A Fuzzy Graph-based Heuristic Algorithm of Possibilistic Clustering

Dmitri A. Viattchenin
United Institute of
Informatics Problems
of the NAS of Belarus

Evgeny Nikolaenya
United Institute of
Informatics Problems
of the NAS of Belarus

Aliaksandr Damaratski
United Institute of
Informatics Problems
of the NAS of Belarus

ABSTRACT
In this paper, a heuristic algorithm of possibilistic clustering based on fuzzy graph decomposition is proposed. For the purpose, concepts of fuzzy graph and fuzzy tolerance relation are considered and basic definitions of the heuristic approach to possibilistic clustering are described. An application of the proposed algorithm to the Tamura's portrait data set is provided and some concluding remarks are stated.

General Terms
Pattern Recognition, Clustering, Algorithm.

Keywords
Fuzzy Graph, Fuzzy Tolerance, Heuristic Possibilistic Clustering, Fuzzy Cluster, Allotment, Tolerance Threshold, Typical Point.

1. INTRODUCTION
Some notes on fuzzy approach to cluster analysis are presented in the first subsection of the section. The second subsection includes a brief review of investigations in the area of graph-theoretical approach to clustering.

1.1 Preliminary Remarks
Determining a partition of given sample data is an important part in data analysis tasks. Clustering methods use a mathematical model based on similarity measure to determine a suitable partition of the data set. In fuzzy clustering the data is not only partitioned in a number of clusters, but each object is assigned a degree of membership for each cluster.

Heuristic methods, hierarchical methods and objective function-based methods are main approaches in fuzzy clustering. In objective function-based clustering the mathematical model is stated in form of an objective function that evaluates the partition of data with respect to the membership degrees and the underlying similarity or dissimilarity measure. Different assumptions and constraints lead to a variety of basic clustering concepts. If the objective function is differentiable, necessary conditions for the membership degrees and other cluster parameters used in the distance or similarity measure can be derived in order to optimize the objective function. The resulting equations are then alternatively applied in an algorithm to determine the data fuzzy partition [1].

A possibilistic approach to clustering was proposed by Krishnapuram and Keller [2] and the approach can be considered as a special case of fuzzy approach to clustering because all methods of possibilistic clustering are objective function-based methods. On the other hand, constraints in the possibilistic approach to clustering are less strong than constraints in the fuzzy objective function-based approach to clustering and values of the membership function of a possibilistic partition can be considered as typicality degrees. So, the possibilistic approach to clustering is more general and flexible approach to clustering than the fuzzy approach.

Objective function-based approach in fuzzy clustering is most common and widespread approach. However, heuristic algorithms of fuzzy clustering display low level of complexity and high level of essential clarity. Some heuristic clustering algorithms are based on a definition of the cluster concept and the aim of these algorithms is cluster detection conform to a given definition. Such algorithms are called algorithms of direct classification or direct clustering algorithms [3]. A heuristic approach to possibilistic clustering is proposed in [4].

1.2 Related Works
The problem of finding structures in graphs is an area that has been applied with success to topics as diverse as social networks, the internet and different military applications, to name a few. The problem was considered in detail, for example, by Schaeffer [5] and Fortunato [6].

Since the fundamental Zadeh's [7] paper was published, fuzzy set theory has been applied to many areas and new concepts were introduced. In particular, fuzzy graphs were considered by Kaufmann [8] and Rosenfeld [9]. So, fuzzy graph-based clustering procedures were elaborated by different researchers.

For example, the method of classification developed from the creation of the subclasses up to their fusion which implies the use of a proximity graph built according to a graduated hierarchy [10]. The method was called unsupervised fuzzy graph clustering and the corresponding UFGC-algorithm is proposed in [10].

On the other hand, an algorithm for hierarchical clustering based on fuzzy graph connectedness algorithm is proposed in [11]. The algorithm applies fuzzy et theory to hierarchical clustering method so as to discover clusters with arbitrary shape. It first partitions the data sets into several sub-clusters using a partitioning method, and constructs a fuzzy graph of sub-clusters by analyzing the fuzzy-connectedness degree among sub-clusters. By computing the α-cut graph, the connected components of the fuzzy graph can be obtained, hence resulting the desired clustering.

The main goal of the present paper is the detailed consideration of the fuzzy graph-theoretical clustering procedure in the framework of the heuristic approach to possibilistic clustering. For this purpose, a short consideration

of fuzzy tolerance relations and fuzzy graphs is presented, basic definitions of the heuristic approach to possibilistic clustering are considered, the general plan of the clustering procedure is proposed, an illustrative example is given and preliminary conclusions are formulated.

2. A PROPOSED ALGORITHM FOR FUZZY GRAPH CLUSTERING

Fuzzy tolerance relations and fuzzy graphs are considered in brief in the first subsection of the section. The second subsection includes a consideration of basic concepts of the heuristic approach to possibilistic clustering. A general plan of the proposed FG-AFC-algorithm is presented in the third subsection of the section.

2.1 Fuzzy Tolerances and Fuzzy Graphs

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. Fuzzy tolerance is the fuzzy binary intransitive relation which possesses the symmetry property $\mu_T(x_i, x_j) = \mu_T(x_j, x_i)$, $\forall x_i, x_j \in X$ and the feeble reflexivity property $\mu_T(x_i, x_j) \le \mu_T(x_i, x_i)$, $\forall x_i, x_j \in X$.

The α-level T_{α_ℓ} of the fuzzy relation T is defined as the set of ordered pairs of objects $(x_i, x_j) \in X \times X$ at which $\mu_T(x_i, x_j) \ge \alpha_\ell$, where $0 \le \alpha_\ell \le 1$. In other words, the α-level T_{α_ℓ} of the fuzzy relation T can be defined as $T_{\alpha_\ell} = \{(x_i, x_j) \in X \times X \mid \mu_T(x_i, x_j) \ge \alpha_\ell\}$.

Let X be a finite universe and T be a fuzzy relation on X with $\mu_T(x_i, x_j)$ being its membership function. The α-level fuzzy relation $T_{(\alpha_\ell)}$ of the fuzzy relation T on the universe X is defined as $T_{(\alpha_\ell)} = \{\langle (x_i, x_j) \in T_{\alpha_\ell}, \mu_{T_{(\alpha_\ell)}}(x_i, x_j) = \mu_T(x_i, x_j)\rangle\}$, where $0 < \alpha_\ell \le 1$. So, the membership function $\mu_{T_{(\alpha_\ell)}}(x_i, x_j)$ of the α-level fuzzy relation $T_{(\alpha_\ell)}$ can be defined as follows:

$$\mu_{T_{(\alpha_\ell)}}(x_i, x_j) = \begin{cases} \mu_T(x_i, x_j), & \text{if } \mu_T(x_i, x_j) \ge \alpha_\ell \\ 0, & \text{if } \mu_T(x_i, x_j) < \alpha_\ell \end{cases}, \qquad (1)$$

where $\alpha_\ell \in (0,1]$. The definition of the α-level fuzzy relation $T_{(\alpha_\ell)}$ was proposed in [4].

