

December 2004

A Clustering Algorithm in Group Decision Making

Xuan-hua Xu
Central South University

Xiao-hong Chen
Central South University

Ding Luo
Central South University

Follow this and additional works at: <http://aisel.aisnet.org/pacis2004>

Recommended Citation

Xu, Xuan-hua; Chen, Xiao-hong; and Luo, Ding, "A Clustering Algorithm in Group Decision Making" (2004). *PACIS 2004 Proceedings*. 68.
<http://aisel.aisnet.org/pacis2004/68>

This material is brought to you by the Pacific Asia Conference on Information Systems (PACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in PACIS 2004 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

A Clustering Algorithm in Group Decision Making[♦]

XU Xuan-hua¹, CHEN Xiao-hong², LUO Ding³

(School of Business, Central South University,
Changsha 410083, Hunan, P.R.C,xuxh@public.cs.hn.cn)

Abstract

The homogeneous requirement of the AHP has stifled its general application and thus hindered the further development of group decision making, which is becoming increasingly popular with the multi-business corporations. In this paper, we adopted the correlation degree of individual preference vectors as the measurement of individuals similarity based on AHP vector space VAHP proposed by Zahir in literature (Saaty 1989). The group in charge of decision making can thus be divided into multiple homogeneous clusters using this measurement, based on whose result the AHP can be applied. A heuristic algorithm is put forward at the end of the paper to facilitate the clustering of the group and the attainment of the group preference, both of which are vital to solve the problem of individual clustering in the group decision making process.

Keyword: VAHP; Group decision making; Group Cluster, Group Coherence

1. Introduction

Along with the development of knowledge economy and network technology, decision process increasingly take the form of group decision making, which can be effectively supported by the network-based group decision making support system[1-2]. The group decision making method becomes one of key problems.

It is well known that Analytic Hierarchy Process(AHP)(Saaty 1990) is usually used in individual decision making or group decision making within small homogeneous groups, which is also one precondition for such applications(Saaty 1989). For instance, when the group is small as in corporate board rooms, the group members sometimes may arrive at a consensus consciousness, and both arithmetic average and geometry average methods can be used in synthesizing individual decisions into a group decision making(Foreman 1996). However, in an intermediate or large-sized group, this homogeneity of individuals within the group cannot be achieved. Such a group of individuals is not expected to form one cohesive collection of similar decisions thinking. But they may form some subgroups (or clusters) while the members of each cluster have homogeneous views. Therefore, in a cognitive sense, the group is a set of clusters. Searching for homogeneous clusters is more important. Zahir extends the conventional AHP formulations to a Euclidean vector space(VAHP)(Zahir 1997) and applies it to both the individual and group decision making process in a homogeneous group.

(70125002)

[♦] **Fund project:** National Science Fund for Distinguished Young Scholars of National Science Foundation of China¹ 1 XU Xuan-hua : School of Business, CSU, Vice-Professor, P.H.D

² Chen XiaoHong : Dean of School of Business, CSU, professor, P.H.D

³ LUO Ding: School of Business, CSU, Management and Engineering major graduate student

Basak and Saaty use statistical methods to divide individuals into homogeneous clusters (Basak 1993), while Sajjad Zahir use an algorithm based on scalar product to form clusters and divide individuals into homogeneous clusters (Zahir 1999). Within the framework of the VAHP, this paper adopts the correlation degree of individual preference vectors as a measure of individual similarity and therefore divides individuals into homogeneous clusters. A heuristic algorithm is generated based on the measurement to produce category sets in the group, where the coherence of each category and that of the whole group set is calculated. The preference of the group is then attained.

2. The similarity degree of individuals in group

Definition 1: Marks group as Ω , in which there are N individuals. E_n is a N -Dimension Euclidean vector space. For the n th criteria, v_j^i ($v_j^i \geq 0$, $j=1,2,\dots,n$) stands for the values of the individual i under this criteria. Then, the value vector $V^i = (v_1^i, v_2^i, \dots, v_n^i)$ is the preference vector of the individual i in the group. Wherein $V_i \in E_n$.

We suppose that n_k is the number of individuals of the k th cluster C_k and there are η clusters in the group. Then, $\sum_{k=1}^{\eta} n_k = N$, where η ($1 \leq \eta \leq N$) is the fragmentation factor of the group. A threshold γ ($1 \leq \gamma \leq N$) is introduced to mark the similarity degree of two preference vectors.

