ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 7, July 2013

Logical Actions of trees for the comparison of Classification Methods

Najma HAMZAOUI[1], Maha AKBIB[2], Wafae BAIDA[3], Abdelfettah SEDQUI[4], Abdelouahid LYHYAOUI[5]

Abdelmalek Essaâdi University, LTI Lab, ENSA of Tangier, Tangier, Morocco

PO Box: 1818, Tangier Principal

Abstract: In this paper, a new approach is presented to compare methods. Many works compare classification methods based directly on generated partitions. We propose, instead, a classification methods comparison, based on the structure of logic trees. As result, we obtain grouping of methods and can even measure the distance between them. In this purpose, we apply this new approach in two cases. The first case is related to improved performance of assembly system at the dynamic management flexibility. The second case deals with the similarity coefficients as key decision able to classify methods. The results are very promising.

Keywords***:*** hierarchical classification, classification methods comparaison, similarity coefficients, structure of tree

I. Introduction

The comparison of classification is one of the open questions in data analysis. The need to compare two classifications occurs when the study of two investigations on the same data set. When there are two partitions made on the same individuals, for example, with two sets of variables or two algorithms, it is necessary to know if these partitions are consistent or if they differ significantly in a sense to be clarified. To address this problem, a wide number of studies related to results of classification comparison have been performed. These works were devoted to the presentation and the definition of different indices to compare partitions. Most of these indices are presented in relational formulations [2,3]. [8,11] cited Cohen's kappa as a coefficient to test the similarity between two partitions by comparing their simulated distribution from the same partition. A well-known Rand index [4, an asymmetric version of Rand [12,13,15] and the corrected Hubert index [5] are used for the comparison of nested partitions with different number of classes. In [14], two indices are inspired by McNemar test and Jaccard index. Whereas some others researchers recommended other coefficients , such as the vectorial correlation index introduced by [6] which is identic to the coefficient S.Janson and Vegelius index redundancy proposed by [7,9] and index of popping [10]. These studies for comparing partitions have achieved success both in theory and practice. However, in these approaches, comparison has been done directly in the resulting classes, and not in the dendrograms of these methods of classification that give these partitions. That is to say, the comparison based on the structure of these partitions is not taken into consideration. Therefore, a new way of comparative research is crucially necessary to evaluate the various methods. For example, if we take similarity coefficients like as an methods of classification, we find researchers have studied differently the comparison of [17,16], investigated two characteristics discriminability and stability to compare the performance of 20 well-known similarity coefficients. Other method was used for comparing two hierarchical classifications by associating each hierarchical structure, an ultrametric matrix, and then calculating the Spearman coefficient between the two ultrametric matrix [14].Our work will focuses on comparing methods of classification based on calculating the distance between their dendrograms from the hierarchical classification.

The objective is to find a formal procedure based on the structure of logic trees from logical actions for comparing two classifications. We obtain these structures of logic trees or hierarchical tree using methods of classifying variables, and compare them from the distance Marczewski-Steinhaus [1]. This proposal comparison procedure of hierarchical trees aims to provide a rational and effective way for grouping different classification methods to be a general approach, including the choice of the closest and the alternative methods. Our goal is to propose an method universal applicable to many fields such as industry, bioinformatics, philology, botany, sciences socials, social studies, astronomy, images, Marketing, diagnosis, medicine, astronomy, economic. It aims to:

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 7, July 2013

- Fill a gap by providing a method for the comparison of several methods of classification based on their dendrograms from the hierarchical classification.

- Compare methods of hierarchical classification (based on measures of similarity) based on their dendrograms or results algorithms turned into dendrogram to situate them in relation to the other on the dendrogram of the proposed method,

- To find classes of methods and recommend the best.

- To classify each new method of literature and relate it to the best and poor classes.

We will take two examples of application for this method. The improved performance of the assembly system at the dynamic management of production flexibility and responsiveness in order to reduce the number of lines by keeping a better range associated with ranges of replacement. The second case is of comparing the 20 well-known similarity coefficients methods which are taken from the literature.

