Building Transitive Groups in Computer-Supported Collaborative Learning Using Fuzzy Clustering

[(1)]Jose C. Romero Cortés, [(2)]Gustavo Núñez Esquer, [(1)]Arturo Aguilar Vázquez
[(1)]Departamento de Sistemas, Universidad Autonoma Metropolitana
[(2)]Centro de Investigacion en Computacion, Instituto Politecnico Nacional
[(1)]Avenida San Pablo 180, Colonia Reynosa Tamaulipas, Azcapotzalco, D.F.
MEXICO

Abstract: - In cluster analysis context is of major interest to induce an order that allow us to compare the groups generated by cluster analysis itself. Using the sample similarities obtained in this analysis like possibilities in fuzzy sets context provide us to induce a transitivity property in cluster analysis, producing a similarities matrix that meets to the transitivity conditions and that resembles as much as it can to the observed similarities matrix. To find out the transitive similarities matrix it is equivalent to find a solution to a mathematical programming problem where the objective function to be minimized is equal to the absolute difference between the similarities matrix and the unknown transitivity matrix entries. By contrast to non-linear programming analytical or heuristical approaches, which entail tedious and/or intricate calculations, the linear programming model we present in this paper is an easy-computing one that, additionally, provide us the explanation of the associated dual problem, something hardly attainable with other approaches. We include an application relevant in collaborative learning study and the results obtained are commented.

K*ey-words*:- Cluster analysis, Transitivity property, Similarity matrix, Transitivity matrix, Collaborative learning, Non linear programming

1 Introduction

Nowadays there exist an important numerical analysis development on cluster analysis, in the same sense the software on this field is abundant, however when we face on problems that involve systems where the application of multi-scaling techniques are required, specifically we talk about opinion studies, there are no means of inducing an order among the clusters estimated, such an order it is necessary because one of the main goals we search for in this kind of studies is to find the "hidden" ordered opinion structure, assuming that this really exists; in these cases the underlying structure of data lacks of probabilistic nature or it is not even approximately stochastic. Then it seems adequate to search for non-probabilistic mathematical models that provide us to analyze this kind of human data structures. Fuzzy sets theory provide us of such non-probabilistic model which appears like a transition and natural way of ordering the preferences among the several categories formed. Some research had been done considering the fuzziness to induce transitivity in cluster analysis using heuristical approaches and/or non-linear mathematical programming, with the computational and algorithmic complexity involved. Our aim in this paper is to use linear programming techniques, in fuzzy sets context, to induce a transitivity property in a

cluster analysis study; these techniques has well-known properties as well as powerful computational algorithms, although the model used here can be applied to similar scenario. In part I we give some basic concepts to handle fuzzy information that arise in this kind of studies, in part II we consider the linear programming mathematical model that provide us to analyze, that is to recognize and classify, opinion studies patterns; that is, we sketch the procedure to find the optimal solution to the linear programming mathematical model, as well as the optimal solution of the dual associated problem. In part III we include a small-scale sampling application example to be analyzed in detail. Finally, we discuss out the results obtained as well as some of its possible implications.

2 Problem formulation

Usually data of dichotomous (binary) relations provided by statistical survey questionnaires does not satisfy transitivity conditions, to see this let us consider three preference choices: C_1 , C_2 , and C_3 ; and assuming that:

$C_i > C_j \Rightarrow$ choice C_i is preferred to choice C_j ...(1)

yet we can observe that:

$C_1 > C_2$; $C_2 > C_3$; $C_3 > C_1$ ...(2)

So (1) is not satisfy by the result shown in (2), this example of non-transitivity show us that observed data structure could prevent us to make a decision about the order of choice alternatives. Hence we need a mathematical tool to be used as a transitive relation, in order to get that tool we are going to review some basic concepts of fuzzy sets theory [2] [3]. The concepts and operations associated with fuzzy sets theory are mostly direct extensions of respective concepts from ordinary (crisp) set theory, in those cases it can be proved that crisp set theory reduces to a particular case of fuzzy set theory.

Definition 1. Fuzzy subset

Let Ω be a universal set, a fuzzy subset A of Ω is one subset that has associated a membership function μ_A defined as:

$\mu_A:\Omega\rightarrow[0,1]$

The membership function $\mu_A(\omega)$ denotes the strength to which ω is an element of the set A .

