993, pp. 98-110.
[6] Łęski J.M., "Robust possibilistic clustering", *Archives of Control Sciences*, vol. 10, 2000, pp. 141-155.
[7] Li R. and Mukaidono M., "Gaussian clustering method based on maximum-fuzzy-entropy interpretation", *Fuzzy Sets and Systems*, vol. 102, 1999, pp. 253-258.
[8] Mandel I.D., *Clustering analysis*, Moscow: "Finansy i Statistica" Publishing House, 1988, (in Russian).
[9] Pedrycz W., "Algorithms of fuzzy clustering with partial supervision", *Pattern Recognition Letters*, vol. 3, 1985, pp. 13-20.
[10] Pedrycz W., "Fuzzy sets in pattern recognition: methodology and methods", *Pattern Recognition*, vol. 23, 1990, pp. 121-146.
[11] Sneath P.H.A. and Sokal R., *Numerical Taxonomy*, San Francisco: Freeman, 1973.
[12] Viattchenin D.A., "On Projections of Fuzzy Similarity Relations", *Proc. of the 5th International Conference on Computer Data Analysis and Modeling CDAM'1998*, Minsk, Belarus, 2003, vol. 2, pp. 91-94.
[13] Viattchenin D.A., "Remarks on Kinds of Fuzzy Clusters", *Proc. of the International Conference "Statistical Data Analysis of Quality of Life"*, Wroclaw, Poland, 1999, pp. 69-79.
[14] Viattchenin D.A., "Criteria of Quality of Allotment in Fuzzy Clustering", *Proc. of the 3rd International Conference on Neural Networks and Artifical Intelligence ICNNAI'2003*, Minsk, Belarus, 2003, pp. 91-94.
[15] Viattchenin D.A., F*uzzy methods of automatic classification*, Minsk: Technoprint Publishing House, 2004, (in Russian).
[16] Viattchenin D.A., "A new heuristic algorithm of fuzzy clustering", C*ontrol & Cybernetics*, vol. 33, 2004, pp. 323-340.
[17] Viattchenin D.A., "Parameters of the AFCmethod of fuzzy clustering", *Bulletin of The Military Academy of The Republic of Belarus*, no. 4, 2004, pp. 51-55, (in Russian).
[18] Yang M.-S. and Wu K.-L., "Unsupervised possibilistic clustering", *Pattern Recognition*, 2006, vol. 39, pp. 5-21.
[19] Zadeh L.A., "Fuzzy Sets", *Information and Contro*l, vol. 8, 1965, pp. 338-353.
[20] Zhang J.-S. and Leung Y.-W., "Improved Possibilistic C-Means Clustering Algorithms", *IEEE Transactions on Fuzzy Systems*, 2004, vol. 12, pp. 209-217.