The rest of this paper is structured as follows. Section 2 presents our proposal method of the classification methods comparison based on their dendrograms with applications examples. Finally, Section 3 concludes this paper with some perspectives.

II. The Proposed Approach

In this section, we present our approach. The theoretical framework of the study is exposed before the presentation of the two applications.

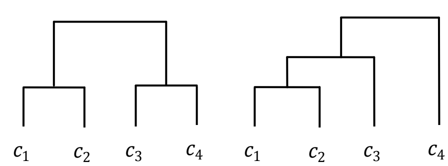

Fig. 1. Trees assembling

Our method aims to compare the partitions derived from methods of hierarchical classification, based on the measures of resemblance on their dendrograms.

For example, we consider in Fig.1 two possible hierarchical classifications for the same data $\{c_1, c_1, c_2, c_3, c_4\}$. We seek for a degree of resemblance between these two dendrograms. In the following, we explain in detail how to do this. In the case if we don't have the hierarchical classification of partitions, results of classification methods must turned into dendrogram to apply our method by ascendant agglomerative methods. Then, to provide the comparison of several methods of classification, our method consists in applying a partitioning algorithm of the distance matrix of these methods of classification. We use a hierarchical clustering algorithm taking into account the criterion of the average link. After that, we get a number of groups by making cuts from the classification tree, the choice of threshold depends on the criteria (number of alternative methods of classification, the closest partitions for a given partition, etc) that determine our goals. This method has the advantage of providing a global view of coherent methods.

A. *Hypergraphs generated by trees*

According to [1] the assembly trees treated are a special case of hypergraphs generated by tree, whose family nodes, has special properties.

Let $X=\{x_1, x_1, x_2, x_3, \ldots, x_n\}$ the set of terminal vertices of a tree.

$d^-(x_i) = 1$, $d^+(x_i) = 0$ for every element of X where $d^-(x_i)$ et $d^+(x_i)$ represent, respectively, interior and exterior halves degree of the node x_i . Let A the class of all trees. Let $A \in A$ represented by the hypergraph $(X, E_A)$where the $E_A$class edges as defined below: each $v \notin X$ (i.e. each non-terminal node in the tree generates $d^+(v) - 1$ edges in E_A . Such an edge consists of those elements of X which are terminal nodes of the subtree generated by v which is obtained by considering as v a root ie, assuming that $d^-(v) = 0$.

The construction method of the hypergraph H_A led to the following insertion:

(i) If $H_A = (X, E_A)$ is the hypergraph generated by a tree $A \in \mathcal{A}$ as described above, then $|E_A| = n - 1$

(ii) The hypergraph H_A generated by A $\in \mathcal{A}$ is not simple if at least one of the nodes $v \in A$ such as $d^+(v) \geq 2$. By definition, a hypergraph is simple if all its edges are distinct.

B. *Distance between trees*

Let X be a finite set such that $|E_A| = n$; where $|.|$ is the cardinality of the set. Let E^* is the class of all subsets of X, and $\mu(E)$ measurement of E on E^* .

Consider $\mu(E) < \infty \; \forall E \in E^*$. The distance Marczewski-Steinhaus [1] between two sets E_1 and E_2 de E^* is:

$$\sigma_\mu(E_1, E_2) = \begin{cases} \dfrac{\rho(E_1, E_2)}{\mu(E_1 \cup E_2)} & \text{si } E_1 \cup E_2 > 0 \\ 0 & \text{si } E_1 \cup E_2 = 0 \end{cases} \quad (1)$$

With$\rho(E_1, E_2) = \mu(E_1 \Delta E_2)$, Δ is the symmetric difference.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 7, July 2013

Consider that$0 \leq \sigma_\mu(E_1, E_2) \leq 1$. Especially, if we consider that $\mu_c(E) = |E|$ and then we take $e_1 = |E_1|$, and $e_2 = |E_2|$ and $d = |E_1 \cap E_2|$.

