

# A Homogeneous Set-Theoretical Frame For Clustering Fuzzy Relational Data

Narcís Clara

Departament d'Informàtica i Matemàtica Aplicada

Universitat de Girona

Campus de Montilivi, 17071 Girona, Catalonia (Spain)

narcis.clara@udg.edu

## Abstract

*Our purpose is to provide a set-theoretical frame to clustering fuzzy relational data basically based on cardinality of the fuzzy subsets that represent objects and their complementaries, without applying any crisp property. From this perspective we define a family of fuzzy similarity indexes which includes a set of fuzzy indexes introduced by Tolia et al, and we analyze under which conditions it is defined a fuzzy proximity relation. Following an original idea due to S. Miyamoto we evaluate the similarity between objects and features by means the same mathematical procedure. Joining these concepts and methods we establish an algorithm to clustering fuzzy relational data. Finally, we present an example to make clear all the process.*

## 1. Introduction

Clustering analysis (CA) is one of the most important applied techniques for pattern recognition. The basic idea of CA is to group together objects closely related. Methods with objective function models using object data have been very developed ([1, 2, 5, 9, 11]) but we will focus our attention in methods with relational clustering. We can distinguish two main types: methods that rely on optimization of an objective function of the relational data and methods that use decompositions of relation matrices, transitive closures or mathematical algorithms. In all approaches is essential to recognize how similar are objects. We intend to put together known results of relational fuzzy clustering with a new family of similarity indexes.

Using matrices of relational data, a fuzzy clustering process consists in three steps. The first one is to define a fuzzy relation between objects and features, after that we have to evaluate the similarity between objects, which can be thought as a fuzzy proximity relation, and, finally, we determine the partitions by its transitive closure for some  $t$ -norm or by means an algorithm ([8]).

If we prefer to apply any hierarchical method based on an algorithm then two clusters are usually grouped to constitute a new cluster when their similarity attains the maximum value. To continue the process we need to define the similarity between this new element and the others. There are many methods to deal with this objective, the most usual is the single linkage, complemented by the complete linkage and the average linkage.

Features Contrast Model (FCM) asserts that similarity should be described as a comparison of features that describe the objects under consideration, and expresses similarity between objects as a function of their common and distinctive features ([10]).

Let  $X = \{A_1, A_2, \dots, A_n\}$  be a set of objects and  $Y = \{P_1, P_2, \dots, P_m\}$  a set of features. FCM assigns to each element  $A_i$  a value  $\mu_{A_i}(P_k) \in \{0, 1\}$  depending on if  $A_i$  verifies or not feature  $P_k$ . Function  $v : X \rightarrow \wp(Y)$  defined as  $P_k \in v(A_i)$  if and only if  $A_i$  verifies  $P_k$  serves to relate objects with features. Fuzzy Features Contrast Model (FFCM) is an extension of FCM when the values belong to the interval  $[0, 1]$  so they are fuzzy. We represent each object  $A_i$  as a fuzzy subset  $\tilde{A}_i$  of the set of features  $Y$  namely

$$A_i \longrightarrow \tilde{A}_i = (\mu_{\tilde{A}_i}(P_1), \mu_{\tilde{A}_i}(P_2), \dots, \mu_{\tilde{A}_i}(P_m)) \quad (1)$$

In these conditions  $card(\tilde{A}_i) = |\tilde{A}_i| = \sum_{j=1}^n \mu_{\tilde{A}_i}(P_j)$  ([4]).

All information determines a matrix  $S \in M_{m \times n}[0, 1]$  representing a fuzzy relation  $\tilde{S}$  between objects and features. Therefore two objects  $A$  and  $B$  are represented by two fuzzy subsets  $\tilde{A}$  and  $\tilde{B}$ . These values are obtained from the inner structure of the data or by experts. An additional hypothesis for FFCM is that similarity between  $A$  and  $B$  has to be a function of  $\tilde{A} \cap \tilde{B}$ ,  $\tilde{A} - \tilde{B}$  and  $\tilde{B} - \tilde{A}$ .

