



HAL
open science

A family of cluster validity indexes based on a l-order fuzzy OR operator

Hoel Le Capitaine, Carl Frelicot

► **To cite this version:**

Hoel Le Capitaine, Carl Frelicot. A family of cluster validity indexes based on a l-order fuzzy OR operator. Joint IAPR International Workshops on Structural and Syntactic Pattern Recognition (SSPR 2008), 2008, Orlando, United States. pp.612-621. hal-00441427

HAL Id: hal-00441427

<https://hal.science/hal-00441427v1>

Submitted on 16 Dec 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A family of cluster validity indexes based on a l -order fuzzy OR operator

Hoel Le Capitaine and Carl Frélicot

MIA Laboratory, University of La Rochelle, La Rochelle 17042 Cedex, France,
hoel.le.capitaine@univ-lr.fr , carl.frelicot@univ-lr.fr

Abstract. Clustering is one of the most important task in pattern recognition. For most of partitional clustering algorithms, a partition that represents as much as possible the structure of the data is generated. In this paper, we address the problem of finding the optimal number of clusters from data. This can be done by introducing an index which evaluates the validity of the generated fuzzy c -partition. We propose to use a criterion based on the fuzzy combination of membership values which quantifies the l -order overlap and the intercluster separation of a given pattern.

1 Introduction

The objective of fuzzy clustering is to partition the data set into c distinct clusters. The fuzzy c -means (FCM) algorithm proposed by Bezdek [3] and its variations [8] are probably the most commonly used fuzzy clustering methods. However, these algorithms require the user to set the number c of clusters although the user do not always know it. A fuzzy c -partition obtained by FCM has to be validated because its quality depends on this number. Many cluster validity indexes have been proposed that evaluate each fuzzy c -partition and determines the optimal number c^* of clusters allowing to obtain the optimal partition of the data.

Compactness and separation of clusters are often considered to validate a partition [4, 12, 7]. In this paper, we propose a new family of indexes that combine an overlap measure and a separation measure both based on aggregation of membership values. Section 2 provides background information on basic fuzzy operators and the l -order fuzzy OR operator our work is based on. Next, in section 3, we briefly describe the FCM algorithm and recall some well-known cluster validity indexes we will use for comparison. The measures we propose to use and the definition of the new index are described in section 4. Experimental results on both synthetic and real data sets that show its efficiency and concluding remarks are given in section 5 and 6 respectively.

2 Mathematical Background

For the applications we have in mind, e.g. cluster validity, we are particularly interested in aggregation functions that map a collection \mathbf{u} of c values in $[0, 1]$

(i.e. a vector $\mathbf{u} = [u_1, \dots, u_c]^T$) to a value in $[0, 1]$. Among the frequently used aggregation operators, one can cite the classes of triangular norms (t-norms) and triangular conorms (t-conorms). They have been introduced to characterize the general multivalued logic *AND* and *OR* operations and are widely used in fuzzy logic and fuzzy set theory to implement conjunctive and disjunctive operator respectively. A triangular norm is a commutative, associative and monotone function \top having for neutral element 1. Alternatively, its dual operator, the triangular conorm, is a commutative, associative and monotone function \perp having for neutral element 0. Examples of triangular norms couples are given in Table 1 for two operands, see [10] for a large survey. Note that various parameterized families have been introduced, e.g. the Hamacher family defined by $u_1 \top_H u_2 = \frac{u_1 u_2}{\gamma + (1-\gamma)(u_1 + u_2 - u_1 u_2)}$ and $u_1 \perp_H u_2 = \frac{u_1 + u_2 - u_1 u_2 - (1-\gamma)u_1 u_2}{1 - (1-\gamma)u_1 u_2}$ where $\gamma \in [0, +\infty[$. The dual couple is generally associated with a fuzzy negation defined as $N(u_1) = 1 - u_1$ and mentioned as the triple (\top, \perp, N) .