$$\sigma_{\mu c}(E_1, E_2) = \frac{e_1 + e_2 - 2d}{e_1 + e_2 - d} \qquad (2)$$

We have also: $0 \leq \sigma_{\mu c}(E_1, E_2) \leq 1$

Consider A_1 andA_2, elements of A respectively represented by hypergraphs $H_{A1} = (X, E_{A1})$ and $H_{A2} = (X, E_{A2})$. The distance between these hypergraphs takes into account the specific stage of edges construction. The distance between trees is given by the following formula:

$$d(A_1, A_2) = \frac{1}{n-1} \min_{p \in P} \sum_{i=1}^{n-1} \sigma_\mu\left(E_{A_1}^i, E_{A_2}^{p_i}\right) \qquad (3)$$

where p_i est le i^{th} element of the permutation p des n-1 integers. P is the set of all permutations, $\sigma_\mu(.,.)$ is given above $E_{A_1}^i \in E_{A_1}$ and $E_{A_2}^{p_i} \in E_{A_2}$ i=1 a n-1.

The following facts are implied by the previous definition:

(i) (A,d) is a metric space
(ii) $d(A_1, A_2) \leq 1$, A_1 and $A_2 \in$ A. the distance d(.,.) is strictly less than 1 if we uses $\sigma_{\mu c}(.,.)$ instead of $\sigma_\mu(.,.)$.

In general, this distance is valid for n-areas trees. We use the binary case because the selected structure for ranges description is that of binary trees [18]

C. *Example of distance between trees*

Let X={c_1, c_1, c_2, c_3, c_4} are all components of a product, and A_1, A_2 and two possible assemblages trees as shown in fig.1. We calculate the proposed distance between the two ranges A_1 and A_2 of $\mathcal{A}$, based on the set of all the components $\mathcal{X}$. $|X| = 4$.

To do this, we seek sets $E_{A_1}$and E_{A_2}sub-trees corresponding to the intermediate stages of the formation of the product. These steps are the edges of the hypergraphs $H_{A1} = (X, E_{A1})$ and $H_{A2} = (X, E_{A2})$. According to proposition 1, the number of intermediate steps for the formation of the product $|E_{A_1}|$=$|E_{A_2}|$=|X|-1=3.
To simplify notations, we set $c_i = i$:

$E_{A_1} = \{\{1,2\}, \{3,4\}, \{1,2,3,4\}\}$ with $E_{A_1}^1 = \{1,2\}$,. $E_{A_1}^2 = \{3,4\}$ and $E_{A_1}^i = \{1,2,3,4\}$.
$E_{A_2} = \{\{1,2\}, \{1,2,3\}, \{1,2,3,4\}\}$

The distance $d(A_1, A_2)$ given by the formula given by [1] is calculated between the components of $E_{A_1}$and components of permutations of E_{A_2}. For this, we look for the set P of permutations p of E_{A_2}. P is described as:

$\{\{1,2\}, \{1,2,3\}, \{1,2,3,4\}\}$

With $E_{A_1}^1 = \{1,2\}$,$E_{A_1}^2 = \{1,2,3\}$

and $E_{A_1}^i = \{1,2,3,4\}$

$\{\{1,2\}, \{1,2,3,4\}, \{1,2,3\}\}$

avec $E_{A_1}^1 = \{1,2\}$,$E_{A_1}^2 = \{1,2,3,4\}$

and $E_{A_1}^i = \{1,2,3\}$

$\{\{1,2,3\}, \{1,2\}, \{1,2,3,4\}\}$

The same approach is applied to search $E_{A_2}^{p_i}$

$\{\{1,2,3\}, \{1,2,3,4\}, \{1,2\}\}$
$\{\{1,2,3,4\}, \{1,2\}, \{1,2,3\}\}$
$\{\{1,2,3,4\}, \{1,2,3\}, \{1,2\}\}$

We calculate the distance between $E_{A_1}$and permutations of E_{A_2}. The minimum values gives us the distance between the trees $A_1$and A_2. In the case treated above, $d(A_1, A_2) = 0.25$

III. APPLICATIONS

A. *Case of assembly*

We take from [20, 21] the case of assembly a ball pen described in Fig.2. It is based mainly here on the determined ranges in the LAB, to perform our work and show the advantages of the method.