### 1.1. Fuzzy similarity measures

A very common approach in engineering and applied sciences is to measure the similarity between objects associ-

ating them to a  $n$ -dimensional vector and calculating their similarity by means a decreasing function  $s = f(d)$  (usually  $s = 1 - d$ ) of their normalized distance  $d$ , so following the first assumption of FFCM but the second one because an object is only thought as a point in a metric space. Very strong properties are fulfilled but when the human judgement has a principal role in the description of the objects can lead to undesirable properties because do not take in account the deep fuzzy structure of the data.

To deal with the fuzzy character of the data sometimes a softer definition is usually taken in account. A function  $s : \tilde{P}(E) \times \tilde{P}(E) \rightarrow [0, 1]$  is a *fuzzy similarity measure* if and only if verifies that for any pair of fuzzy subsets  $\tilde{A}, \tilde{B}$   $0 \leq s(\tilde{A}, \tilde{B}) \leq 1$  and  $s(\tilde{A}, \tilde{B}) = s(\tilde{B}, \tilde{A})$  (symmetry). If moreover  $\forall \tilde{A} \quad s(\tilde{A}, \tilde{A}) = 1$  (reflexivity), it can be interpreted as a *proximity fuzzy relation* in  $X$ . These kind of fuzzy relations are fundamental to carry out a fuzzy clustering process ([12]).

For binary relational data and in order to analyze how similar are two objects we define similarity and dissimilarity crisp parameters ([8]):

$$\begin{aligned} a &= |v(A_i) \cap v(A_j)| & b &= |v(A_i) \cap v(A_j)^c| \\ c &= |v(A_i)^c \cap v(A_j)| & d &= |v(A_i)^c \cap v(A_j)^c| \end{aligned} \quad (2)$$

Parameters  $a$  and  $d$  evaluate common features and  $b$  and  $c$  distinctive features. From similarity and dissimilarity parameters are defined many similarity indexes for crisp sets ([8]).

## 1.2. Transitive closures

Let  $\tilde{R}$  be a fuzzy relation in an universe of objects  $X$ , then  $\tilde{R}$  is  $t$ -transitive if and only if

$$\forall a, b \in A \quad \mu_{\tilde{R}}(a, b) \geq \max_{c \in A} t(\mu_{\tilde{R}}(a, c), \mu_{\tilde{R}}(c, b))$$

A proximity fuzzy relation  $\tilde{R}$  is a  $t$ -equivalence fuzzy relation if verifies the  $t$ -transitive condition. The concept of closure  $t$ -transitive  $\tilde{R}^*$  of a proximity relation  $\tilde{R}$  is defined as the *smaller*  $t$ -equivalence relation that includes  $\tilde{R}$  namely

$$\tilde{R}^* = \min\{\tilde{S} : \tilde{R} \subset \tilde{S} \text{ and } \tilde{S} \text{ is } t\text{-equivalence}\} \quad (3)$$

When the  $t$ -norm is not explicitly mentioned means that we use the  $t$ -norm of the minimum. Designating by  $R$  the matrix representation of  $\tilde{R}$ , we calculate its  $t$ -transitive closure by means  $R^* = R^n$  ([2]). Other methods most performing to calculate the transitive closure have been found, which reduce its computing time from order  $O(n^5)$  to order  $O(n^2)$  ([6]).

## 1.3. Hierarchical methods for fuzzy clustering relational data

Hierarchical methods consist in not considering fix beforehand how may clusters there are and are divided in ag-

glomerative and divisive methods. In this paper we only refer to agglomerative methods in which we start from the most finer partition composed by all the singletons until we merge all the objects in an unique cluster. It seems very logical that at each step we group elements with the maximum similarity (minimum distance). When two objects have been grouped our set of reference changes and its cardinal diminishes in one unit. In some manner we have to define the similarity between this new element and the others. The form in which is carried out this process determines all the clustering process because different definitions for similarity between clusters induce different algorithms. The most known is given by the single linkage: let  $\Delta$  and  $\Gamma$  be two clusters, if we group  $\Delta$  and  $\Gamma$  so  $\Sigma = \Delta \cup \Gamma$  therefore

$$\forall \Phi, \Phi \neq \Delta, \Gamma \quad s(\Sigma, \Phi) = \max(s(\Delta, \Phi), s(\Gamma, \Phi)) \quad (4)$$

Others, as complete linkage or average linkage simply change the maximum for the minimum or an average respectively.