Definition 2: The similarity degree $r_{ij} = r(V_i, V_j)$ between two preference vectors is defined as

$$r(V^i, V^j) = \frac{(|V^i - \bar{V}^i|) \cdot (|V^j - \bar{V}^j|)^T}{\|V^i - \bar{V}^i\|_2 \cdot \|V^j - \bar{V}^j\|_2}.$$

Where the $\|\cdot\|_2$ is the 2-norm of the vector, $\bar{V}^i = \frac{1}{n} \sum_{l=1}^n v_l^i$, $\bar{V}^j = \frac{1}{n} \sum_{l=1}^n v_l^j$,

and we have the inequality $0 \leq r_{ij} \leq 1$.

The condition of $r(V^i, V^j) \geq \gamma$ is set for such situation, which means that the similarity degree between any two preference vectors is large or equal to the threshold γ . It, therefore, can also be viewed as a qualification parameter of an individual, which is used to determine whether the member should be contained in a certain cluster.

Definition 3: The preference of a cluster. For the k^{th} cluster C^k , the preference can be calculated by adding the preference vector of each member in this cluster, that is:

$$G^k = \sum_{V^i \in C^k} V^i.$$

We can get the unit vector \hat{G}^k by standardizing G^k using the following formula:

$$\hat{G}^k = \sum_{V^i \in C^k} V^i / \left\| \sum_{V^i \in C^k} V^i \right\|_2, \text{ where } \hat{G}_j^k = \sum_{V^i \in C^k} v_j^i / \left\| \sum_{V^i \in C^k} V^i \right\|_2, \quad j = 1, 2, \dots, n,$$

and $(\hat{G}^k)^T \cdot \hat{G}^k = 1$

Definition 4: The preference of a group. Through the weighted sum of the preferences of all the clusters in the group, the preference E of the group can be obtained as:

$$E = \sum_{k=1}^n \frac{n_k}{N} \hat{G}^k, \text{ where } E_j = \sum_{k=1}^n \frac{n_k}{N} \hat{G}_j^k, \quad j = 1, 2, \dots, n,$$

Unit vector \hat{E} , the preference of the group, can also be formed by standardizing E as:

$$\hat{E}^k = E^k / \|E^k\|_2, \text{ and } (\hat{E})^T \cdot \hat{E} = 1.$$

3. The Algorithm of individuals clustering

3.1 Clustering

For a formed (existing) cluster, select the preference vector of an individual. If the correlation degree of the linear combination between the vector and those of other individuals in the cluster is larger than threshold γ , then this individual can be distributed to the cluster. Otherwise, this vector will be distributed to one of the other clusters. When all the vectors in the group have been distributed to corresponding clusters, the algorithm stops. The threshold γ marks the minimum value of correlation degree within a cluster, rely on which whether an individual can be distributed to the cluster is decided.

3.2 Algorithm of clustering

Step1. All preference vectors are collected into set U. Set a temporary set T. Meanwhile, initialize the counter $k = 1$.

Step2. Randomly select vector V^i from U ($V^i \in U$), move it from set U to cluster C^k , and initialize this counter $n_k=1$ of the individuals of the cluster.

Step3. The linear combination is carried out for all the vectors in C^k , we can obtain the β marked as following.

$$\beta = \sum_{l=1}^{n_k} \alpha_l \cdot V^l, \text{ where } \alpha_1, \alpha_2, \dots, \alpha_{n_k} \text{ are non- negative real coefficients.}$$

Step4. If U is not empty, randomly select another vector V^j ($V^j \in U$). Otherwise, go to the step6.

Step5. Calculate $r_j = r(\beta, V^j) = \frac{(|\beta - \bar{\beta}|) \cdot (|V^j - \bar{V}^j|)^T}{\|\beta - \bar{\beta}\|_2 \cdot \|V^j - \bar{V}^j\|_2}$

If $r_j \geq \gamma$, then distribute V^j to cluster C^k . At the same time, move V^j from U to C^k and $n_k = n_k + 1$. Append T to U, that is $U = U \cup T$ and set $T = \Phi$ (empty set). Otherwise, move V^j from U to T and go to step3.

Step6. If T is not empty, set $U = T$, $T = \Phi$, $k = k + 1$, and go to step two. Otherwise, go to step seven.

Step7. Record results.

$k =$ the number of clusters, $n_i =$ the number of individuals of cluster i , where $i = 1, 2, \dots, k$; C^i is the cluster i .