When Ω is the real line, in many cases, the membership function adopts two important forms: triangular and trapezoidal.

Example . In the context of Computer Supported Collaborative Learning (CSCL), let $\Omega = \{\omega_1, \omega_2 , \omega_3 , \omega_4 , \omega_5 \}$ be the learners set and $A_i = \{\mu_{i1}/\omega_1 , \mu_{i2} / \omega_2 , \mu_{i3} / \omega_3 , \mu_{i4} / \omega_4 , \mu_{i5} / \omega_5 \}$ the fuzzy set such as the learner i perceives the other learners, for any variable.

Example . Let us suppose that the vector $A_1 = \{1.0/\omega_1, 0.4/\omega_2, 0.2/\omega_3, 0.5/ \omega_4, 0.6/\omega_5 \}$ represents the perceptions that learner 1 has of each learner, referred to message clearly variable. The learner2 belong to A_1 with a intensity of 0.4. We can build the fuzzy set for any variable in the analysis, all variables will be studied in the section IV.

Definition 2. Enclosure of fuzzy sets

Let A_1 and A_2 be two fuzzy subsets of Ω, then we say that A_1 it is contained in A_2 , denoted as, $A_1 \subset A_2$,when: $A_2(\omega) \geq A_1(\omega)$, for each $\omega \in \Omega$.

Definition 3. Equality of fuzzy subsets

Let A_1 and A_2 be two fuzzy subsets of Ω, then we say that A_1 is equal to A_2, $A_1 = A_2$,whenever $A_1 \subset A_2$ and $A_2 \subset A_1$.

Definition 4. Union of fuzzy subsets

Let A_1 and A_2 be two fuzzy subsets of Ω, then the union of A_1 and A_2 , denoted as $A_1 \cup A_2$, is defined as:

$A_1 \cup A_2(\omega) = \text{Max}\{A_1(\omega), A_2(\omega)\} = V\{A_1(\omega), A_2(\omega)\}$, for each $\omega \in \Omega$.

Definition 5. Intersection of fuzzy sets

Let A_1 and A_2 be two fuzzy subsets of Ω, then the intersection of A_1 and A_2 , denoted as $A_1 \cap A_2$, is given by: $A_1 \cap A_2(\omega) = \text{Min}\{A_1(\omega) , A_2(\omega)\} = \Lambda\{A_1(\omega), A_2(\omega)\}$, for each $\omega \in \Omega$.

Definition 6. Fuzzy relative complement

Let A_1 and A_2 be two fuzzy subsets of Ω, then the relative complement of A_2 with respect to A_1, which it is denoted as $A_1 - A_2$, is defined by:

$A_1 - A_2(\omega) = V\{0, [A_1(\omega) - A_2(\omega)]\}$, for each $\omega \in \Omega$.

Definition 7. Fuzzy relationship

Let Ω_1 and Ω_2 be two universal crisp sets, then a fuzzy relationship R over Ω_1 , Ω_2 is a fuzzy subset of the Cartesian product $\Omega_1 \times \Omega_2$.

When $\Omega_1 = \Omega_2 = \Omega$, we say that R is a fuzzy relation on Ω. As it was mentioned before, the classical concept of ordinary (crisp) set relationship is a particular case of a fuzzy relationship when membership function values are constrained to be one or zero.

Definition 8. Fuzzy composition

Let R_1 and R_2 be two fuzzy relationships on Ω, then we define the fuzzy composition of R_1 and R_2 , denoted as $R_1 \circ R_2$, as given by:

$R_1 \circ R_2(i,j) = V_k\{R_1(i,k) \Lambda R_2(k,j)\}$, for each i,j in Ω .

Definition 9. Fuzzy identity relationship

We say that R defines a fuzzy identity relationship, denoted by I, when: R (i,j) is defined as one (1) if i = j , and zero (0) otherwise. It is clear from this definition that: $R \circ I = I \circ R = R$, for any fuzzy relationship R.