Another approach is based in the theory of fuzzy relations. We calculate the transitive closure with the  $t$ -norm of the minimum, which is the greatest  $t$ -norm, and for each  $\alpha$ -level we obtain a partition. A very important theorem proves that this partition coincides with the results of the simple linkage clustering process and the connected components of a fuzzy graph ([8]). All that makes advisable to use this method for a great set of applications. In fact what happens is that the condition of transitivity for the function  $s$  is equivalent to the ultrametric property. Unfortunately, we obtain undesirable results in some applications what means that for all  $\alpha$ -level the partition is formed only by the whole universe or by  $n$  clusters all them constituted by only one element.

A more general procedure is calculating the closure  $t$ -transitive for some  $t$ -norm greater than the bounded product ( $t(x, y) = \max(0, x + y - 1)$ ) because for this  $t$ -norm the  $t$ -transitivity condition is equivalent to the triangular inequality, and by means an iterative process obtaining a partition. Following this strategy Miin-Shen Yang et al have implemented an algorithm that in certain domains improve the single linkage method but loosing uniqueness what is a great inconvenient ([12]). The algorithm proposed at subsection 2.2. wants to be another alternative to these kind of procedures.

## 2. A homogenous fuzzy set-theoretical frame

The aim of this section is to provide a homogeneous structure for clustering depending basically on the cardinality of the fuzzy subsets. That is achieved using a set of similarity indexes which are a generalized set of known fuzzy indexes ([10]) and defining the fuzzy structure of clusters

what allow us to determine similarities between objects and between clusters with the same mathematical method.

## 2.1. Generalized fuzzy indexes

We wish to generalize crisp indexes making an extension of (2), without applying any crisp property and proving that reliable properties are fulfilled. Following definitions and properties are a generalization of some concepts due to Toliás et al which define a set of fuzzy similarity indexes -Generalized Tversky index ([10])- that represents a fuzzified restraint set of crisp indexes. This set contains as a particular case Jackard's coefficient but simple matching coefficient or Rao's coefficient. What we propose has a similar structure but including all the indexes.

FFCM assumes that similarity between two objects depends on their common and distinctive features. We think that this affirmation can be interpreted in a general way in the sense that it is also relevant in which degree two objects share the negation of a specific feature. Assuming this point of view, similarity between  $A$  and  $B$  is a function of  $\tilde{A} \cap \tilde{B}$ ,  $\tilde{A} - \tilde{B} = \tilde{A} \cap \tilde{B}^c$ ,  $\tilde{B} - \tilde{A} = \tilde{B} \cap \tilde{A}^c$  and  $\tilde{A}^c \cap \tilde{B}^c$ . From now on  $t$  and  $n$  mean a  $t$ -norm and a negation respectively ([7]). Making a fuzzy generalization of (2) we introduce the fuzzy similarity and dissimilarity parameters in the following form:

$$\begin{aligned} a &= |\tilde{A}_i \cap \tilde{A}_j| = \sum_{k=1}^m t(\mu_{\tilde{A}_i}(P_k), \mu_{\tilde{A}_j}(P_k)) \\ b &= |\tilde{A}_i \cap \tilde{A}_j^c| = \sum_{k=1}^m t(\mu_{\tilde{A}_i}(P_k), n(\mu_{\tilde{A}_j}(P_k))) \\ c &= |\tilde{A}_i^c \cap \tilde{A}_j| = \sum_{k=1}^m t(n(\mu_{\tilde{A}_i}(P_k)), \mu_{\tilde{A}_j}(P_k)) \\ d &= |\tilde{A}_i^c \cap \tilde{A}_j^c| = \sum_{k=1}^m t(n(\mu_{\tilde{A}_i}(P_k)), n(\mu_{\tilde{A}_j}(P_k))) \end{aligned} \quad (5)$$