A fuzzy graph, which was considered by Kaufmann [8] and Rosenfeld [9], is a weighted graph. There is a fuzzy tolerance graph $G = (X, T)$, where $X = \{x_1, \ldots, x_n\}$ is the set of vertices and T is fuzzy tolerance on X.

Let $G = (X, T)$ is a fuzzy tolerance graph, $\mu_T(x_i, x_j)$ is membership function of T, and T_{α_ℓ} is the α-level of T.

So, a usual graph $G_{\alpha_\ell} = (X, T_{\alpha_\ell})$ is the α-level of the fuzzy tolerance graph $G = (X, T)$.

An α-level fuzzy graph $G_{(\alpha_\ell)} = (X, T_{(\alpha_\ell)})$ can be defined as a fuzzy tolerance graph, where $X = \{x_1, \ldots, x_n\}$ is the set of vertices and $T_{(\alpha_\ell)}$ is the α-level fuzzy tolerance on X.

Let $0 < \alpha_0 < \alpha_1 < \ldots < \alpha_\ell < \ldots < \alpha_Z \le 1$ be an ordered sequences of threshold values. Thus, we have an ordered sequence of fuzzy graphs $G_{(\alpha_Z)} \subseteq \ldots \subseteq G_{(\alpha_\ell)} \subseteq \ldots \subseteq G_{(\alpha_1)} \subseteq G_{(\alpha_0)}$, because we have the ordered sequence of fuzzy tolerances $T_{(\alpha_Z)} \subseteq \ldots \subseteq T_{(\alpha_\ell)} \subseteq \ldots \subseteq T_{(\alpha_1)} \subseteq T_{(\alpha_0)}$ and $T_{(\alpha_0)} = T$ [4].

Let $A_{\alpha_\ell}^1, \ldots, A_{\alpha_\ell}^c$ are subsets of vertices, $A^l, A^m \subset X$, $l, m \in \{1, \ldots, c\}$, $A^l \cap A^m = \varnothing$, $l \ne m$, and $\bigcup_{l=1}^c A_{\alpha_\ell}^l = X$.

So, a connected fuzzy sub-graph $G_{(\alpha_\ell)}^l = (A_{\alpha_\ell}^l, T_{(\alpha_\ell)})$ is a component of the α-level fuzzy graph $G_{(\alpha_\ell)} = (X, T_{(\alpha_\ell)})$. That is why a family of connected fuzzy sub-graphs $G_{(\alpha_\ell)}^l = (A_{\alpha_\ell}^l, T_{(\alpha_\ell)})$, $l = 1, \ldots, c$, $\bigcup_{l=1}^c G_{(\alpha_\ell)}^l = G_{(\alpha_\ell)}$ is a classification on the level $\alpha_\ell \in (0,1]$.

2.2 Basic Concepts of the Heuristic Approach to Possibilistic Clustering

Let us remind basic concepts of a heuristic method of possibilistic clustering which was proposed in [4]. The essence of the heuristic approach to possibilistic clustering is that the sought clustering structure of the set of objects is formed based directly on the formal definition of fuzzy cluster and possibilistic memberships are determined also directly from the values of the pair wise similarity of objects.

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 the 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. The α-level fuzzy set $A_{(\alpha)}^l = \{(x_i, \mu_{A^l}(x_i)) \mid \mu_{A^l}(x_i) \ge \alpha\}$, $l \in [1, n]$ is fuzzy α-cluster or, simply, fuzzy cluster. So $A_{(\alpha)}^l \subseteq A^l$, $\alpha \in (0,1]$, $A^l \in \{A^1, \ldots, A^n\}$ and μ_{li} is the membership degree of the element $x_i \in X$ for some fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1, n]$. 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 $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1, n]$ can be defined as a

$$\mu_{li} = \begin{cases} \mu_{A^l}(x_i), & x_i \in A^l_\alpha \\ 0, & otherwise \end{cases}, \qquad (2)$$

where an 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.

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

$$\tau^l_e = \arg\max_{x_i} \mu_{li}, \ \forall x_i \in A^l_\alpha, \qquad (3)$$

is called a typical point of the fuzzy cluster $A^l_{(\alpha)}$, $\alpha \in (0,1]$, $l \in [1,n]$. A fuzzy cluster $A^l_{(\alpha)}$ can have several typical points. That is why symbol e is the index of the typical point.

Let $R^\alpha_{c(z)}(X) = \{A^l_{(\alpha)} \mid l = \overline{1,c},\ 2 \leq c \leq n, \alpha \in (0,1]\}$ 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 $X = \{x_1,\dots,x_n\}$. If a condition

$$\sum_{l=1}^{c} \mu_{li} > 0, \ \forall x_i \in X \qquad (4)$$

is met for all fuzzy clusters $A^l_{(\alpha)} \in R^\alpha_{c(z)}(X)$, $l = \overline{1,c}$, $c \leq n$, then the family is the allotment of elements of the set $X = \{x_1,\dots,x_n\}$ among fuzzy clusters $\{A^l_{(\alpha)}, l = \overline{1,c}, 2 \leq c \leq n\}$ for some value of the tolerance threshold α. It should be noted that several allotments $R^\alpha_{c(z)}(X)$ can exist for some tolerance threshold α. That is why symbol z is the index of an allotment.

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 $\alpha \in (0,1]$ is the initial allotment of the set $X = \{x_1,\dots,x_n\}$. 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 α-level fuzzy sets $A^l_{(\alpha)}$, $l = \overline{1,c}$, $\alpha \in (0,1]$ are fuzzy clusters. These fuzzy clusters constitute an initial allotment for some tolerance threshold α and they can be considered as clustering components.

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

$$\bigcup_{l=1}^{c} A^l_\alpha = X \qquad (5)$$

and a condition

$$card(A^l_\alpha \cap A^m_\alpha) = 0, \ \forall A^l_{(\alpha)}, A^m_{(\alpha)}, \ l \neq m, \ \alpha \in (0,1] \qquad (6)$$

are met for all fuzzy clusters $A^l_{(\alpha)}$, $l = \overline{1,c}$ of some allotment $R^\alpha_{c(z)}(X) = \{A^l_{(\alpha)} \mid l = \overline{1,c}, c \leq n\}$ for a value $\alpha \in (0,1]$, then the allotment is the allotment among fully separate fuzzy clusters.