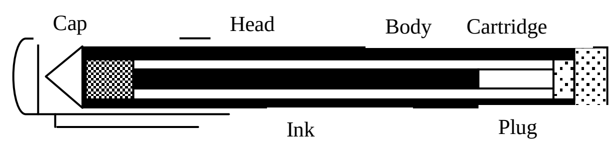

Fig.2. The ball pen

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 7, July 2013

Fig.3.The 10 ranges assembly pen

III.A.1 Notations

First note the correspondence between the different assertions above and the physical or logical properties about the product:

X ------> the set of all product components to assemble

$\mathcal{A}$ -----> the set of all eligible ranges

E_{A_1}----> the set of subtrees of tree A_1, describing intermediate steps of assembling $H_{A1} = (X, E_{A1})$ implies the definition above for the product with components belong to X. Fig. 3 shows all eligible ranges of pen obtained from LEGA software developed in LAB and describing the set $\mathcal{A}$.

Remarks:

- Two-disjoint subsets correspond to the elements of product assembled in parallel,
- Two identical subsets correspond to the case where the assembling tree is not binary,
- If we class elements E_{A_1}, then, the latter element describes the finished product.

III.A.2 Calculation of distances:

The same reasoning use in II.C, applies to the calculation of distances between the pen ranges. The ball pen is constitued by a head (he), a cartridge (cr), ink (in), a body (bd), a plug (pl), and a cap (cp), .

All components of the pen is:

X={he, cr, in, bd, pl, cp}

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 7, July 2013

For example, for the first two trees of assembling in the Fig. 3, the sets intermediate steps for assembling the pen E_{A_1} and E_{A_2} are given. They are directly involved in the calculation of the distance between A_1 and A_2 as described in the example in paragraph 2.2.

$|E_{A_1}|=|E_{A_2}|=|X|-1=5.$

$E_{A_1} = \{\{he, cr\}, \{bd, pl\}, \{he, cr, in\}, \{he, cr, in, bd, pl\}, \{te, cr, in, cp, co, bo\}\}$

$TE_{A_2} = \{\{he, cr\}, \{he, cr, in\}, \{he, cr, in, bd\}, \{he, cr, in, bd, pl\}, \{he, cr, in, bd, pl, cp\}\}$

$d(A_1, A_2)=0.16$

The calculated distances between the different assembly trees are given in Table I.

TABLE I. MATRIX OF DISTANCES BETWEEN THE TREES OF ASSEMBLY

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}
A_1	0.000	0.160	0.386	0.326	0.200	0.470	0.316	0.266	0.250	0.226
A_2		0.000	0.380	0.246	0.100	0.440	0.300	0.166	0.233	0.066
A_3			0.000	0.133	0.140	0.100	0.080	0.213	0.146	0.319
A_4				0.000	0.146	0.233	0.213	0.080	0.280	0.180
A_5					0.000	0.380	0.200	0.066	0.313	0.166
A_6						0.000	0.180	0.313	0.246	0.313
A_7							0.000	0.133	0.066	0.233
A_8								0.000	0.200	0.100
A_9									0.000	0.300
A_{10}										0.000

III.A.3 *Exploitation*

From the distance matrix, two methods can be applied to obtain results:

From the distance matrix, two methods can be applied to obtain results:

- The direct method which is to choose first the best range according to a criterion such as the presence of maximum parallelism between operations (this feature allows the reactivity). In this case, the range ensuring maximum parallelism is the range of minimum depth. The application to the case of the pen, gives the range A_1 as the best in the sense of parallel operations. Alternative ranges are then selected based on their proximity to the sense of distance. Their number is usually set arbitrarily by the company, commonly 0-3. If we determine the three ranges closest to A_1 range, we obtain a sequence of ranges entirely ordered in the proximity described by the set $Groupe(A_1) = \{A_2, A_{10}, A_5\}$
- The second method which is based on the classification is to apply a partitioning algorithm of the distance matrix. We use a hierarchical clustering algorithm taking into account the criterion of the maximum diameter. Then, from the classification tree given in Fig.4 and by making cuts in the tree, we get a number of families of assembly ranges. The choice of the threshold depends on the criteria (number of ranges replaced, the closest ranges to a given range, etc), which determines the objectives of the company. This method has the advantage of providing a global view of coherent families' ranges.