Definition 10. Reflexive, Symmetric and Transitive Relationships [1]

Let R be any relationship, then we say that R is:

(i) Reflexive, if and only if $R \geq I$

(ii)Symmetric, if and only if R (i,j) = R (j,i)

(iii)Transitive, if and only if $R(i,j) \geq R \circ R(i,j)$

(iv)Anti-symmetric, if and only if $R(i,j) = 0 \Leftrightarrow R(j,i) \neq 0$

When a fuzzy relationship satisfies (i), (ii), and (iii) we say that it defines a fuzzy equivalence relationship, and if satisfies (i), (iii) and (iv) then we say that it defines a fuzzy semi-ordering relationship.

In clustering and ordering problems transitive relationships are very important: to make a decision about equivalent classes of objects pertain to clustering problems, and to make a decision about ordering relationships into several choices of alternatives that pertain to ordering problems, in this last case we must define fuzzy semi-ordering relationships.

In this point, R(i,j) represents the similarity between learners i,j in relation to their performance in the work group, for example R(2,3)= 0.18 indicates the learners 2 and 3 share marginally in the CSCL.

3 Problem solution

3.1 A Linear Programming Model To Induce Transitivity

Let us suppose that raw data has not a transitive structure as it is usual in opinion studies context, then given the data matrix we face the problem of finding a fuzzy transitive matrix T, from the constructed similarities matrix S, such that the sum of their squared observed differences be equal to a minimum; this is equivalent to require that the difference between S and T would be the minimum quantity available, mathematically this implies that a non-linear programming is formulated [1] in such a way that the objective function Z is

$$Min\ z = \sum_{i=1}^{L}\sum_{j=1}^{L}(S_{ij} - T_{ij})^2 \qquad \cdots(3)$$

defined as:

where the entries of matrix T are the non-negative fuzzy decision variables to be determined and must satisfy the transitive condition (iii), which is equivalent to:

$$T_{ij} \geq T_{ik} \quad and T_{ij} \geq T_{kj,} with \quad T_{ij} \geq 0 \qquad ..(4)$$

or equivalently :

$$T_{ij} \geq \bigcup_{k} \left(\bigcap_{i,j} (T_{ik}, T_{kj}) \right) \quad , with \quad T_{ij} \geq 0, \forall i, j$$

This quadratic non-linear programming has been solved previously , and it is equivalent to require that absolute difference between the correspondent entries of matrices S and T be equal to the minimum possible value. So this requirement implies that (3) can be defined as:

$$Min\ z = \sum_{i} \sum_{j} \left| (S_{ij} - T_{ij}) \right| \qquad \cdots(5)$$

subjected to the set of constraints given by (4) and although:

$$S_{ij} - T_{ij} = g_{ij} - f_{ij}, \quad with \quad i, j = 1,..., L \ ...(6)$$

The solution of the latter model is obtained by solving the next linear programming model [4][8]:

$$Min\ z = \sum_{i} \sum_{j} (g_{ij} + f_{ij})$$

Subject to :

$$T_{ij} \geq T_{ik}, \quad i, j, k = 1,..., L \qquad \cdots(7)$$
$$T_{ij} \geq T_{kj}, \quad i, j, k = 1,..., L$$
$$S_{ij} - T_{ij} = g_{ij} - f_{ij}$$
$$with \quad T_{ij} \geq 0, \quad g_{ij} \geq 0, \quad f_{ij} \geq 0$$

The variables g_{ij} and f_{ij} defined in (7) is an artifice to avoid using absolute values in (5), where:

$$(S_{ij} - T_{ij}) = g_{ij} \ for \ (S_{ij} - T_{ij}) > 0 \quad and$$
$$(S_{ij} - T_{ij}) = f_{ij} \ for \ (S_{ij} - T_{ij}) < 0$$

as it is the case when we have an unconstrained variable. Hence, any standard linear programming software can be used to find the optimal solution of (7), that is, the entries T_{ij} of the transitive matrix T that has the most resemblance to the respective entries S_{ij} of the observed similarities matrix S.The dual associated problem to (7) is given by:

$$Max \quad y = S_{11} y_1 + ... + S_{LL} y_{L^2}$$

Subject to :

$$A' y' \leq 0$$
$$Ay'' \leq gf \qquad ...(8)$$

$$with \quad y' \in R_{+}^{L}, \ y'' \in R^{L}$$

Where :

$$A' = \begin{pmatrix} 1-100...00...00...00 \\ 10-10...00...00...00 \\ \vdots \\ 1000...0-1...00...00 \\ \vdots \\ 0000...00...1-1...00 \\ 0000...00...10...-10 \end{pmatrix}$$

$$A'' = \begin{pmatrix} 11-1.....0 \\ 01-1.....0 \\ \vdots \\ 000.....11-1 \end{pmatrix}$$

$$gf = \begin{pmatrix} g_{11} + f_{11} \\ \vdots \\ g_{LL} + f_{LL} \end{pmatrix}$$