Table 1. Basic t-norms and t-conorms couples

Standard	$u_1 \top_S u_2 = \min(u_1, u_2)$
	$u_1 \perp_S u_2 = \max(u_1, u_2)$
Algebraic	$u_1 \top_A u_2 = u_1 u_2$
	$u_1 \perp_A u_2 = u_1 + u_2 - u_1 u_2$
Lukasiewicz	$u_1 \top_L u_2 = \max(u_1 + u_2 - 1, 0)$
	$u_1 \perp_L u_2 = \min(u_1 + u_2, 1)$

Assume that the values u_i ($i = 1, \dots, c$) to be aggregated represent the degree to which an object \mathbf{x} satisfies each group description, i.e. its similarity to the prototypes describing each group. Using this knowledge contained in \mathbf{u} , clustering consists in selecting the most appropriate group that the objects will be assigned to. The maximum operator is commonly used in this situation, but we may be interested in the lower values, which interact with the greatest value. In particular, if an object satisfies more than one group description, such an exclusive partitioning is not efficient. A fundamental issue becomes the determination of the overall degree of exclusive belongingness to a group or cluster. In [11], the authors define the l -order fuzzy OR operator. This operator evaluates degrees of satisfaction at a given order by combination of triangular norms. Let \mathcal{P} be the powerset of $C = \{1, 2, \dots, c\}$ and $\mathcal{P}_l = \{A \in \mathcal{P} : |A| = l\}$ where $|A|$ denotes the cardinality of subset A , then the fOR- l is defined by:

$$\bigoplus_{i=1, \dots, c}^l u_i = \bigtop_{A \in \mathcal{P}_{l-1}} \left(\bigoplus_{j \in C \setminus A} u_j \right) \quad (1)$$

It must be viewed as some kind of generalization of the notion of “ l^{th} highest” value, with l in C . In particular, with standard triangular norms, $\bigoplus^l(\mathbf{u})$ is ex-

actly the “ l^{th} highest” element of \mathbf{u} . For instance, let us take $C = \{1, 2, 3\}$, $l = 2$ and use standard triangular norms. We have $\mathcal{P}_1 = \{\{1\}, \{2\}, \{3\}\}$ and

$$\bigwedge_{i=1, \dots, 3}^2 u_i = \min(\max(u_2, u_3) \max(u_1, u_3) \max(u_1, u_2))$$

so that if $u_2 < u_1 < u_3$, $\bigwedge_{i=1, \dots, 3}^2 u_i = u_1$.

This operator satisfies nice mathematical properties such as monotony, symmetry, see [11] for proofs and details.

3 Cluster Validity for Fuzzy Clustering

3.1 Fuzzy c -means algorithm

Clustering is an instance of unsupervised classification which aims at finding a structure of groups in set of n p -dimensional patterns $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$. In this framework, the label vectors $\mathbf{u}_k = u(\mathbf{x}_k)$ do not exist and clustering algorithms can be used to obtain them from X . For instance, the *fuzzy c -means* (FCM) algorithm partitions X into $c > 1$ clusters by minimizing the following objective function [3]:

$$J_m(U, V) = \sum_{k=1}^n \sum_{i=1}^c u_{ik}^m \|\mathbf{x}_k - \mathbf{v}_i\|^2 \quad (2)$$

where u_{ik} is the membership degree of \mathbf{x}_k to the i^{th} cluster represented by its centroid $\mathbf{v}_i \in \mathbb{R}^p$. Centroids are gathered into a $(c \times p)$ matrix $V = [\mathbf{v}_1, \dots, \mathbf{v}_c]$. Degrees u_{ik} are subject to $\sum_{i=1}^c u_{ik} = 1$ for all \mathbf{x}_k in X and to $0 < \sum_{k=1}^n u_{ik} < n$ ($\forall i = 1, \dots, c$). In addition, they are elements of the fuzzy c -partition matrix U ($c \times n$). The so-called *fuzzifier* $m > 1$ is a weighting exponent which makes the resulting partition more or less fuzzy [12]. The higher m is, the softer the clusters’ boundaries are. Minimization of (2) is obtained by iteratively updating (U, V) as follows:

$$u_{ik} = 1 / \sum_{j=1}^c \left(\frac{\|\mathbf{x}_k - \mathbf{v}_i\|}{\|\mathbf{x}_k - \mathbf{v}_j\|} \right)^{2/(m-1)} \quad (3)$$

$$\mathbf{v}_i = \frac{\sum_{k=1}^n u_{ik}^m \mathbf{x}_k}{\sum_{k=1}^n u_{ik}^m} \quad (4)$$

The usual euclidian norm $\|\cdot\|$ induces hyperspherical clusters, hence FCM can only detect clusters with the same shape and orientation.