We determine a family of generalized fuzzy similarity indexes, which includes a generalization of the simple matching coefficient for  $\lambda = \mu = 1$ , defined by

$$s_g(\tilde{A}, \tilde{B}) = \frac{|\tilde{A} \cap \tilde{B}| + |\tilde{A}^c \cap \tilde{B}^c|}{|\tilde{A} \cap \tilde{B}| + |\tilde{A}^c \cap \tilde{B}^c| + \lambda|\tilde{A} - \tilde{B}| + \mu|\tilde{B} - \tilde{A}|} \quad (6)$$

This family verifies two expected conditions. Exists  $f : \tilde{P}(Y) \times \tilde{P}(Y) \times \tilde{P}(Y) \times \tilde{P}(Y) \rightarrow \mathbf{R}$  that for any elements  $\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}$  belonging to  $\tilde{P}(E)$  therefore  $s_g(\tilde{A}, \tilde{B}) = f(\tilde{A} \cap \tilde{B}, \tilde{A}^c \cap \tilde{B}^c, \tilde{A} - \tilde{B}, \tilde{B} - \tilde{A})$ . On the other hand, monotony is also fulfilled in the sense that if  $\tilde{A} \cap \tilde{C} \subset \tilde{A} \cap \tilde{B}$ ,  $\tilde{A}^c \cap \tilde{C}^c \subset \tilde{A}^c \cap \tilde{B}^c$ ,  $\tilde{A} - \tilde{C} \subset \tilde{A} - \tilde{B}$  and  $\tilde{C} - \tilde{A} \subset \tilde{B} - \tilde{A}$  therefore  $s_g(\tilde{A}, \tilde{B}) \geq s_g(\tilde{A}, \tilde{C})$  [3].

It is clear that  $s_g$  is a fuzzy similarity measure. A relevant property issues from the fact that  $s_g$  is a proximity relation if the associated  $t$ -norm verifies the noncontradiction principle ([7]):  $s_g(\tilde{A}, \tilde{A}) = 1$  if and only if  $|\tilde{A} - \tilde{A}| = |\tilde{A} \cap \tilde{A}^c| = 0$  what means that for all

$x \in [0, 1]$   $t(x, 1 - x) = 0$ , thus the  $t$ -norm have to verify the noncontradiction principle. Moreover, under prototypical conditions  $s_g((1, 1, \dots, 1), \tilde{B})$  does not depend on the  $t$ -norm, and  $s_g((1, 1, \dots, 1), (1, 1, \dots, 1)) = 1$ , what is very logical. Under maximum uncertainty conditions  $s_g(0.5, 0.5, \dots, 0.5), \tilde{B}) = 0.5$ . This last property solves the question about the inconsistency in considering absolutely similar two objects with the maximum uncertainty conditions.

## 2.2. Clustering method defining the fuzzy structure of clusters

This method has been conceived from some observations due to S. Miyamoto. It seems logical to use similars methodologies in the diferents steps of the process seeking a most homogenous theoretical frame.

Let  $\tilde{\wp}(Y)$  be the power set of fuzzy subsets of  $Y$  therefore there is a map between objects and fuzzy subsets of features

$$\begin{aligned} X &\longrightarrow \tilde{\wp}(Y) \\ A_i &\longrightarrow \tilde{A}_i \quad \text{so that} \quad Y \xrightarrow{\mu_{\tilde{A}_i}} [0, 1] \\ & \quad \quad \quad \quad \quad \quad \quad \quad P_j \longrightarrow \mu_{\tilde{A}_i}(P_j) \end{aligned}$$

When we form a new cluster we make the union of two precedent clusters. It is necessary to define the membership function of a cluster without confusing it with the union of two fuzzy subsets. The difference consists in the fact that if  $\Delta = \{A, B\}$  then  $\tilde{A} \cup \tilde{B} \neq \tilde{\Delta} = \{\tilde{A}, \tilde{B}\}$ . We need to define in some manner the membership function of  $\tilde{\Delta} \in \tilde{P}(Y)$ .