Where dual variables y'' define an order of preference within each cluster involved; so we can induce an order not just among the several clusters formed given by means of the optimal solution of (7), but in addition to into each one of them, given by means of the optimal solution of (8).

3.2 Building Groups in CSCL

The potential of fuzzy clustering is illustrated generating work groups with similar performance in CSCL. It is important to mention that the approach used induces an order between clusters and the lecturers in each cluster are liked , the preferences opinion of the learners by pairs will be noted in the order of the clusters. The first cluster will contain the group of learners with the best grades for all the variables considered, following the second cluster and so on. The measures used in the analysis can be similarities, distances, correlations etc. The CSCL has many application dominions as learning, training and knowledge discovery in systems as education, business, government , science and engineering. The interaction between learners is a central point in CSCL, it is measured considering the opinion by pairs of learners for the next variables [6]:

1) Importance of the proposals
2) Justify of arguments
3) Relation with the themes
4) Clearly in the messages
5) Contributions
6) Originality in contributions

The nature of these variables is fuzzy, any of these can be measure in terms of categories. For example, in relation to variable 4), the answer could be: no clearly, some clearly and clearly, each learner will be qualify in any of these categories with one grade of membership.

Each learner will qualify the others learners in scale of 1 to L, where L is the size group, and the grade L indicates the most preferred learner, and the answer 1 indicates the lowest one. Matrix similarities S is created by using the raw data set obtained, the corresponding entries S_{ij} of matrix S are the respective parameters of the linear program formulated in (7) where the decision variables T_{ij} to be determined when the linear program optimal solution is reached. S_{ij} can be calculated using many methods as correlations, distances , absolute deviations etc., we use the following expression:

$$S_{ij} = 1 - \sum_{K=1}^{m} \left| x_{ijk} - x_{jik} \right| / (Lm), \quad i, j = 1,..., L$$

where m is the number of variables ...(9)

and x_{ijk} *is the* grade *from learner i to learner j in the* variable *k*.

We are searching for a transitive matrix T obtained from the non-transitive data similarities matrix S, such that we can use that matrix T as a transitive fuzzy relation. Though some authors have had use heuristic methods to induce transitivity in hierarchical clustering, Watada et al (1982), in this paper we found out the elements of transitive matrix T directly, providing meet comparative aims in similar studies avoiding, as given in (3), quadratic errors considerations. When we apply this linear program approach to grades data obtained in the study, we get the matrix T, through the optimal solution of (7) and we include the ordered clusters recognized according to specific similarity levels, as well as the order induced into each one of the clusters obtained when the optimal solution to (8) is attained. The aim of this application is to generate ordered groups from five learners, considering the grades by pairs for the six variables mentioned. The grades observations are sketched in Table 1, for variable 2, each entry is the average by pair of learners.

Learner	L1	L2	L3	L4	L5
L1	5.0	2.0	1.0	2.5	3.0
L2	2.0	5.0	1.2	2.7	2.5
L3	1.0	1.2	5.0	1.2	2.8
L4	2.5	2.7	1.2	5.0	1.8
L5	3.0	2.5	2.8	1.8	5.0

Table1. Grades observations for variable 2

The non-transitive similarities matrix S is created from the observed data shown on Table 1, by using (9), and it is shown on Table 2:

Learner	L1	L2	L3	L4	L5
L1	1	.312	.776	.360	.801
L2	.312	1	.180	.600	.501
L3	.776	.180	1	.248	.633
L4	.360	.600	.248	1	.545
L5	.801	.501	.633	.545	1