3.2 Classical indexes

Validating the provided clustering of X consists in assessing whether the resulting partition reflects the data structure or not. Since c is a user-defined parameter of clustering algorithms such as FCM, most of works on cluster validity focus on the number of clusters problem. Many validity indexes have been proposed for fuzzy clustering (refer to [4, 9, 14, 13, 16] for comparative studies). They can be classified in two main categories. The first one is composed of indexes that only use membership degrees (U). Let us cite the *Partition Coefficient* [3], taking values in $[\frac{1}{c}, 1]$:

$$PC(c) = \frac{1}{n} \sum_{k=1}^n \sum_{i=1}^c u_{ik}^2 \quad (5)$$

or the *Partition Entropy* [2], taking values in $[0, \log(c)]$:

$$PE(c) = -\frac{1}{n} \sum_{k=1}^n \sum_{i=1}^c u_{ik} \log(u_{ik}) \quad (6)$$

Both PC to be maximized and PE to be minimized are monotonic with c , as well as their bounds. Normalized versions have been proposed to reduce this monotonic tendency, *e.g.* in [6]. We will use these normalized versions in the experiments: $NPC(c) = \frac{cPC(c)-1}{c-1}$ and $NPE(c) = \frac{nPE(c)}{n-c}$. The second category consists of indexes that use membership degrees but also some information about the geometrical structure of the data (U, V, X), *e.g.* the Xie-Beni index [12, 15]:

$$XB(c) = \frac{J_m(U, V) / n}{\min_{i,j=1,\dots,c;j \neq i} \|\mathbf{v}_i - \mathbf{v}_j\|^2} \quad (7)$$

or the Fukuyama-Sugeno index [7]:

$$FS(c) = J_m(U, V) - \sum_{k=1}^n \sum_{i=1}^c u_{ik}^m \|\mathbf{v}_i - \bar{\mathbf{v}}\|^2 \quad (8)$$

where $\bar{\mathbf{v}}$ is the mean of centroids. Both XB and FS combine the FCM objective function (2) which measures the degree of compactness of the clusters and an additional term which measures the degree of their separation. Combination indicates that both indexes are to be minimized. The more compact and separated the clusters are, the more optimal c is.

4 The Proposed Index

Let $U = [\mathbf{u}_1, \dots, \mathbf{u}_c]$ be a fuzzy c -partition provided by a fuzzy clustering algorithm, *e.g.* FCM. An overlap measure between l fuzzy clusters for each point \mathbf{x}_k in X described by its membership degrees can be obtained by Eq. (1). By successively computing $\perp^l(\mathbf{u}_k)$ for different values of c , we get a combination

of l -order overlap measure for \mathbf{x}_k . The most satisfied order is obtained by the disjunction of these measures, and we define the overall overlap measure as:

$$O(\mathbf{u}_k, c) = \underset{l=2,c}{\perp} \left(\underset{i=1,\dots,c}{\perp} u_{ik} \right) \quad (9)$$

In [4], Bezdek and Pal show that intercluster separation plays a more important role in cluster validity than cluster diameters. We propose to introduce such a measure by quantifying the separation of each point \mathbf{x}_k with $\underset{1}{\perp}(\mathbf{u}_k)$ which is the overlap measure between 1 fuzzy cluster, i.e. its separation from the other fuzzy clusters, since \mathbf{u}_k sum up to one. Note that if standard triangular norms are used, $\underset{1}{\perp}(\mathbf{u}_k)$ is the maximum coordinate of \mathbf{u}_k . Finally, we define the family of l -order Fuzzy OR Indexes as:

$$lFORI(c, \top, \perp) = \frac{1}{n} \sum_{k=1}^n \frac{O(\mathbf{u}_k, c)}{\underset{i=1,\dots,c}{\perp} u_{ik}} \quad (10)$$

Given U , the less overlapping and separated clusters are, the lower the value of $lFORI$ is expected to be and minimizing Eq. (10) will give the optimal number c^* of clusters.

