Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 162 (2019) 316–323

Procedia Computer Science

www.elsevier.com/locate/procedia

7th International Conference on Information Technology and Quantitative Management (ITQM 2019)

Using clustering methods to deal with high number of alternatives on Group Decision Making

J. A. Morente-Molinera*[a], S. Ríos Aguilar[b], R. González-Crespo[b], E. Herrera-Viedma[a]

[a]Andalusian Research Institute in Data Science and Computational Intelligence, University of Granada, jamoren@decsai.ugr.es,viedma@decsai.ugr.es, Granada, Spain.
[b]School of Engineering and Technology, Universidad Internacional de la Rioja (UNIR), Logroño, Spain.

Abstract

Novel Group Decision Making methods and Web 2.0 have augmented the quantity of data that experts have to discuss about. Nevertheless, experts are only capable of dealing with a reduced set of information. In this paper, a novel method for dealing with decision environments that include a large set of alternatives is presented. By the use of clustering methods, the available alternatives are combined into clusters according to their similarity. Afterwards, one Group Decision Making process is employed for choosing a cluster and another one for selecting the final alternative.

Keywords: Clustering methods, large-scale Group Decision Making, linguistic modelling ;

1. Introduction

Web 2.0 technologies have provoked a profound impact on how users see Internet today. This change mainly concerns the users which have now a main role in the providing and consuming information tasks related to the Internet. The Internet users' number has increased exponentially and so has the information available on it.

When applying these technologies to traditional Group Decision Making (GDM) methods [1, 2, 3], several challenges appear. One of them consists in finding new GDM methods capable of managing the large amount of data that some experts have to analyze. Although the information available is high, experts cannot really deal with too many information at the same time. Therefore, there is a need of develop new methods that allow experts to work in an organized manner [4]. It should be noticed that if the experts are forced to deal with too much information, then it is quite difficult for them to make reliable decisions.

Consequently, it is necessary to aid the experts by allowing them to work with a reduced amount of information. One way of achieving this task is by deleting alternatives from the main alternatives set. This way, experts

*Corresponding author. Tel.: +34958241773.
E-mail address: jamoren@decsai.ugr.es.

10.1016/j.procs.2019.11.290

will only have to deal with a reduced set of alternatives. Nevertheless, this approach implies that the experts do not directly decide which alternatives are discussed and which are discarded. In order to solve this, a novel GDM approach based on clustering methods [5, 6] is proposed. Experts will group the alternatives in different categories and a GDM process is carried out in order to select the most appropriate alternatives set.

This article proposes a new GDM method that performs two GDM processes. In the first one, the experts choose the set of alternatives that have the most desirable characteristics. In the second, a ranking over the alternatives of the chosen set is calculated. It should be noticed that the alternatives groups are created using similarity based on expert's suggestions. This way, the experts are the ones comparing the alternatives and establishing their similarity.

Paper sections are organized as follows. In subsection 2, basis required to comprehend the presented method are exposed. In subsection 3, the proposed method is thoroughly described. In subsection 4, an example in which the method is applied is shown. Finally, the paper ends with a conclusion section.

2. Preliminaries

This section expose all the technologies and procedures that were used in the development of the novel method. They are briefly introduced in the next subsections.

2.1. Basis of Group Decision Making

This field is a quite popular topic in recent research [7, 8, 9, 10]. It is possible to formally define a GDM problem as follows:

Let have sets of experts and alternatives, $E = \{e_1, \dots, e_n\}$ and $X = \{x_1, \dots, x_m\}$, respectively. A GDM problem consists on ranking the set of alternatives by using the set of preferences, $P = \{p^k_{ij}\}$, that are generated and provided to the system by the experts. p^k_{ij} indicates how much expert e_k prefers x_i over x_j.

In order to carry out a GDM process, the next steps can be followed:

- **Providing preferences**: Experts interact with the system in order to share their preferences values. They must indicate how much they prefer each alternative over the others. Multi-granular fuzzy linguistic modelling methods [11, 12] could be used in order to allow experts to provide the information using the mean that better fit them.
- **Obtaining the collective preference matrix**: All the preferences that are received by the system are aggregated into a collective preference matrix. This matrix represents the overall opinion of all of them.
- **Ranking alternatives according to expert's preferences**: Using the preference collective matrix, the available alternatives in the decision process are ranked.
- **Calculating consensus**: Searching for a consensual decision in Decision Making procedures is an important issue [13, 14]. If the experts agree on a specific solution to the problem, then, the chosen ranking is ratified by all of them. Therefore, GDM methods should encourage experts to go on a thorough debate in order to reach an agreement. For this purpose, it is possible to perform several GDM rounds. The process will end if the consensus degree is high enough or if too many time have passed and they cannot find a common solution.