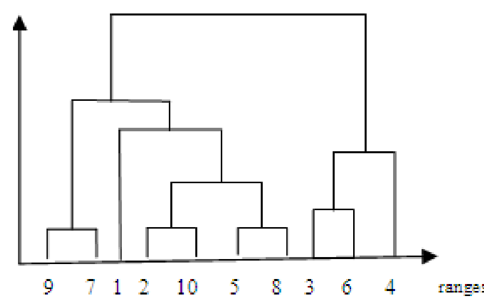

Fig.4. Classification tree ranges, (A_i denote i;i =1,10)

In search of a family which forms the best range and ranges replacement, it is necessary to cut the classification tree at an appropriate level. For example, a group made up of a single range represents little interest. The A_1 range being here the best, the cut must be made at a level $\alpha > 0.266$ to a level $0.266 < \alpha < 0.313$, we obtain the three families of ranges in Fig. 5.

$$\begin{cases} Family\ F_1 : F_1 = \{A_1, A_2, A_5, A_8, A_{10}\} \\ Family\ F_2 : F_2 = \{A_7, A_9\} \\ Family\ F_3 : F_3 = \{A_3, A_4, A_6\} \end{cases}$$

Fig.5. coherent family ranges

Level

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 7, July 2013

We find that the family F_1 contains the best range and range replacement which number depends on the level of the cut.

In this first case, the method used to solve the general problem of classification ranges for the analysis of different logical ranges of a product for their selection. The distance Marczewski-Steinhaus is particularly adapted for solving this problem. To improve the reactivity of the assembly system and thereby its robustness, we propose a general approach, including the choice of a better range in the sense of a criterion such as the maximum parallelism of operations and of its range of replacement to overcome the inevitable vicissitudes of production, such as ,component failure or unavailability of a machine. So, the method exploits the classification tree to give families, thereby, one of them contains the best range and the range of alternatives.

B. *Case of comparing similarity coefficients:*

III.B.1Data Base:

We take from [16,19] the initial data matrix (Table II). It has 8 machines which groups should be identified in order to create the production cells. Each cell will contain a number of machines that processes a product family. It is based here on the 20 similarity indices compared in [16] to classify the machines matrix in families.

TABLE II. THE INITIAL MATRIX OF 8 MACHINES

		Parts																			
Machines		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	1	0	1	1					1	1		1		1	1		1	1		1	
	2			1	1		1	1							1				1		1
	3		1						1	1		1		1	1		1	1		1	
	4			1	1		1	1			1								1		1
	5	1				1	1				1		1			1		1			
	6	1				1				1	1		1			1					1
	7			1	1		1	1				1	1						1		1
	8			1	1		1	1											1		1

Figure 6 shows all obtained classifications/partitions.