If U is hard, i.e. $u_{ik} \in \{0, 1\}$, then one value equals 1 while the others are 0, say $u_{1k} = 1$ and $u_{2k} = \dots = u_{ck} = 0$. Since 0 is the absorbing element of \top , it is easy to verify that $O(\mathbf{u}_k, c) = 0$ for all \mathbf{u}_k , whatever (\top, \perp) , therefore $lFORI(c, \top, \perp) = 0$:

$$\begin{aligned} \underset{l=2,c}{\perp} \left(\underset{i=1,\dots,c}{\perp} u_{ik} \right) &= \left(\underset{i=1,\dots,c}{\perp} (1, 0, \dots, 0) \right) \underset{1}{\perp} \dots \underset{1}{\perp} \left(\underset{i=1,\dots,c}{\perp} (1, 0, \dots, 0) \right) \quad (11) \\ &= \underset{1}{\perp} (\underbrace{0, \dots, 0}_{c-1 \text{ times}}) = 0. \quad (12) \end{aligned}$$

and $\underset{1}{\perp}(1, 0, \dots, 0) = 1$, since 1 is the absorbing element of \perp .

On the other hand, if U is totally fuzzy, i.e. $u_{ik} = \frac{1}{c}$ ($\forall i = 1, \dots, c$), the resulting $lFORI$ value depends on the couple (\top, \perp) because Eq. (9) only reduces to

$$\underset{l=2,c}{\perp} \left(\underset{i=1,\dots,c}{\perp} u_{ik} \right) = \left(\underset{i=1,\dots,c}{\perp} \left(\frac{1}{c}, \dots, \frac{1}{c} \right) \right) \underset{1}{\perp} \dots \underset{1}{\perp} \left(\underset{i=1,\dots,c}{\perp} \left(\frac{1}{c}, \dots, \frac{1}{c} \right) \right) \quad (13)$$

in the general case. If standard t-norms are used, we have:

$$\bigoplus_{l=2,c}^1 \left(\bigoplus_{i=1,\dots,c}^l u_{ik} \right) = \bigoplus \left(\underbrace{\frac{1}{c}, \dots, \frac{1}{c}}_{c-1 \text{ times}} \right) \quad (14)$$

$$= \frac{1}{c} \quad (15)$$

which is the value of $\bigoplus(\mathbf{u}_k)$ for all \mathbf{u}_k , therefore $lFORI(c, \top_S, \perp_S) = 1$ for all c . Unfortunately, it is not possible to give a simple value for $lFORI$ for other couples (\top, \perp) , but an upper bound can be found. Due to lack of space, the proof is postponed to a forthcoming long paper, as well as properties that could help the user to choose the couple (\top, \perp) .

5 Experiments

5.1 Behavior according to clusters' separability

First, we generated a series of 10 data sets, each composed of 800 points drawn from a mixture of $c = 4$ bivariate normal distributions. The covariance matrix of each component is the same $\Sigma_i = I$ ($\forall i = 1, \dots, c$) and the mean vectors are: $\mu_1 = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\mu_2 = \alpha \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\mu_3 = \alpha \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ and $\mu_4 = \alpha \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, for increasing values of $\alpha = 1, 2, \dots, 10$. This successively moves the clusters in opposite directions, creating less overlap as the clusters become more and more separated. The first and last data sets are shown in Figure 1-*left*. Each data set was then clustered using FCM with $c = 4$, providing a fuzzy partition matrix U_α . Corresponding values of $lFORI$ for the different basic norms are plotted in Figure 1-*right* as a function of α . As expected, the proposed validity index decreases towards 0 as α increases whatever the couple (\top, \perp) .