Let  $\Delta = \{A_i\}_{i \in I^*} \in \wp(X)$  be a cluster and  $I^* \subset I$  with  $I = \{1, 2, \dots, n\}$ . We define a correspondence between clusters and fuzzy subsets of features by means the following map

$$\begin{aligned} \wp(X) &\longrightarrow \tilde{\wp}(Y) \\ \Delta &\longrightarrow \tilde{\Delta} \quad \text{so that} \\ Y &\xrightarrow{\mu_{\tilde{\Delta}}} [0, 1] \\ P_j &\longrightarrow \mu_{\tilde{\Delta}}(P_j) = \frac{\sum_{i \in I^*} \mu_{\tilde{A}_i}(P_j)}{|\Delta|} \end{aligned} \quad (7)$$

If  $\Gamma$  and  $\Lambda$  are two clusters represented by fuzzy subsets  $\tilde{\Gamma} = \{g, g \in \Gamma\}$  and  $\tilde{\Lambda} = \{h, h \in \Lambda\}$ , we define

$$s(\Gamma, \Lambda) = \frac{|\Gamma \cap \Lambda| + |\Gamma^c \cap \Lambda^c|}{|\Gamma \cap \Lambda| + |\Gamma^c \cap \Lambda^c| + \lambda|\Gamma \cap \Lambda^c| + \mu|\Gamma^c \cap \Lambda|} \quad (8)$$

Unfortunately, do not exist formula to update the similarity measures from one step to another. It is necessary to calculate the membership function of the new cluster determining the similarity measure between this new element

and the others. After that we need an algorithm to establish the partition. We present a procedure that has a similar structure to the classical algorithms of hierarchical clustering; obviously, other algorithms can be introduced but all the ideas lead us to an algorithm similar to this one. We do not use a programming language, is only a description of the necessary steps to achieve the hierarchical clustering. The iterative process of the classical methods is modified by the following form:

**Input:** A set of objects  $X$  with  $|X| = n$ , a set of features  $Y$  with  $|Y| = m$  and a fuzzy relation  $\tilde{S}$  between  $X$  and  $Y$  represented by a matrix  $S \in M_{m \times n}([0, 1])$ .

**Output:** A directed tree that describes the process of generating clusters called dendrogram.

*Algorithm*

- For  $i = 1, \dots, n$  do  $\Delta_i = \{A_i\}$ ,  
 $Z^{(0)} = \{\Delta_1, \dots, \Delta_n\}$  and  $I = \{1, \dots, n\}$ .
- For  $i, j \in I, i \neq j$  do  $s(\Delta_i, \Delta_j) = s_g(A_i, A_j)$
- $l = n, k = 0$  ( $l$  enumerate clusters and  $k$  partitions).
- $l = l + 1, k = k + 1$
- Find  $\alpha = \max_{G, H \in Z^{(k)}; G \neq H} s(G, H) = s(\Delta_i, \Delta_j)$  (If there is more than a pair of elements that verify this condition we select between them randomly).
- $\Delta_l = \Delta_i \cup \Delta_j$
- $\mu_{\tilde{\Delta}_l}(P_j) = \frac{\sum_{A_i \in \Delta_l} \mu_{\tilde{A}_i}(P_j)}{|\Delta_l|}$
- $I = I - \{i, j\} \cup \{l\}, Z^{(k)} = \bigcup_{i \in I} \{\Delta_i\}$
- For  $i \in I, i \neq l$  do  $s(\Delta_i, \Delta_l) = s_g(\Delta_i, \Delta_l)$  (Updating similarities).
- Repeat the process until  $|I| = 1$  ( $l = 2n - 1, k = n - 1$  or  $|Z^{(k)} = 1$ ).