Fuzzy clusters in the sense of definition (2) can have an intersection area. If the intersection area of any pair of different fuzzy clusters is an empty set, then conditions (5) and (6) are met and fuzzy clusters are called fully separate fuzzy clusters. Otherwise, fuzzy clusters are called particularly separate fuzzy clusters and $w \in \{0,\dots,n\}$ is the maximum number of elements in the intersection area of different fuzzy clusters. For $w = 0$ fuzzy clusters are fully separate fuzzy clusters. Thus, the conditions (5) and (6) can be generalized for a case of particularly separate fuzzy clusters. So, a condition

$$\sum_{l=1}^{c} card(A^l_\alpha) \geq card(X), \ \forall A^l_{(\alpha)} \in R^\alpha_{c(z)}(X),$$
$$\alpha \in (0,1], \ card(R^\alpha_{c(z)}(X)) = c \qquad (7)$$

and a condition

$$card(A^l_\alpha \cap A^m_\alpha) \leq w, \ \forall A^l_{(\alpha)}, A^m_{(\alpha)},$$
$$l \neq m, \alpha \in (0,1] \qquad (8)$$

are generalizations of conditions (5) and (6). Obviously, if $w = 0$ in conditions (7) and (8) then conditions (5) and (6) are met.

The adequate allotment $R^\alpha_{c(z)}(X)$ for some value of tolerance threshold $\alpha \in (0,1]$ is a family of fuzzy clusters which are elements of the initial allotment $R^\alpha_I(X)$ for the value of α and the family of fuzzy clusters should satisfy the conditions (7) and (8). So, the construction of adequate allotments $R^\alpha_{c(z)}(X) = \{A^l_{(\alpha)} \mid l = \overline{1,c}, c \leq n\}$ for every α is a trivial problem of combinatorics.

Allotment $R^\alpha_P(X) = \{A^l_{(\alpha)} \mid l = \overline{1,c}\}$ of the set of objects among the minimal number c, $2 \leq c \leq n$ of fully separate fuzzy clusters for some tolerance threshold $\alpha \in (0,1]$ is the principal allotment of the set $X = \{x_1,\dots,x_n\}$.

Several adequate allotments can exist. Thus, the problem consists in the selection of the unique adequate allotment $R^*_c(X)$ from the set B of adequate allotments, $B = \{R^\alpha_{c(z)}(X)\}$, which is the class of possible solutions of the concrete classification problem. The selection of the unique adequate allotment $R^*_c(X)$ from the set

$B = \{R_{c(z)}^{\alpha}(X)\}$ of adequate allotments must be made on the basis of evaluation of allotments. In particular, the criterion

$$F(R_{c(z)}^{\alpha}(X), \alpha) = \sum_{l=1}^{c} \frac{1}{n_l} \sum_{i=1}^{n_l} \mu_{li} - \alpha \cdot c \qquad (9)$$

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

Maximum of criterion (9) corresponds to the best allotment of objects among c fuzzy clusters. So, the classification problem can be characterized formally as determination of the solution $R_c^*(X)$ satisfying

$$R_c^*(X) = \arg \max_{R_{c(z)}^{\alpha}(X) \in B} F(R_{c(z)}^{\alpha}(X), \alpha). \qquad (10)$$

The problem of cluster analysis can be defined in general as the problem of discovering the unique allotment $R_c^*(X)$, resulting from the classification process and detection of fixed or unknown number c of fuzzy clusters can be considered as the aim of classification.

2.3 A Plan of the FG-AFC-Algorithm

Let introduce some important concepts. Let $G_{(\alpha_\ell)} = (X, T_{(\alpha_\ell)})$ is the α-level fuzzy graph for some value of tolerance threshold $\alpha_\ell \in (0,1]$ and fuzzy sub-graphs $G_{(\alpha_\ell)}^{l} = (A_{\alpha_\ell}^{l}, T_{(\alpha_\ell)})$, $l = 1, \dots, c$ are components of the α-level fuzzy graph $G_{(\alpha_\ell)} = (X, T_{(\alpha_\ell)})$. So, the condition (5) and the condition (6) are met.

On the other hand, let $R_c^{(\alpha_\ell)*}(X) = \{A_{(\alpha_\ell)}^{1}, \dots, A_{(\alpha_\ell)}^{c}\}$ be the unique adequate allotment among c fully separate fuzzy clusters for the value $\alpha_\ell \in (0,1]$. Obviously, that a condition $card(R_c^*(X)) = card(\{A_{\alpha_\ell}^{l}\}_{l=1}^{c})$ is met. If a condition $card(Supp(A_{(\alpha_\ell)}^{l})) = card(A_{\alpha_\ell}^{l})$, $\forall l \in \{1, \dots, c\}$ is met, then the allotment $R_c^{(\alpha_\ell)*}(X)$ is the corresponding allotment for the α-level fuzzy graph $G_{(\alpha_\ell)} = (X, T_{(\alpha_\ell)})$.

Let $G_{(\alpha_\ell)}^{l} = (A_{\alpha_\ell}^{l}, T_{(\alpha_\ell)})$, $l \in \{1, \dots, c\}$ is a component of the α-level fuzzy graph $G_{(\alpha_\ell)} = (X, T_{(\alpha_\ell)})$ for some value of tolerance threshold $\alpha_\ell \in (0,1]$. Obviously, that a fuzzy cluster $A_{(\alpha_\ell)}^{l} \in R_c^{(\alpha_\ell)*}(X)$ corresponds to the component $G_{(\alpha_\ell)}^{l} = (A_{\alpha_\ell}^{l}, T_{(\alpha_\ell)})$. Thus, constructing α-level fuzzy graphs $G_{(\alpha_\ell)}$ and corresponding allotments $R_c^{(\alpha_\ell)*}(X)$ for different values of tolerance threshold $\alpha_\ell \in (0,1]$ is the aim of classification.

The corresponding FG-AFC-algorithm for detecting components of α-level fuzzy graphs $G_{(\alpha_\ell)}$ and the corresponding allotments $R_c^{(\alpha_\ell)*}(X)$ is a six-step procedure of classification.