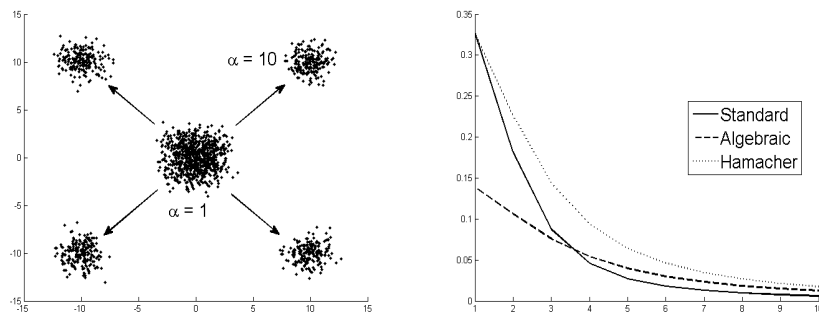


Fig. 1. α -separated data sets – $\alpha = 1$ and 10 (*left*) and values of $lFORI(c = 4)$ as a function of α for various t-norms (*right*).

5.2 Robustness to outliers

In order to compare the proposed index to the classical ones, we generated a data set D_1 containing $n = 200$ points consisting of 50 points each drawn from a mixture of $c = 4$ bivariate normal distributions, see Figure 2-*left*. FCM was used with $m = 2$ for c varying from $c_{min} = 2$ to $c_{max} = 10$. A second artificial data set D_2 was generated. It is similar to D_1 except that 100 points drawn from a uniform distribution were added, as shown in Figure 2-*right*. These additional points act as noise and can make the FCM algorithm partitioning the data set into three clusters because the less separated groups in the right-lower area tend to become only one cluster. Values of the tested validity indexes on D_1 and D_2 are given in Table 2 and Table 3 respectively, where optimal values are bold faced and acceptable ones are italicized. Even if the classical indexes are known to be efficient, most of them fail in giving the right number of clusters on D_1 and with even stronger reason in presence of noise (D_2), whereas *lFORI* always gives the right number $c^* = 4$ whatever the couple (\top, \perp) for both data sets.

Moreover, multiple runs of FCM with random initializations on data set D_2

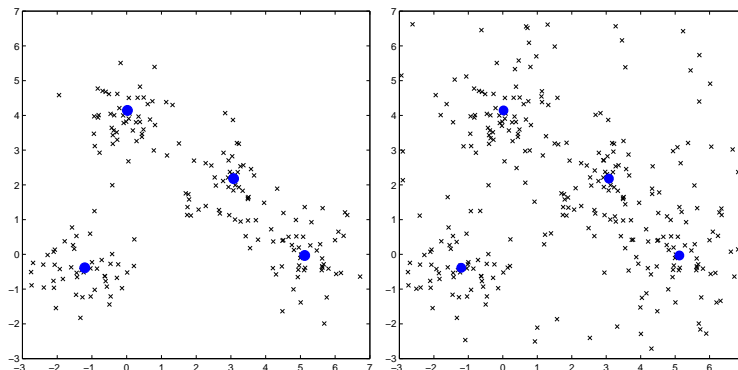


Fig. 2. Centroids of optimal clusters for data sets D_1 (*left*) and D_2 (*right*).

showed us that the proposed index gives more stable results compared to the others, thus demonstrating a significantly higher robustness to noisy data.

5.3 Sensitivity to the fuzzifier m

The FCM objective function depends on the fuzzy exponent m , see Eq. (2). Since the resulting fuzzy c -partition U is sensitive to this parameter, a cluster validity index can also be analysed with respect to m . In [4], Pal and Bezdek have shown that FCM provides best results for m lying in $[1.5, 2.5]$. We compare the different indexes in this range on data sets D_1 and D_2 . Again, c is varying from $c_{min} = 2$ to $c_{max} = 10$. The selected number of clusters are reported in Table 4. As it can be seen, the proposed index *lFORI* is less sensitive to m whatever the couple (\top, \perp) for both data sets.