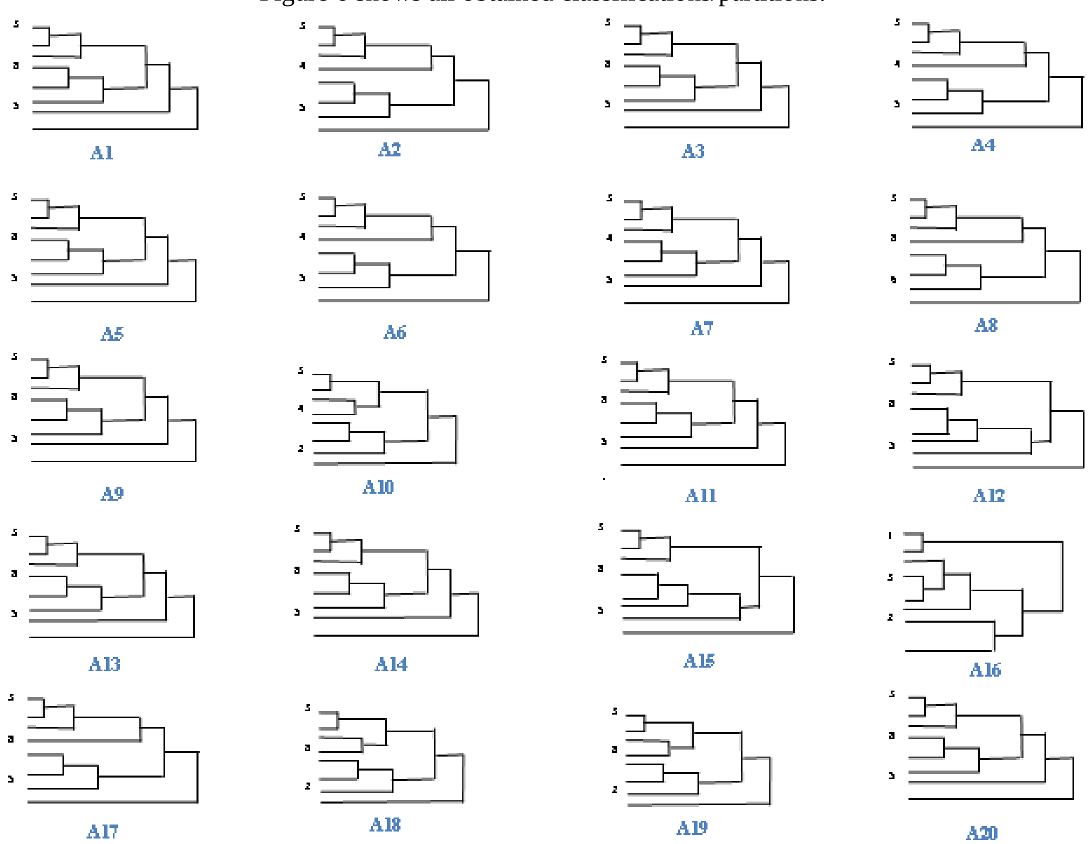

Fig.6. The classification of 20 similarity indices

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 7, July 2013

TABLE III. MATRIX OF DISTANCES BETWEEN THE TREES OF INDICES OF COEFFICIENTS

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}	A_{16}	A_{17}	A_{18}	A_{19}	A_{20}
A_1	0	0.3810	0	0.3810	0.0357	0.3095	0.3146	0.3478	0.0357	0.3748	0.0357	0.0816	0.0357	0.0357	0.0816	0.5722	0.1097	0.3387	0.3388	0
A_2		0	0.3810	0	0.3810	0.1071	0.0816	0.5000	0.3810	0.1857	0.3809	0.3963	0.3810	0.3810	0.3963	0.5607	0.3238	0.4976	0.4976	0.3810
A_3				0.3810	0.0357	0.3095	0.3146	0.3478	0.0357	0.3748	0.0357	0.0816	0.0357	0.0357	0.0816	0.5723	0.1097	0.3388	0.3388	0
A_4					0.3810	0.1071	0.0816	0.5000	0.3809	0.1857	0.3809	0.3963	0.3810	0.3810	0.3963	0.5607	0.3238	0.4976	0.4976	0.3810
A_5					0	0.3095	0.3121	0.3197	0	0.3476	0	0.1097	0	0	0.1097	0.5697	0.0816	0.3048	0.3048	0.0357
A_6						0	0.1862	0.5000	0.3095	0.2571	0.3095	0.3172	0.3095	0.3095	0.3172	0.5607	0.2524	0.4677	0.4677	0.3095
A_7							0	0.4578	0.3120	0.1811	0.3120	0.3810	0.3121	0.3121	0.3810	0.5723	0.3044	0.4701	0.4701	0.3146
A_8								0	0.3197	0.4000	0.3197	0.3963	0.3197	0.3197	0.3963	0.5488	0.2952	0.2214	0.2214	0.3478
A_9									0	0.3476	0	0.1097	0	0	0.1097	0.5697	0.0816	0.3048	0.3048	0.0357
A_{10}										0	0.3476	0.4024	0.3476	0.3476	0.4024	0.5440	0.3310	0.4571	0.4571	0.3748
A_{11}											0	0.1097	0	0	0.1097	0.5697	0.0816	0.3048	0.3048	0.0357
A_{12}												0	0.1097	0.1097	0	0.5607	0.1786	0.3833	0.3833	0.0816
A_{13}													0	0	0.1096	0.5697	0.0816	0.3048	0.3048	0.0357
A_{14}														0	0.1096	0.5697	0.0816	0.3048	0.3048	0.0357
A_{15}															0	0.5607	0.1786	0.3833	0.3833	0.0816
A_{16}																0	0.5488	0.5440	0.5440	0.5723
A_{17}																	0	0.3143	0.3143	0.1097
A_{18}																		0	0	0.3388
A_{19}																			0	0.3388
A_{20}																				0