1. Calculate α-level values of the fuzzy tolerance T and construct the sequence $0 < \alpha_0 < \alpha_1 < \dots < \alpha_\ell < \dots < \alpha_Z \leq 1$ of α-levels; set $\ell := 0$;

2. Construct the α-level fuzzy graph $G_{(\alpha_\ell)}$;

3. The following condition is checked:

   if a condition $\ell < Z$ is met

   then the following condition is checked:

   > **if** a condition is met $\mu_T(x_i, x_j) < \alpha_\ell$ for some arc (x_i, x_j), $i, j \in \{1, \dots, n\}$

   > **then** the arc should be deleted from the fuzzy graph $G_{(\alpha_\ell)}$

   else stop;

4. The following condition is checked:

   if the fuzzy graph $G_{(\alpha_\ell)}$ can be divided into unknown number c of components $G_{(\alpha_\ell)}^{l} = (A_{\alpha_\ell}^{l}, T_{(\alpha_\ell)})$, $l = 1, \dots, c$

   then go to step 5

   else set $\ell := \ell + 1$ and go to step 2;

5. Construct allotment $R_c^{(\alpha_\ell)*}(X)$ among unknown number c of fully separated fuzzy clusters as follows:

   5.1 Set $l = 1$

   5.2 Construct fuzzy sub-relation $T_{(\alpha_\ell)}^{l}$ for a component $G_{(\alpha_\ell)}^{l} = (A_{\alpha_\ell}^{l}, T_{(\alpha_\ell)})$, $l \in \{1, \dots, c\}$

   5.3 Columns or lines of the fuzzy sub-relation $T_{(\alpha_\ell)}^{l}$ matrix are fuzzy sets $\{A^{l(1)}, \dots, A^{l(j)}\}$ on $A_{\alpha_\ell}^{l} \subset X$;

   5.4 Calculating the cardinality $\mathcal{F}(A^{l(k)})$ of each fuzzy set $A^{l(k)}$, $k \in \{1, \dots, j\}$ according to a formula $\mathcal{F}(A^{l(k)}) = \sum_{i=1}^{n} \mu_{A^{l(k)}}(x_i)$, $k \in \{1, \dots, j\}$, $x_i \in A_{\alpha_\ell}^{l}$;

   5.5 The following condition is checked:

if for some unique fuzzy set $A^{l(k)}$ a condition $A^{l(j)} = \arg\max_{A^{l(k)}} \mathcal{F}(A^{l(k)})$ is met

then the fuzzy set $A^{l(k)}$ should be selected as a basic fuzzy set for constructing the fuzzy cluster $A_{(\alpha_\ell)}^l$ and the element of the basic fuzzy set $A^{l(j)}$ for which $\tau^l = \arg\max_{x_i} \mu_{A^{l(k)}}(x_i)$, $\forall x_i \in A_{\alpha_\ell}^l$ is the typical point of the fuzzy cluster $A_{(\alpha_\ell)}^l$

else the fuzzy cluster $A_{(\alpha_\ell)}^l$ is not exists and

stop;

5.6 The membership degree of the element $x_i \in A_{\alpha_\ell}^l$ for the fuzzy cluster $A_{(\alpha_\ell)}^l$ should be constructed as follows:

if the element $x_i \in A_{\alpha_\ell}^l$ is a vertex of the shortest path (τ^l, x_i)

then the value of the membership degree μ_{li} of

the element $x_i \in A_{\alpha_\ell}^l$ for the fuzzy cluster $A_{(\alpha_\ell)}^l$ is equal to the value of minimal value of

the membership function $\mu_{T_{(\alpha_\ell)}^l}(x_i, x_j)$ of arc

from the shortest path (τ^l, x_i);

5.7 The following condition is checked:

if the membership degree assigned to each element $x_i \in A_{\alpha_\ell}^l$

then the fuzzy cluster $A_{(\alpha_\ell)}^l$ is formed and $A_{(\alpha_\ell)}^l \in R_c^{(\alpha_\ell)*}(X)$

else if all elements $x_i \in A_{\alpha_\ell}^l$ are not verified

then $l := l + 1$ and go to step 5.2

else go to step 6;

6. The following condition is checked:

if the unique allotment $R_c^{(\alpha_\ell)*}(X)$ which correspond to the α-level fuzzy graph $G_{(\alpha_\ell)}$ is constructed

then the α-level fuzzy graph $G_{(\alpha_\ell)}$ and the corresponding allotment $R_c^{(\alpha_\ell)*}(X)$ are a solution of the classification problem for the value α_ℓ; set $\ell := \ell + 1$ and go to step 2;

else the α-level fuzzy graph $G_{(\alpha_{\ell-1})}$ and the corresponding allotment $R_c^{(\alpha_{\ell-1})*}(X)$ for the value $\alpha_{\ell-1}$ are resulting solution of the classification problem and stop.

It should be noted, that the set B of adequate allotments is not constructed in the FG-AFC-algorithm and the criterion (9) is not used in the clustering procedure.

So, a family of α-level fuzzy graph $G_{(\alpha_\ell)} = (X, T_{(\alpha_\ell)})$ and corresponding allotments $R_c^{(\alpha_\ell)*}(X)$ for different values of the tolerance threshold $\alpha_\ell \in (0,1]$ is a result of classification process.

The proposed FG-AFC-algorithm should be explained by illustrative examples.

3. EXPERIMENTAL RESULTS

A simple example is considered in the first subsection of the section. The second subsection includes a consideration of application of the proposed FG-AFC-algorithm to Tamura's portrait data set.

3.1 A Simple Illustrative Example

Let us consider a simple numerical example [12]. Let $X = \{x_1, \ldots, x_5\}$ be the initial set of elements and T is a fuzzy tolerance relation on X. So, $G = (X, T)$ is a fuzzy tolerance graph and $\mu_T(x_i, x_j)$, $i, j = 1, \ldots, 5$ is membership function of T. The fuzzy graph $G = (X, T)$ is shown in Fig. 1.
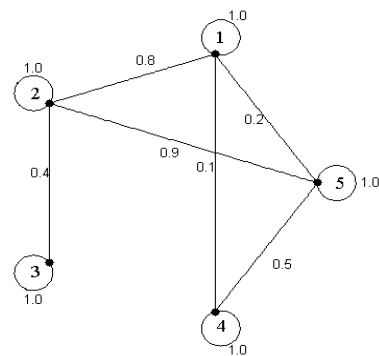

Fig 1: An initial fuzzy graph

The corresponding fuzzy tolerance relation T is presented in Table 1.

Table 1. An initial fuzzy tolerance relation

T	x_1	x_2	x_3	x_4	x_5
x_1	1.0				
x_2	0.8	1.0			

x_3	0.0	0.4	1.0	
x_4	0.1	0.0	0.0	1.0
x_5	0.2	0.9	0.0	0.5	1.0

So, an ordered sequence of values of tolerance threshold $0 < 0.1 < 0.2 < 0.4 < 0.5 < 0.8 < 0.9 < 1$ can be constructed.