Table 2. Validity indexes on data set D_1 .

c	NPC	NPE	XB	FS $\times 10^{-3}$	$lFORI$		
					$(\top, \perp)_S$	$(\top, \perp)_A$	$(\top, \perp)_H$
2	0.57	0.52	0.13	-1.57	0.17	0.14	0.15
3	0.70	0.56	0.07	-3.39	0.12	0.12	0.14
4	0.74	0.59	<i>0.08</i>	-0.74	0.07	0.07	0.12
5	0.67	0.80	0.31	-4.25	0.12	0.11	0.17
6	0.61	1.00	0.45	-2.98	0.18	0.13	0.18
7	0.56	1.15	0.37	-1.53	0.19	0.10	0.20
8	0.530	1.27	0.30	-1.60	0.23	0.09	0.19
9	0.532	1.37	0.31	-2.43	0.29	0.10	0.19
10	0.50	1.47	0.28	-0.15	0.28	0.11	0.17

Table 3. Validity indexes on data set D_2 .

c	NPC	NPE	XB	FS $\times 10^{-3}$	$lFORI$		
					$(\top, \perp)_S$	$(\top, \perp)_A$	$(\top, \perp)_H$
2	0.49	0.58	0.23	-0.85	0.25	0.17	0.19
3	0.60	0.71	0.11	-2.33	0.23	0.11	0.20
4	<i>0.59</i>	0.87	0.13	-0.75	0.17	0.09	0.15
5	0.55	1.08	0.24	-5.21	0.18	0.13	0.21
6	0.50	1.26	0.35	-0.06	0.21	0.14	0.22
7	0.50	1.36	0.36	-0.12	0.21	0.17	0.22
8	0.46	1.50	0.44	-2.41	0.26	0.16	0.21
9	0.45	1.61	0.46	-2.61	0.30	0.16	0.19
10	0.44	1.70	0.36	-1.81	0.30	0.15	0.18

5.4 Benchmark data sets

We finally compare the different indexes on benchmark data sets:

- *Iris* [5], composed of three classes of 50 flowers each described by 4 physical attributes. Two classes have a substantial overlap in the feature space and the optimal number of clusters to be found is debatable: 2 or 3, see [4].
- *Wine* [5], which consists of 13 chemical attributes for $n = 178$ italian wines, divided into three classes. Classes are well separable, so the indexes found in the literature generally give the right number of clusters.
- *Wisconsin Breast Cancer* [5], composed of $n = 699$ malignant/benign cells described by 9 features computed from digitized images.
- The artificial data set *X30* introduced in [4], consisting in $n = 30$ observations in \mathbb{R}^2 , for which three clusters are expected.
- The bidimensional artificial data set *Bensaid* [1] characterized by 3 classes of very different cardinalities (6, 3 and 40).
- The original *Starfield* [15], which contains the position and light intensity of $n = 51$ bright stars near Solaris. The expected number of clusters is 8 or 9, depending on the papers.

Table 4. Selected number of clusters in data sets D_1 and D_2 for different values of m .

data set	m	NPC	NPE	XB	FS	$lFORI$		
						$(\top, \perp)_S$	$(\top, \perp)_A$	$(\top, \perp)_H$
D_1	1.5	4	4	3	5	4	4	4
	1.7	4	4	4	4	4	4	4
	1.9	4	2	3	4	4	4	4
	2.1	4	2	3	5	4	4	4
	2.3	4	2	3	5	4	4	4
	2.5	3	2	4	6	4	3	4
D_2	1.5	5	4	3	5	4	4	4
	1.7	4	4	3	3	4	4	4
	1.9	4	2	3	4	4	4	4
	2.1	3	2	3	5	4	4	4
	2.3	4	2	3	3	4	3	4
	2.5	4	2	4	4	4	3	4

Table 5 summarizes the results obtained on these data sets. The c^* column gives the expected number of clusters and the other columns show the optimal number of clusters obtained using the validity indexes. The proposed index always finds the optimal number of clusters whatever the couple (\top, \perp) while some others do not.