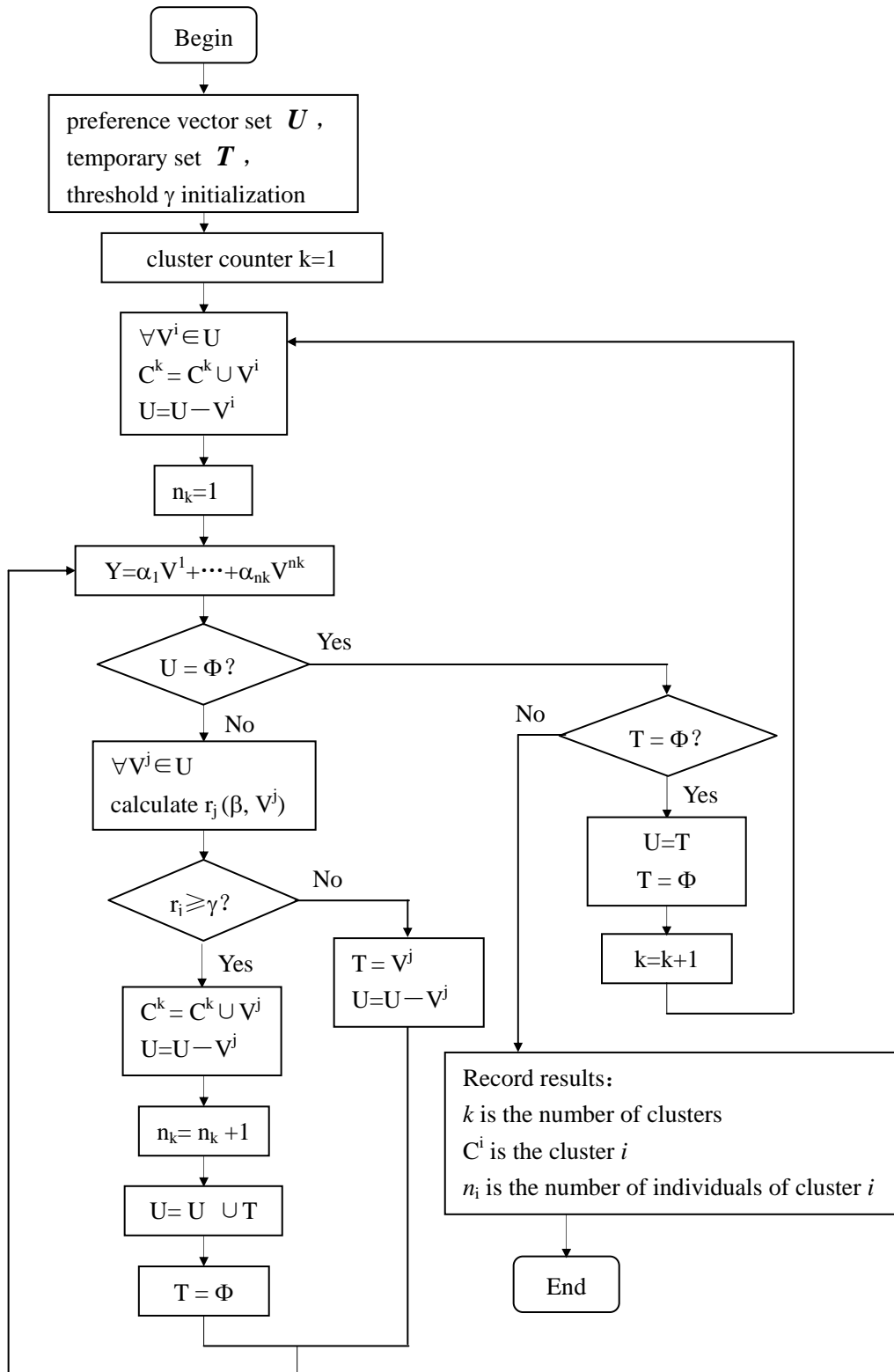
It is obvious that $\sum_{i=1}^k n_i = N$, k is a measure of the group component.

If $k = 1$, then there only has one cluster in the group. Then it is a homogeneous group.

The value of the threshold γ is very important, because whether a certain decision member i can enter into a certain cluster C^k , depend on whether the similarity degree r_{ik} between the preference vector V^i of member i and the linear combination V^k of the preference vectors of all the members in the cluster C^k is larger than or equal to the

threshold γ , that is whether $r_{ik} \geq \gamma$. If this inequality is right, the member i can enter into the cluster C^k , that is the member i is close to preference of all the members in cluster C^k and the close degree is larger than or equal to this threshold γ . Otherwise, the member i can't enter into the cluster C^k . The more the threshold γ is large, the more the member i is difficult to enter into the cluster C^k . Contrarily, the member i is more easy to enter into the cluster C^k . The value of the threshold γ can be manually taken a greater number at first, then, the algorithm is carried out. If all the members in the Ω can't completely enter into all the clusters formed, to reduce the value of the threshold value γ proportionately and to carry out the algorithm continuously until all the members in the Ω have completely enter into all the cluster formed.