The initial graph is divided into two components for the value of tolerance threshold $\alpha_3 = 0.5$. So, the α-level fuzzy graph $G_{(\alpha_\ell)}$ is shown in Fig. 2.
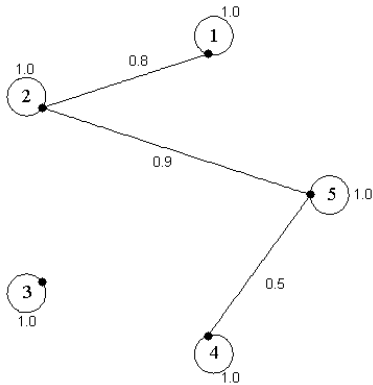

Fig 2: An α-level fuzzy graph obtained from the FG-AFC-algorithm for the value α=0.5

The allotment among two fully separated fuzzy clusters corresponds to the fuzzy graph. The fuzzy cluster $A^1_{(0.5)} = \{(x_1, 0.8), (x_2, 1.0), (x_4, 0.5), (x_5, 0.9)\}$ corresponds to the first class and the object x_2 is the unique typical point τ^1 of the fuzzy cluster. The fuzzy cluster $A^3_{(0.5)} = \{(x_3, 1.0)\}$ corresponds to the second class and the unique object which belongs to the class is its typical point, $x_3 = \tau^3$.

The corresponding allotment $R^{(\alpha_\ell)*}_c(X)$ for the value $\alpha = 0.5$ is presented in Fig. 3, where membership values of the first class are represented by ○ and membership values of the second class are represented by ■.
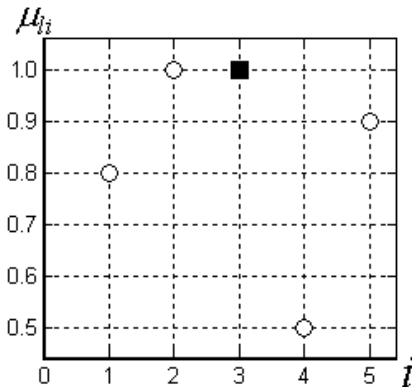
Fig 3: Membership functions of two fuzzy clusters obtained from the FG-AFC-algorithm

It should be noted, that the allotment among two fully separated fuzzy clusters obtained by the FG-AFC-algorithm is differ from the allotment among two fully separated fuzzy clusters obtained by the direct D-AFC(c)-algorithm of possibilistic clustering [4]. By executing the D-AFC(c)-algorithm for $c = 2$, the allotment was obtained for the tolerance value $\alpha = 0.1$. The corresponding allotment $R^*_c(X)$ is presented in Fig. 4, where membership values of the first class are represented by ○ and membership values of the second class are represented by ■.
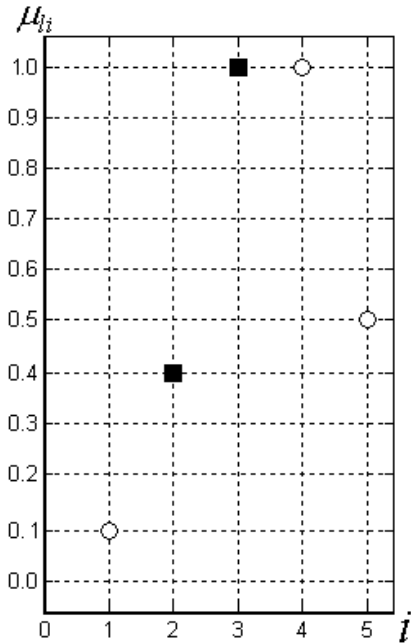

Fig 4: Membership functions of two fuzzy clusters obtained from the D-AFC(c)-algorithm

By executing the FG-AFC-algorithm, the resulting graph was obtained for the value of tolerance threshold $\alpha_4 = 0.8$. The graph is shown in Fig. 5.
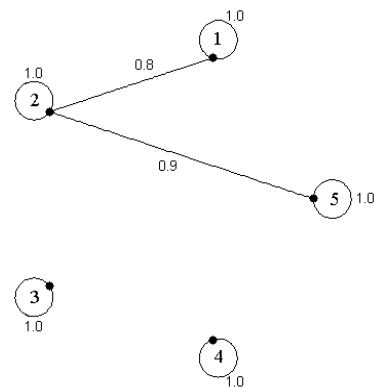

Fig 5: A resulting graph obtained from the FG-AFC-algorithm

The allotment among three fully separated fuzzy clusters corresponds to the fuzzy graph. The fuzzy cluster $A^1_{(0.5)} = \{(x_1, 0.8), (x_2, 1.0), (x_5, 0.9)\}$ corresponds to the first class and the object x_1 is the unique typical point τ^1 of the

fuzzy cluster. The fuzzy cluster $A^3_{(0.5)} = \{(x_3, 1.0)\}$ corresponds to the second class and the unique object which belongs to the class is its typical point, $x_3 = \tau^3$. The fuzzy cluster $A^4_{(0.5)} = \{(x_4, 1.0)\}$ corresponds to the third class and the unique object x_4 is the typical point τ^4 of the class. The result can be presented by the diagram of Fig. 6 where membership functions of three fuzzy clusters of the allotment $R^{(\alpha_\ell)*}_c(X)$ are shown.
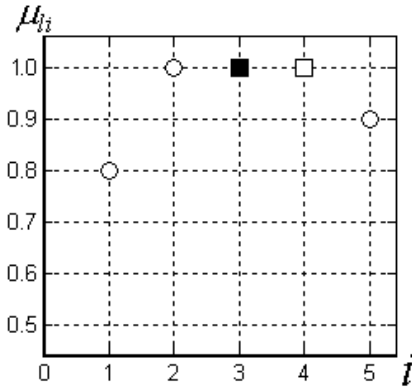

Fig 6: Membership functions of three fuzzy clustersobtained from the FG-AFC-algorithm

Membership values of the first class are represented by ○, membership values of the second class are represented by ■ and membership values of the third class are represented by □ in Fig. 6.

The allotment $R^*_c(X)$ obtained from the D-AFC(c)-algorithm for $c = 3$ is presented in Fig. 7.

The allotment was obtained for the value of tolerance threshold $\alpha = 0.5$. Membership values of the first class are represented by ○, membership values of the second class are

represented by ■ and membership values of the third class are represented by □ in Fig. 7. So, the allotment obtained from the FG-AFC-algorithm of fuzzy graph clustering is different from the allotment obtained by using the direct D-AFC(c)-algorithm.
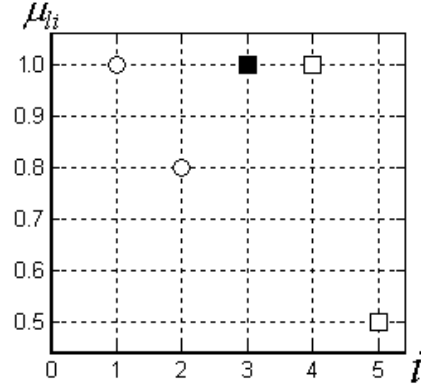

Fig 7: Membership functions of three fuzzy clusters obtained from the D-AFC(c)-algorithm

A performance of the proposed FG-AFC-algorithm can be explained by other illustrative example.