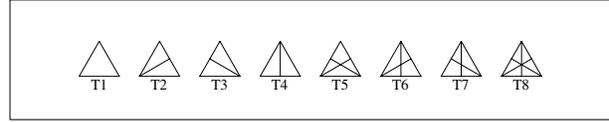
*end of the algorithm*

As the fuzzy relation is changed from a step to another step the maximum values of the similarity (or minimums of the distance) refer to different matrices so we do not have an order of the similarity levels as we have for instance in the simple linkage method. We can order only the steps.

### 3. Example

With the aim to clarify all the process by means an application we have chosen a complete geometrical example. Like this we have the great advantage of checking the results in a unmistakable way. On the other hand, the fuzzy

relation between objects and features becomes crisp. Calculations have been reduced to the minimum expression because they can be easily checked. Our objects are equilateral triangles with all the possible combinations of bisectors therefore  $X = \{T1, T2, T3, T4, T5, T6, T7, T8\}$ . One triangle has no bisectors, three have one bisector, three have two bisectors and one has three bisectors.



**Figure 1. Equilateral triangles.**

Our objective is to group objects more similar in the sense that they have the same bisectors (not the same number of bisectors). Let  $b_i$  be the bisectors with positive, negative and vertical slope respectively. Let  $Y = \{P_1, P_2, P_3\}$  be the set of features defined by  $P_i =$  to have the bisector  $b_i$ , and  $T_i$  the equilateral triangles represented at Figure 1. The fuzzy relation  $\tilde{S}$  between  $X$  and  $Y$  is determined by the matrix

$$S = \begin{matrix} & T1 & T2 & T3 & T4 & T5 & T6 & T7 & T8 \\ \begin{matrix} P_1 \\ P_2 \\ P_3 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix} \end{matrix}$$

As similarity index we have chosen a member of the family of generalized fuzzy indexes defined by  $\lambda = \mu = 1$ . That index is a fuzzy generalization of the simple matching coefficient and we call it  $s_{gsm}$ . Therefore

$$s_{gsm}(\tilde{A}, \tilde{B}) = \frac{|\tilde{A} \cap \tilde{B}| + |\tilde{A}^c \cap \tilde{B}^c|}{|\tilde{A} \cap \tilde{B}| + |\tilde{A}^c \cap \tilde{B}^c| + |\tilde{A} - \tilde{B}| + |\tilde{B} - \tilde{A}|} \quad (9)$$

In order to ensure the reflexive hypothesis, cardinals have been calculated by the  $t$ -norm of the bounded product which verifies the non-contradiction principle. Let  $R = (r_{ij})$  be the matrix of similarities then

$$R = \begin{matrix} & T1 & T2 & T3 & T4 & T5 & T6 & T7 & T8 \\ \begin{matrix} T1 \\ T2 \\ T3 \\ T4 \\ T5 \\ T6 \\ T7 \\ T8 \end{matrix} & \begin{pmatrix} 1 & 0.67 & 0.67 & 0.67 & 0.33 & 0.33 & 0.33 & 0 \\ 0.67 & 1 & 0.33 & 0.33 & 0.67 & 0.67 & 0 & 0.33 \\ 0.67 & 0.33 & 1 & 0.33 & 0.67 & 0 & 0.67 & 0.33 \\ 0.67 & 0.33 & 0.33 & 1 & 0 & 0.67 & 0.67 & 0.33 \\ 0.33 & 0.67 & 0.67 & 0 & 1 & 0.33 & 0.33 & 0.67 \\ 0.33 & 0.67 & 0 & 0.67 & 0.33 & 1 & 0.33 & 0.67 \\ 0.33 & 0 & 0.67 & 0.67 & 0.33 & 0.33 & 1 & 0.67 \\ 0 & 0.33 & 0.33 & 0.33 & 0.67 & 0.67 & 0.67 & 1 \end{pmatrix} \end{matrix}$$

We will apply three methods in order to notice the differences.