Table2. Non-transitive similarity matrix

According to (7) the linear programming model to be solved in this case is:

$$
\begin{aligned}
&Min \quad Z = \sum_{i=1}^{5}\sum_{j=1}^{5}(g_{ij} + f_{ij}) \\
&subject\ to: \qquad\qquad (10) \\
&T(i,j) \geq T(i,k), \quad i,j,k = 1,...,5 \\
&T(i,j) \geq T(k,j), \quad i,j,k = 1,...,5 \\
&S_{ij} - T_{ij} = g_{ij} - f_{ij} \\
&with \quad T(i,j) \geq 0,\ g_{ij} \geq 0,\ f_{ij} \geq 0
\end{aligned}
$$

The optimal solution of (10), applied to Table 2 leads to the transitive similarity matrix T which it is shown on the Table 3:

Learner	L1	L2	L3	L4	L5
L1	1	.479	.704	.479	.801
L2	.479	1	.479	.600	.479
L3	.704	.479	1	.479	.704
L4	.479	.600	.479	1	.479
L5	.801	.479	.704	.479	1

Table 3. Transitive similarity Matrix

From this transitive similarity matrix we can infer an ordered cluster set with respect to some relevant similarity levels, as is indicated in Table 4:

Similarity Level	**Cluster**
.8010	(L1,L5)
1.0000	(L3)
.6000	(L2,L4)

Table 4. Ordered cluster set

The first cluster contains the learners with the best grades and they have performance liked in the group, the others present these attributes with values lesser one before, and so on. This result is also shown in the following dendogram obtained by SAS, the average linkage distance between two groups is the average distance between pair of observations, one in each group, this method is debt to Sokal and Michener [7]:

Average Linkage Cluster Analysis
Average Distance Between Clusters

```
          1.2  1.1  1.0  0.9  0.8  0.7  0.6  0.5
Learner---+----------------------------------+-
1         xxxxxxxxxxxxxxxxxxxxxxxxxxxxx
          xxxxxxxxxxxxxxxxxxxxxxxxxxxxx
5         xxxxxxxxxxxxxxxxxxxxxxxxxxxxx
          xxxxxxxxxxxxxxxxxxx
          xxxxxxxxxxxxxxxxxxx
3         xxxxxxxxxxxxxxxxxxx
          x
          x
2         xxxxxxxxx
          xxxxxxxxx
4         xxxxxxxxx
```

4 Conclusions

The paper only presents some variables of the analysis and the results are partial, however we are working in a complete set of variables with the purpose to evaluate the system in a summary form. Usually this

evaluation is assigning weights to the performance of the learners in the different subsystems, however these are arbitrary, but with our approach of fuzzy clustering is not necessary fixed any weight. In particular, we present the potential of this approach building work groups in CSCL. Perhaps the real world complexity is even most drastically reflected in the so called *soft areas* such as social, economic and like, where their inherent feedback and/or hierarchical structures and/or dynamic evolution characteristics are taking into account to get a better model to represent them. We hope this paper contributes to elucidate the dynamic nature of opinion patterns structures, another real system where its complex nature drive us to develop out complex systems approach to examine it closely.

References:

1 Watada, J; Tanaka, H., and Asai K. *.A heuristic method of hierarchical clustering for fuzzy intransitive relations*, In, *Fuzzy Set and Possibility Theory,* Yager, R.R., Editor, New York, Oxford, Toronto, 1982

2 Yager, R.R., and Filev, D., *Essentials of Fuzzy Modeling and Control,* Wiley, 1994

3 Zimmermann, H.J., *Fuzzy Set Theory - and Its Applications,* Second, Revised Edition. Kluwer Academic Publishers, 1991.

4 Fisher,W.D. *A note on curve fitting with minimum deviations by linear programming*, JASA, Vol. 56, 1962, 359-262.

5 Jang,J.S.R.;Sun,C.T.,Mizutani,E., *Neuro Fuzzy and Soft Computing: A Combinatorial Approach to Learning and Machine Intelligence*. Prentice-Hall, 1997.

6 Rodríguez,G.J.L., *Modelo de trabajo grupal y evaluación en aprendizaje colaborativo personalizado asistido por computadora.* IPN.CIC., 1999

7 SAS (1997).*Statistiscal Analysis System. Version 6.0*

8 Aguilar V.A., Romero C.J.C. *A model for inducing transitivity in hierarchical clustering: an application to a consumer preference study.* Joint Conference IASS/IAOS-INEGI. 1998