3.2 An Example of Tamura's Portrait Data Set

Let us consider an application of the proposed FG-AFC-algorithm to the classification problem for the following illustrative example. The problem of classification of family portraits coming from three families was considered by Tamura, Higuchi and Tanaka in [12]. The number of portraits was equal to 16 and the real portrait assignment among three classes is presented in Fig. 8.
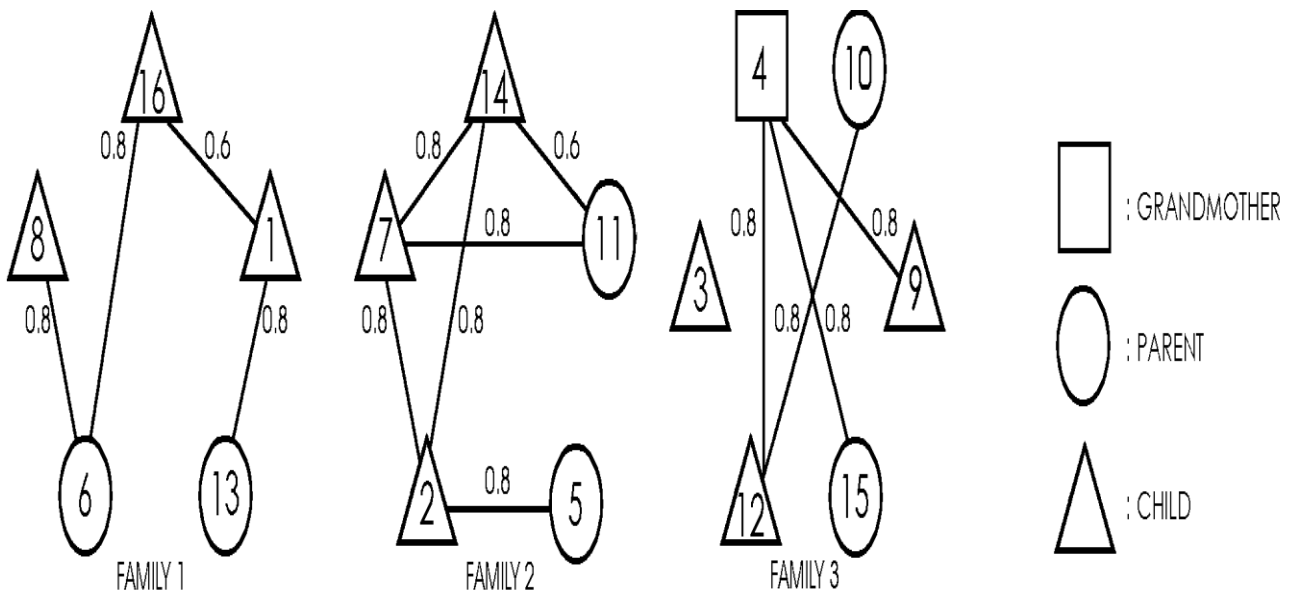

Fig 8: Real portraits classification

The data were originally analyzed in order to identify families with the technique of first transforming the matrix of a fuzzy tolerance into a matrix of a fuzzy similarity relation and then taking an appropriate α -level of the fuzzy similarity relation [12].

The partition proved to be obtained with α -level equal to 0.6. The partition identified the three families $A^1 = \{x_1, x_6, x_8, x_{13}, x_{16}\}$, $A^2 = \{x_2, x_5, x_7, x_{11}, x_{14}\}$ and $A^3 = \{x_4, x_9, x_{10}, x_{12}, x_{15}\}$.

However, person x_3 is not a member of any of the three families. The subjective similarities assigned to the individual pairs of portraits collected in the tabular format are presented in Table 2.

The presented matrix of subjective similarities is the matrix of the fuzzy tolerance relation on $X = \{x_1, \dots, x_{16}\}$. So, the initial fuzzy graph $G = (X, T)$ is shown in Fig. 9

Table 2. The matrix of subjective similarities

| T	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}
x_1	1.0															
x_2	0.0	1.0														
x_3	0.0	0.0	1.0													
x_4	0.0	0.0	0.4	1.0												
x_5	0.0	0.8	0.0	0.0	1.0											
x_6	0.5	0.0	0.2	0.2	0.0	1.0										
x_7	0.0	0.8	0.0	0.0	0.4	0.0	1.0									
x_8	0.4	0.2	0.2	0.5	0.0	0.8	0.0	1.0								
x_9	0.0	0.4	0.0	0.8	0.4	0.2	0.4	0.0	1.0							
x_{10}	0.0	0.0	0.2	0.2	0.0	0.0	0.2	0.0	0.2	1.0						
x_{11}	0.0	0.5	0.2	0.2	0.0	0.0	0.8	0.0	0.4	0.2	1.0					
x_{12}	0.0	0.0	0.2	0.8	0.0	0.0	0.0	0.0	0.4	0.8	0.0	1.0				
x_{13}	0.8	0.0	0.2	0.4	0.0	0.4	0.0	0.4	0.0	0.0	0.0	0.0	1.0			
x_{14}	0.0	0.8	0.0	0.2	0.4	0.0	0.8	0.0	0.2	0.2	0.6	0.0	0.0	1.0		
x_{15}	0.0	0.0	0.4	0.8	0.0	0.2	0.0	0.0	0.2	0.0	0.0	0.2	0.2	0.0	1.0	
x_{16}	0.6	0.0	0.0	0.2	0.2	0.8	0.0	0.4	0.0	0.0	0.0	0.0	0.4	0.2	0.0	1.0
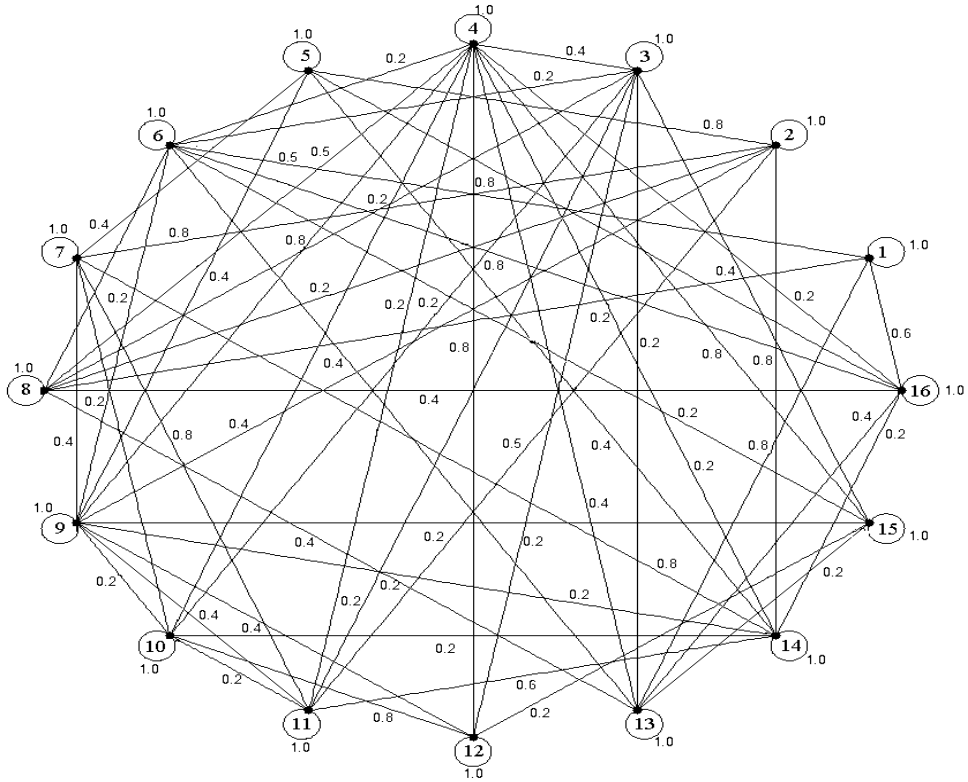
Fig 9: The initial fuzzy graph