#### 3.1. Transitive closures

As  $0.33 = r_{51} < 0.67 = \max_{j=1 \dots 5} \min\{r_{5j}, r_{j1}\}$ , results that  $\tilde{R}$  is not min-transitive. Since  $R \neq R^2, R^2 \neq R^3$

but  $R^3 = R^4$  we deduce  $R^* = R^4$ . Making the necessary calculations we find  $R_{ii}^* = 1$  and  $R_{ij}^* = 0.67$  if  $i \neq j$ .

We obtain the different partitions in function of the  $\alpha$ -levels. If  $0 < \alpha \leq 0.67$ , only one cluster is formed: the whole set  $X$ . If  $0.67 < \alpha \leq 1$ , eight clusters are formed:  $\{T1\}$ ,  $\{T2\}$ ,  $\{T3\}$ ,  $\{T4\}$ ,  $\{T5\}$ ,  $\{T6\}$ ,  $\{T7\}$  and  $\{T8\}$ .

Now we check another possibility. We will apply a method proposed by M. S. Yang and H. M. Shih ([12]) using the  $t$ -transitive closure with the  $t$ -norm of the bounded product which is the smallest that verify the triangular inequality. We obtain  $R = R^2$  so  $\tilde{R}$  is  $t_{bp}$ -transitive. Choosing  $\alpha = 0.55$  and applying the algorithm following its nomenclature, which do not verify uniqueness, we find the maximum value at 0.67. Randomly we select  $r_{21}^{(2)} = 0.67 \neq 0$ ; so  $C = \{1, 2\}$  is our first set of indexes, and  $I - C = \{3, 4, 5, 6, 7, 8\}$ . As  $r_{31}^{(2)} + r_{32}^{(2)}$ ,  $r_{41}^{(2)} + r_{42}^{(2)}$ ,  $r_{51}^{(2)} + r_{52}^{(2)}$ ,  $r_{61}^{(2)} + s_{62}^{(2)}$ ,  $r_{71}^{(2)} + r_{72}^{(2)}$  and  $r_{81}^{(2)} + r_{82}^{(2)}$  have some element equal to 0 therefore our first cluster is  $\Delta_1 = \{T1, T2\}$ . Following the same procedure  $\Delta_2 = \{T3, T5\}$ ,  $\Delta_3 = \{T4, T6\}$  and  $\Delta_4 = \{T7, T8\}$ . We achieve a partition in which the elements with more differences have not been grouped. These results are better than those of the transitive closure.

### 3.2. Method exposed at subsection 2.2.

At the first step  $\Delta_i = \{Ti\}$  for  $i = 1, \dots, 8$  so  $Z^{(0)} = X$ . As some objects attain the maximum value (0.67) we choose between them randomly, for instance  $s_{78} = 0.67$  and we put them together so  $\Delta_9 = \{T7, T8\}$  and  $Z^{(1)} = \{\{T1\}, \{T2\}, \{T3\}, \{T4\}, \{T5\}, \{T6\}, \{T7, T8\}\}$ . At this moment we have to calculate the similarity between this new clusters and the others by means (9). From (7) the membership function of  $\tilde{\Delta}_1$  is defined by the values  $\mu_{\tilde{\Delta}_1}(P_1) = 0.5$ ,  $\mu_{\tilde{\Delta}_1}(P_2) = 1$  and  $\mu_{\tilde{\Delta}_1}(P_3) = 1$ . Following the same methodology we obtain the new matrix of similarities

$$S_1 = \begin{matrix} & \begin{matrix} T1 & T2 & T3 & T4 & T5 & T6 & \Delta_1 \end{matrix} \\ \begin{matrix} T1 \\ T2 \\ T3 \\ T4 \\ T5 \\ T6 \\ \Delta_1 \end{matrix} & \begin{pmatrix} 1 & 0.67 & 0.67 & 0.67 & 0.33 & 0.33 & 0.17 \\ 0.67 & 1 & 0.33 & 0.33 & 0.67 & 0.67 & 0.17 \\ 0.67 & 0.33 & 1 & 0.33 & 0.67 & 0 & 0.5 \\ 0.67 & 0.33 & 0.33 & 1 & 0 & 0.67 & 0.5 \\ 0.33 & 0.67 & 0.67 & 0 & 1 & 0.33 & 0.5 \\ 0.33 & 0.67 & 0 & 0.67 & 0.33 & 1 & 0.5 \\ 0.17 & 0.17 & 0.5 & 0.5 & 0.5 & 0.5 & 1 \end{pmatrix} \end{matrix}$$