Its process picture of block diagram as follows.



4. The coherence of group

4.1 The coherence of cluster

Definition 5: The coherence of cluster. The coherence ρ^k of the k^{th} cluster C^k (in which there is at least one individual) is defined as:

$$\rho^k = \frac{1}{n_k} \sum_{\substack{i,j=1 \\ i \neq j}}^{n_k} r(V^i, V^j), \text{ where } V^i, V^j \in C^k, (i, j = 1, \dots, n_k, i \neq j)$$

The coherence ρ^k of any a cluster is considered meaningful only when it has more than one member. Otherwise, it would have a value of zero.

4.2 The coherence of group

For a group, two types of coherences are defined: the general coherence and the representative coherence.

Definition 6: The general coherence of a group. It represents the consistency of all the individuals in the group and does not consider any structure of clusters. The general coherence of group is defined as following

$$\rho = \frac{1}{N} \sum_{\substack{i,j=1 \\ i \neq j}}^N r(V^i, V^j), (i, j = 1, \dots, N, i \neq j)$$

Definition 7: The representative coherence of group. It represents the consistency among the clusters. With \hat{G}^k represents the clustered preference vector of each cluster. Therefore, the representative coherence of group is defined as following.

$$\rho^R = \frac{1}{\eta} \sum_{\substack{i,j=1 \\ i \neq j}}^{\eta} (\hat{G}^i)^T \cdot \hat{G}^j$$

5. Example

We use the following data set composed of only three different vectors , each data set has the randomly multi-copy mixed. These three vectors are:

V_1	0.9659	0.7071	0.2588
V_2	0.2588	0.7071	0.9659
n	23	15	12
γ	0.25	0.5	0.75

Their threshold γ in the 2-dimensions vector space are 0.25,0.5 and 0.75. There respectively are 23,15 and 12 members in this data set. Therefore, we expect the algorithm to produce 3 clusters. We set the threshold $\gamma = 0.99$ and get the first output of cluster is:

Number of cluster = 1, the coherence of the cluster 1 is 1:

45	0.2588	0.9659	29	0.2588
3	0.2588	0.9659	32	0.2588
39	0.2588	0.9659	14	0.2588
4	0.2588	0.9659		
40	0.2588	0.9659		
25	0.2588	0.9659		
6	0.2588	0.9659		
48	0.2588	0.9659		
7	0.2588	0.9659		

6. Conclusions

The adoption of correlation degree preference vector enabled “intelligent”

distribution of individuals in a decision group into various clusters. An algorithm based on this idea was provided in our paper. Yet any appearance of new individual in a cluster would like change the attribute of the cluster as a whole, the threshold % for instance. Thus, if a once rejected individual becomes candidate in a later selection process, it may find its way into the cluster. Therefore, the further task will involve revision of the qualification of rejected individuals once the clusters are changed and the most efficient way of the second round distribution.

Reference

- Cheng Xiao-hong, "Research of Enterprise Management Decision Support System"
Journal of Central South Institute of Mining and Metallurgy 1993, Vol.24 (5) P711
- Cheng Xiao-hong, ZHU Yan-ju. A Layer Model Method—Based Group decision making Support System in Internet or Intranet Environment. Chinese Journal of Management Science. 2001. Vol.9(6): 49-57(in Chinese)
- Sajjad Zahir, Decision making in the vector space formulation of the analytic hierarchy process, European Journal of Operational Research 112(1999)620-634.
- Saaty, T.L., 1990. Multicriteria Decision Making: The Analysis Hierarchy Process. RWS Publications, Pittsburgh, PA, USA.
- Saaty, T.L., 1989. Group decision making making and the AHP. In: Golden, B.L., Wasil, E.A., Harker, P.T. (Eds.), The Analytic Hierarchy Process: Application and Studies. Springer, New York, pp. 59±67.
- Foreman, E.H., Peniwati, K., 1996. Aggregating individual judgements and priorities with the analytic hierarchy process. In: Proceedings of the Fourth International Symposium on the Analytic Hierarchy Process, Vancouver, Canada, 12±15 July, pp. 383±391.
- Zahir, S., 1997. Geometry of decision making and the vector space formulation of the AHP. Working Paper, University of Lethbridge, July.
- Basak, I., Saaty, T.L., 1993. Group decision making Making Using the Analytic Hierarchy Process. Mathematical & Computer Modelling 17, 101-109.