An ordered sequence of values of tolerance threshold $0 < 0.2 < 0.4 < 0.5 < 0.6 < 0.8 < 1$ can be constructed. The initial graph cannot be divided into components for values $\alpha = 0.2$ and $\alpha = 0.4$. The α-level fuzzy graph $G_{(\alpha_\ell)}$ for the value $\alpha = 0.5$ is shown in Fig. 10. Membership functions of corresponding fuzzy clusters of the obtained allotment are presented in Fig. 11.
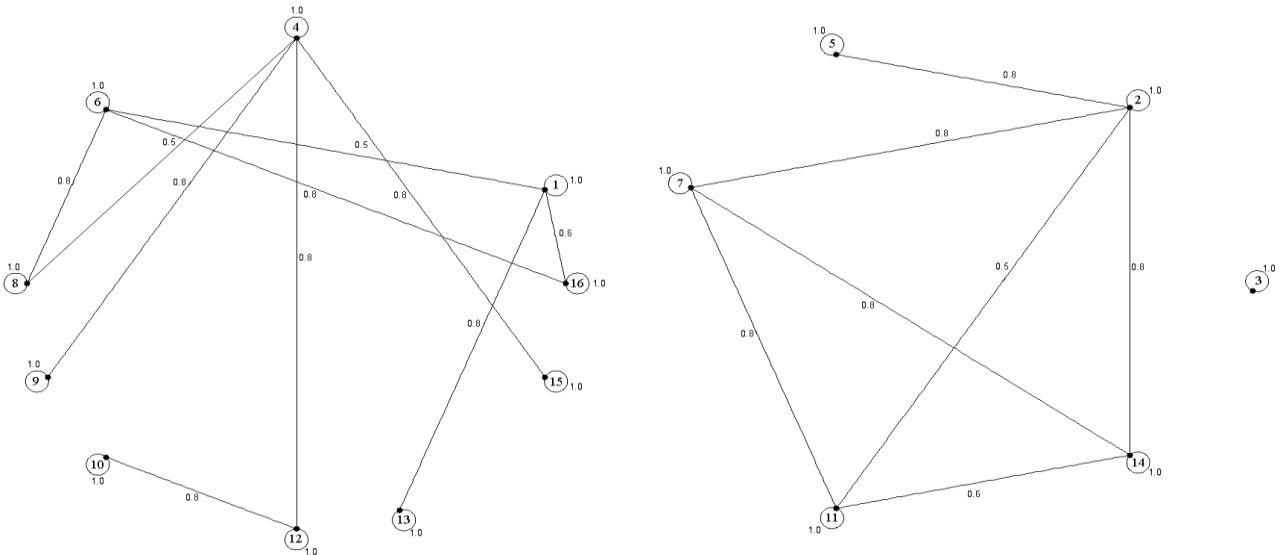

Fig 10: An α-level fuzzy graph obtained from the FG-AFC-algorithm for the value α=0.5
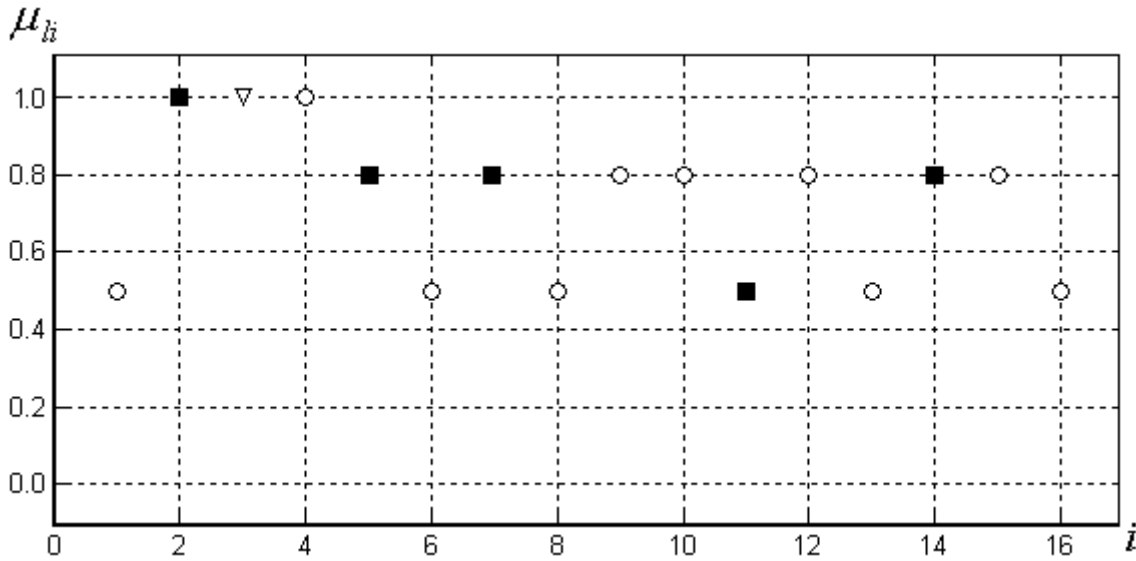
Fig 11: Membership functions of three fuzzy clusters obtained from the FG-AFC-algorithm, which correspond to the obtained allotment

Membership values of the first class are represented by $\circ$, membership values of the second class are represented by $\blacksquare$ and membership values of the third class are represented by ∇ in Fig. 11. So, the fourth object is the typical point of the first class, the second object is the typical point of the second class, and the third object is the typical point of the third class.