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 7, July 2013

III.B.2. Calculation of distances:

To compare the similarity indices, we compare the above-obtained partitions using the distance Marczewski – Steinhaus[1].

For example, for the first two dendrograms of Fig. 1, the sets of intermediate steps for classification E_{A_1} and E_{A_2} are given. They are directly involved in the calculation of the distance between A_1 and A_2 as described in the example in paragraph II

$|E_{A_1}|=|E_{A_2}|=|X|$-1=7.

$$E_{A_1} = \begin{Bmatrix} \{3,8\},\{1,6\},\{3,8,5\},\{1,6,2\}, \\ \{3,8,5,1,6,2\},\{3,8,5,1,6,2,7\}, \\ \{3,8,5,1,6,2,7,4\} \end{Bmatrix}$$

$$E_{A_2} = \{\{3,4\},\{1,6\},\{3,4,5\},\{1,6,8\},\{3,4,5,7\}, \{3,4,5,7,1,6,8\},\{3,4,5,7,1,6,8,2\}\}$$

$d(A_1, A_2)$=0.3810

The calculated distances between the different trees, are given in Table III.

III.B.3 Operating:

From the distance matrix and applying the algorithm of hierarchical classification analysis, a dendrogram was obtained in fig.7. By making cuts in this latter according to the fixed objectives, we obtain a number of families of indices.

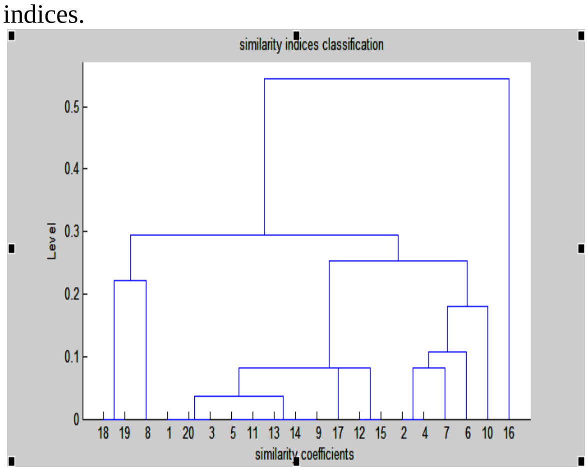

Fig.7. Similarity indices classification

To find family of indices, it is necessary to cut the classification tree at an appropriate level. We will choose $0.221 < \alpha < 0.252$.

For this level, we obtain the four families of indices as shown in fig.8.

$$\begin{cases} \text{Famille} F_1 : F_1 = \{18,19,8\} \\ \text{Famille } F_2 : F_2 = \{1,20,3,5,11,13,14,9,17,12,15\} \\ \text{Famille } F_3 : F_3 = \{2,4,7,6,10\} \\ \text{Famille } F_4 : F_4 = \{16\} \end{cases}$$

Fig.8. groups of coherent similarity indices

We deduce that we found the same results that are in the article [16] and we can say that the family F2 contains the most powerful indices as Jaccard, Sorenson, Kulczynski and Sokal and Sneath 2. However, the F3 family contains inefficient indices, namely: Hamann, Simple matching, Rogers and Tanimoto.

IV. Conclusion

In this paper, we presented a universal approach that aims to provide an efficient and effective tool for grouping different methods for classification. This approach gives the closest and the alternative families. We have presented its application to solve the general problem of classification ranges for the analysis of different logical ranges for selection of a product and compare different similarity coefficients methods. Finally, it is interesting to continue the application of this method of classification in different contexts in which the goal is to generalize classification method comparison for any partitioned data. Simply just have their hierarchical structures.