Table 5. Selected number of clusters in benchmark data sets.

Data set	c^*	NPC	NPE	XB	FS	$lFORI$		
						$(\top, \perp)_S$	$(\top, \perp)_A$	$(\top, \perp)_H$
Iris	2 or 3	2	2	2	3	2	2	2
Wine	3	3	2	3	5	3	3	3
Breast	2	2	2	2	3	2	2	2
X30	3	3	2	3	7	3	3	3
Bensaid	3	3	2	3	7	3	3	3
Starfield	8 or 9	2	2	6	7	8	8	8

6 Conclusion

A new family of cluster validity indexes for fuzzy partitions has been proposed. These indexes combine a new measure of overlap of clusters and a separation measure. The novelty of the approach is that, for each data point to be clustered, the relative importance of each membership degree and the relationship of the degrees are taken into account through a combination of triangular norms (\top, \perp) . Results obtained on artificial and benchmark data sets have shown that the proposed family of indexes is most of time more efficient than well-known

ones, less sensitive to the fuzzifier exponent and particularly robust in noisy environments.

Further results about properties of the proposed indexes based on the mathematical properties of the aggregation operators involved as well as guidelines to choose a member (\top, \perp) of the family appropriated to specific situations will come soon.

References

1. A. M. Bensaid, L. O. Hall, J. C. Bezdek, L. P. Clarke, M. L. Silbiger, J. A. Arrington, and R. F. Murtagh. Validity-guided (re)clustering with applications to image segmentation. *IEEE Transactions on Fuzzy Systems*, 4(2):112–123, 1996.
2. J. C. Bezdek. Cluster validity with fuzzy sets. *Journal of Cybernetics*, 3:58–72, 1974.
3. J. C. Bezdek. *Pattern Recognition with fuzzy objective function algorithm*. Plenum Press, 1981.
4. J.C. Bezdek and N.R. Pal. Some new indexes of cluster validity. *IEEE Transactions on Systems, Man and Cybernetics*, 23(3):301–315, 1998.
5. C. Blake and C. Merz. Uci repository of machine learning databases, 1998. <http://www.ics.uci.edu/mllearn/MLRepository.html>.
6. R. N. Dave. Validating fuzzy partitions obtained through c -shells clustering. *Pattern Recognition Letters*, 17(6):613–623, 1996.
7. Y. Fukuyama and M. Sugeno. A new method for choosing the number of clusters for the fuzzy c -means method. In *Proc. 5th Fuzzy Systems Symposium*, pages 247–250, 1989.
8. D.E. Gustafson and W.C. Kessel. Fuzzy clustering with fuzzy covariance matrix. In *Proc. IEEE Conference on Decision and Control*, pages 761–766, San Diego, California, 1979.
9. D-W. Kim, K.H. Lee, and D. Lee. On cluster validity index for estimating the optimal number of fuzzy clusters. *Pattern Recognition*, 37(3):2009–2025, 2004.
10. E.P Klement and R. Mesiar. *Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms*. Elsevier, 2005.
11. L. Mascarilla, M. Berthier, and C. Frélicot. A k -order fuzzy or operator for pattern classification with k -order ambiguity rejection. *Fuzzy Sets and Systems*, 159(15):2011–2029, 2008.
12. N.R. Pal and J.C. Bezdek. On cluster validity for the fuzzy c -means model. *IEEE Transactions on Fuzzy Systems*, 3(3):370–379, 1995.
13. W. Wang and Y. Zhang. On fuzzy cluster validity indices. *Fuzzy Sets and Systems*, 158(19):2095–2117, 2007.
14. K.L. Wu and M.S. Yang. A cluster validity index for fuzzy clustering. *Pattern Recognition Letters*, 26(9):1275–1291, 2005.
15. X. L. Xie and G. Beni. A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8):841–847, 1991.
16. M. H. F. Zarandi, E.Neshat, and I. B. Türksen. A new cluster validity index for fuzzy clustering based on similarity measure. In *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, 11th International Conference*, pages 127–135, 2007.