The second class is corresponds to Family 2 and all members of Family 1 is included to the first class. So, the first class is the union of the Family 1 and Family 3 and the third object is a unique element of the third class. So, the results obtained by the proposed FG-AFC-algorithm seem to be appropriate.

It should be noted, that the allotment among particularly separated fuzzy clusters was obtained by using the D-AFC(c)-algorithm for $c = 3$ [4]. The corresponding membership functions are presented in Fig. 12.
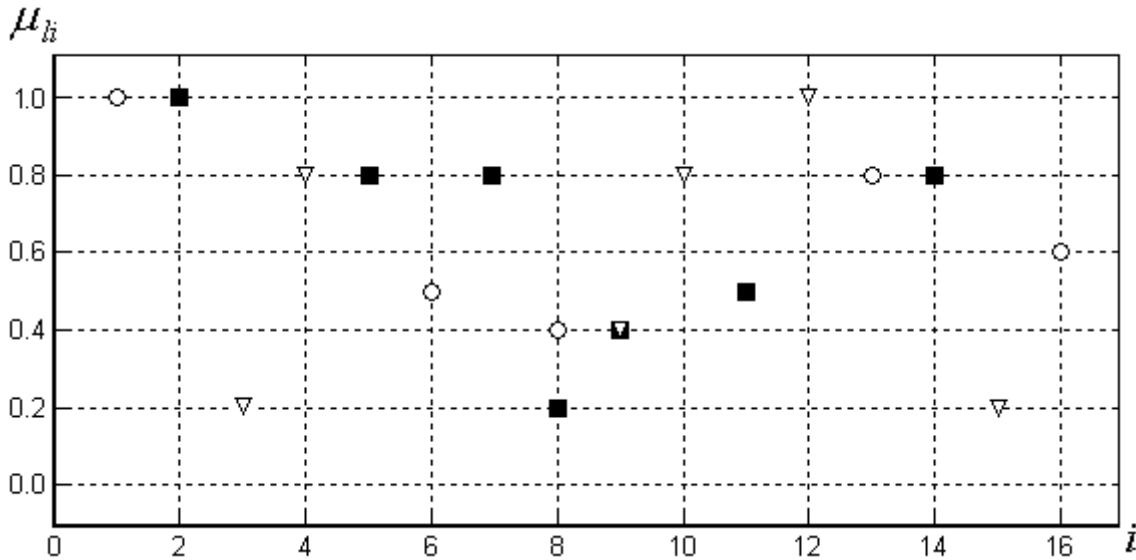

Fig 12: Membership functions of three partially separated fuzzy clusters obtained from the D-AFC-algorithm

So, by executing the D-AFC(c)-algorithm for three classes, we obtain the allotment $R_c^*(X)$ among particularly separated fuzzy clusters, which corresponds to the result, is received for the tolerance threshold $\alpha = 0.2$. The ninth element of the set of objects is belonging to the second class and to the third class and membership values are equal, $\mu_{2\,9} = \mu_{3\,9} = 0.4$.

4. FINAL REMARKS

Results of experiments are summarized and discussed in the first subsection of the section. The second subsection deals with the perspectives on future investigations.

4.1 Discussions

The novel FG-AFC-algorithm of possibilistic clustering is proposed in the paper. The algorithm is based on the idea of detecting of components of the initial fuzzy graph and constructing fuzzy clusters of the sought allotments from the components directly. So, the criterion of quality of the sought allotment is not used in the FG-AFC-algorithm. The results of application of the proposed FG-AFC-algorithm to the data sets show that the algorithm is an effective tool for solving the classification problem.

4.2 Perspectives

A transitive approximation of the initial fuzzy tolerance relation can be obtained by the TAGA-algorithm [13]. So, the TAGA-algorithm can be implemented into the FG-AFC-algorithm and a new fuzzy graph-based heuristic algorithm of possibilistic clustering can be elaborated.

5. REFERENCES

[1] Höppner, F., Klawonn, F., Kruse, R. and Runkler, T. 1999. Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition. Chichester: John Wiley & Sons.

[2] Krishnapuram, R. and Keller, J.M. 1993. A Possibilistic Approach to Clustering. IEEE Transactions on Fuzzy Systems. 1, 98-110.

[3] Mandel, I.D. 1988. Clustering Analysis. Moscow: Finansy i Statistica. (in Russian)

[4] Viattchenin, D.A. 2013. A Heuristic Approach to Possibilistic Clustering: Algorithms and Applications. Heidelberg: Springer.

[5] Schaeffer, S. 2007. Graph Clustering. Computer Science Review. 1, 27-64.

[6] Fortunato, S. 2010. Community Detection in Graphs. Physics Reports. 486, 75-174.

[7] Zadeh, L.A. 1965. Fuzzy Sets. Information and Control. 8, 338-353.

[8] Kaufmann, A. 1975. Introduction to the Theory of Fuzzy Subsets. New York: Academic Press.

[9] Rosenfeld, A. 1975. Fuzzy Graphs. In Fuzzy Sets and Their Applications to Cognitive and Decision Processes. New York: Academic Press, 77-95.

[10] Devillez, A., Billaudel, P. and Villermain Lecolier, G. 2002. A Fuzzy Hybrid Hierarchical Clustering Method with a New Criterion Able to Find the Optimal Partition. Fuzzy Sets and Systems. 128, 77-95.

[11] Dong, Y., Zhuang, Y., Chen, K. and Tai, X. 2006. A Hierarchical Clustering Algorithm Based on Graph Connectedness. Fuzzy Sets and Systems. 157, 1760-1774.

[12] Tamura, S., Higuchi, S. and Tanaka, K. 1971. Pattern Classification Based on Fuzzy Relations. IEEE Transactions on Systems, Man, and Cybernetics. 1, 61-66.

[13] Dawyndt, P., De Meyer, H. and De Baets, B. 2006. UPGMA Clustering Revisited: A Weight-Driven Approach to Transitive Approximation. International Journal of Approximate Reasoning. 42, 174-191.