V. References

[1] Karonski M; and Palka Z.. ‘’On MarZeweski-Steinhaus Distance between Hypergraphs’’, Expositione Mathematicae,XVI, 1, 1977.
[2] KENDALL M.G, STUART A. The Advanced Theory of Statistics, Griffin, Londre, vol. 2, 1961
[3] MARCOTORCHINO F., Utilisation des Comparaisons par Paires en Statistique des Contingences (Partie II), Etude du Centre Scientifique IBM France, No F069, 1984.
[4] HUBERT L., ARABIE P., Comparing Partitions, Journal of Classification, vol. 2, 193-198, 1985.
[5] CHAVENT, M., ET AL., Critère de Rand Asymétrique, in Proceedings SFC 2001, 8èmes rencontres de la Société Francophone de Classification, Pointe à Pitre, 2001.
[6] ROBERT P., ESCOUFIER, Y. A Unifying Tool for Linear Multivariate Statistical Methods: the RV-coefficient. Appl. Statist., vol. 25, 257-265, 1976.
[7] JANSON S., VEGELIUS J. The J-index as a Measure of Association for Nominal Scale Response Agreement. Applied psychological measurement, vol. 16, 243-250, 1982.
[8] COHEN J., A Coefficient of Agreement for Nominal Scales., Educ. Psychol. Meas., vol. 20, 27-46, 1960.
[9] STEWART D., LOVE W., A General Canonical Correlation index, Psychological Bulletin, vol. 70, 160- 163, 1968.
[10] POPPING, R. Traces of agreement. On the Dot- Product As a Coefficient of Agreement. Quality and Quantity, vol. 17, No1, 1-18, 1983
[11] *Sur des indices de comparaison de deux classifications.* Saporta, G. Youness et G. 2003, Xèmes rencontres de la Société francophone de classification, pp. 169-172.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 7, July 2013

[12] Contribution aux méthodes de classification non supervisée via des approches prétopologiques et d'agrégation d'opinions. Boubou, M. 2007, Thèse de doctorat de l'université Claude Bernard .

[13] *Objective Criteria for the Evaluation of Clustering Methods.* Rand, W. M. 336, 1971, Journal of the American Statistical Association, Vol. 66, pp. 846-850.

[14] *Contributions à une méthodologie de comparaison de partitions.* YOUNESS, G. 2004, Thèse de doctorat de l'université Paris 6.

[15] *Critère de Rand Asymétrique.* CHAVENT, M., ET AL. in Proceedings SFC 2001.

[16] Similarity coefficient methods applied to the cell formation problem: a comparative investigation. Yasuda, Y. Yin et K. 48 , 2005, Computers & Industrial Engineering , pp. 471–489.

[17] *Evolving Similarity Coefficients for the Solution of Cellular Manufacturing Problems.* Mort, C. Dimopoulos and N. 2000. Proceedings of the 2000 Congress on Evolutionary Computation. Vol. 1, pp. 617 - 624 .

[18] Boujault A., Favrel J., Baptiste P. et al. :’’Ingenieurie des systemes flexibles d’assemblage’’, Compte-rendu de fin de recherche MRT sur l’ingenierie des systemes flexibles d’assemblage, institut Productique Besancon, Fevrier 1991.

[19] *A close neighbouralgorithm for designing cellular manufacturing systems.* Cheng, W.J. Boe and C. H. 10, 1991, INT. J. PROD. RES., Vol. 29, pp. 2097 - 2116.

[20] Boujault A.’’Contribution a une approche methodologique de l’assemblage automatise: elaboration automatique des sequences operatoires’’, these d’etat, Universite Franche-Comite,1984.

[21] Sedqui A., Martinez M., Campagne J.P. et Favrel J." Analyse des gammes de production en vue de la conception des cellules flexibles d'assemblage", 5ème Congrès de Génie Industriel GI5-IPI 95, 2-4 Avril 1996, Grenoble France