The maximum value is 0.67. We build  $\Delta_{10} = \{T4, T6\}$  so  $Z^{(2)} = \{\{T1\}, \{T2\}, \{T3\}, \{T5\}, \{T4, T6\}, \{T7, T8\}\}$ . For the next steps we find that  $\Delta_{11} = \{T3, T5\}$ ,  $\Delta_{12} = \{T1, T2\}$ ,  $\Delta_{13} = \Delta_5 \cup \Delta_2 = \{T4, T6, T7, T8\}$ ,  $\Delta_{14} = \Delta_3 \cup \Delta_4 = \{T1, T2, T3, T5\}$ , and, finally,  $\Delta_{15} = X$ . We represent all these results in a dendrogram in Figure 2.

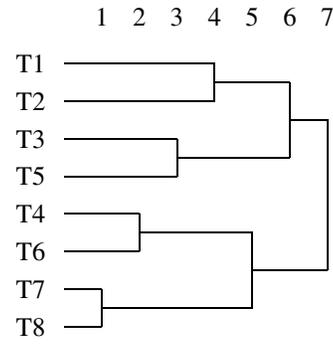


Figure 2. Dendrogram of the clusters.

## 4. Conclusions

Combining techniques on fuzzy similarities, without considering any a priori crisp simplification, and fuzzy clustering basically based on the fuzzy structure of the data given by membership functions, we have proposed an algorithm that shows reliable results when other classical methods do not achieve a good partition of the universe.

## References

- [1] J. C. Bezdek, J. Keller, R. Krishnapuram, and N. R. Pal. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*. Kluwer Academic Publishers, Boston, 1999.
- [2] D. Boixader, J. Jacas, and J. Recasens. Transitive closure and betweenness relations. *Fuzzy Sets and Systems*, 120(7):415–422, 2001.
- [3] N. Clara. Generalized fuzzy similarity indexes. *Neural Nets. Lectures Notes in Computer Science*, 3931:163–170, Springer 2006.
- [4] D. Dubois and H. Prade. *Fuzzy Sets and Systems, Theory and Applications*. Academic Press, San Diego, 1980.
- [5] R. Krishnapuram and J. M. Keller. A possibilistic approach to clustering. *IEEE Trans, Fuzzy Systems*, 2(1):98–110, 1993.
- [6] H. S. Lee. An optimal algorithm for computing the max-min transitive closure of a fuzzy similarity matrix. *Fuzzy Sets and Systems*, 1(123):129–136, 2001.
- [7] R. Lowen. *Fuzzy Set Theory*. Kluwer Academic Publishers, Dordrecht, 1996.
- [8] S. Miyamoto. *Fuzzy Sets on Information Retrieval and Cluster Analysis*. Kluwer Academic Publishers, Dordrecht, 1990.
- [9] P. K. Simpson. Fuzzy min-max neural networks - part 2: Clustering. *IEEE Trans, Fuzzy Systems*, 1(1):32–45, 1993.
- [10] Y. A. Tolias, S. M. Panas, and L. H. Tsoukalas. Generalized fuzzy indexes for similarity matching. *Fuzzy Sets and Systems*, 1(120):255–270, 2001.
- [11] M. S. Yang. A survey of fuzzy clustering. *Math. Comput. Modelling*, 18(11):1–16, 1993.
- [12] M. S. Yang and H. M. Shih. Cluster analysis based on fuzzy relations. *Fuzzy Sets and Systems*, 1(120):197–212, 2001.