3. A Group Decision Making method which handles a large set of alternatives

The designed GDM procedure for environments with a high number of alternatives converts a large-scale GDM process with a high number of alternatives into two GDM processes with a reduced set of alternatives. Thanks to this, experts can comfortably tackle both of them without getting lost among all available alternatives. The novel presented method has the following steps:

- **Creating the alternatives groups**: Each expert categorizes all the alternatives if different groups.
- **Carrying out the consensual grouping**: By using an hierarchical clustering method, a set of groups of alternatives is generated. All experts provided information is used in order to create these groups. Therefore, each final generated group is created by using the all the experts' judgments.
- **Selecting the most preferred group**: Experts debate about which alternatives' group has the most promising alternatives. A GDM process is carried out in order to determine which is the best one.
- **Selecting the most preferred alternative inside the chosen group**: Experts debate about the alternatives that are contained in the chosen group. A GDM method is used in order to rank them.

All this process are detailed exposed in the next subsections.

3.1. Creating the consensual grouping

In this step, alternatives are classified into several groups according to their similarity. First of all, experts have to provide information about how they would classify the alternatives into groups. They can build any number of groups and they can classify together any number of alternatives. Using that information, a co-occurrence matrix, CO, is built. Each position, co_{ij}, indicates how many times the experts have classified elements i and j together in the same group.

By using the generated co-occurrence matrix, hierarchical clustering methods [15, 16] are employed in order to generate different groups of alternatives. Experts can specify the number of groups of alternatives that they want to discuss about and that level of the generated classification tree is chosen.

3.2. Carrying out the two Group Decision Making processes

Once that the groups of alternatives are created, experts can carry out the decision process and select the best alternative going through two different GDM processes. In the first one, experts will decide which group of alternatives is the best one. Finally, alternatives of the selected group are discussed in order to select the best option. The GDM scheme exposed in subsection 2.1 will be followed. The following parameters are taken into account:

- **Preference providing means**: Experts use preference relations matrices in order to provide their preferences to the system. A preference relation matrix can be defined as a matrix, P where p^k_{ij} refers to the amount of preference of alternative x_i over alternative x_j that expert e_k has provided.
- **Aggregation operator**: In order to aggregate the information, the mean operator is used. Therefore, it is possible to aggregate the preferences provided by the experts by following the next expression:

$$c_{ij} = \sum_k w_k \cdot p^k_{ij} \tag{1}$$

  where w_k is the weight associated to expert e_k. It is important that the weights fulfill that $\sum w_i = 1$.
- **Ranking alternatives**: Selection operators [17, 18] are used in order to generate the final ranking of alternatives. The following expressions can be applied to calculate the non guided dominance and guided dominance degree. GNDD and GDD respectively:

$$GNDD_i = \phi(c^s_{1i}, c^s_{2i}, \ldots, c^s_{(i-1)i}, c^s_{(i+1)i}, \ldots, c^s_{ni}) \tag{2}$$

  taking into account that

$$c^s_{ji} = \max\{c_{ji} - c_{ij}, 1\}$$

$$GDD_i = \phi(c_{i1}, c_{i2}, \ldots, c_{i(i-1)}, c_{i(i+1)}, \ldots, c_{in}) \tag{3}$$

  In order to calculate the position in the alternatives ranking, RV, GDD and $GNDD$ operators results can be combined as follows:

$$RV_i = (GDD_i + GNDD_i)/2, \forall i \in [0, m] \tag{4}$$

- **Measuring consensus**: In order to measure consensus among the experts, the following expressions can be used [19]:
  - **Consensus between pair of alternatives,** cp: First of all, let calculate the similarity between two preference relation matrices:
  $$sm_{ij}^{lk} = s(p_i^{lk}, p_j^{lk}) = 1 - |p_i^{lk} - p_j^{lk}| \quad (5)$$
  where p_i^{lk} refers to the preference relation matrix value of e_i for alternatives x_l and x_k. By aggregating all these matrices, it is possible to calculate the global consensus matrix as follows:
  $$cm^{lk} = \phi(sm_{ij}^{lk})\ i, j = 1, \ldots, m;\ l, k = 1, \ldots, n;\ i < j \quad (6)$$
  where ϕ is the mean operator. Using cm, it is possible to calculate cp_{lk} for the alternatives x_l and x_k using the following expression:
  $$cp^{lk} = cm^{lk},\ \forall l, k = 1, \ldots, n;\ l \neq k \quad (7)$$
  - **Consensus on an specific alternative,** ca: By aggregating cp values of an specific alternative, it is possible to calculate the consensus reached on an specific alternative:
  $$ca^l = \frac{\sum_{k=l,k\neq l}^{n}(cp^{lk} + cp^{lk})}{2(n-1)} \quad (8)$$
  - **Global consensus value**: Finally, the global consensus value is calculated by aggregating all the ca values calculated on the previous step:
  $$cr = \frac{\sum_{l=1}^{n} ca^l}{n} \quad (9)$$

4. Example

Imagine that three experts, $E = \{e_1, e_2, e_3\}$, need to decide where to invest some funds. They have in total 14 different choices according to their advisers: $X = \{x_1, \cdots, x_{14}\}$. First of all, experts must create their own groups of similar alternatives. Obtained results can be seen on Tables 1, 2 and 3.

1	2	3	4	5
x_3	x_4	x_7	x_1	x_{11}
x_5	x_8	x_9	x_2	
x_6	x_{13}		x_{10}	
x_{12}	x_{14}			

Table 1. Classification in groups performed by e_1.

1	2	3	4
x_3	x_4	x_7	x_1
x_5	x_8	x_9	x_2
x_6	x_{11}		x_{10}
x_{12}	x_{14}		x_{13}

Table 2. Groups made by e_2.

Once that each expert has made his/her own groups, the co-occurrence matrix is calculated. This matrix represents the number of times that two alternatives have been classified together by the experts. The co-occurrence matrix generated can be seen in Table 4.

1	2	3	4	5
x_3	x_1	x_7	x_4	x_{11}
x_5	x_2	x_9	x_8	x_{14}
x_6	x_{13}	x_{12}	x_{10}	

Table 3. Groups made by e_3.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}
x_1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
x_2	3	-	-	-	-	-	-	-	-	-	-	-	-	-
x_3	0	0	-	-	-	-	-	-	-	-	-	-	-	-
x_4	0	0	0	-	-	-	-	-	-	-	-	-	-	-
x_5	0	0	3	0	-	-	-	-	-	-	-	-	-	-
x_6	0	0	3	0	3	-	-	-	-	-	-	-	-	-
x_7	0	0	0	0	0	0	-	-	-	-	-	-	-	-
x_8	0	0	0	3	0	0	0	-	-	-	-	-	-	-
x_9	0	0	0	0	0	0	3	0	-	-	-	-	-	-
x_{10}	2	2	0	1	0	0	0	1	0	-	-	-	-	-
x_{11}	0	0	0	1	0	0	0	1	0	0	-	-	-	-
x_{12}	0	0	2	0	2	2	1	0	1	0	0	-	-	-
x_{13}	2	2	0	1	0	0	0	1	0	1	0	0	-	-
x_{14}	0	0	0	2	0	0	0	2	0	0	2	0	1	-

Table 4. Co-occurrence matrix.

Experts decide that they want to make 5 different alternatives groups. Therefore, the first GDM process will be carried out with 5 different alternatives. The final assigned categories are shown in Table 5. Also, the resulting dendogram given by R software is shown in Figure 1.

Cat. 1	Cat. 2	Cat. 3	Cat. 4	Cat. 5
x_8	x_6	x_4	x_3	x_1
x_{11}	x_{12}	x_7	x_5	x_2
x_{14}		x_9		x_{10}
				x_{13}

Table 5. Categories obtained by the applied hierarchical clustering procedure.

Once that the groups are made, experts start carrying out the first GDM process. They decide to discuss until a consensus value of 0.60 is reached. For carrying out this process, the next steps are followed:

- **Preference providing step**: Experts provide their preference value by using the linguistic label set $S = \{s_1, s_2, s_3, s_4, s_5\}$. After carrying out the debate, they provide the following preference values:

$$P_1 = \begin{pmatrix} - & s_2 & s_1 & s_1 & s_3 \\ s_2 & - & s_2 & s_2 & s_2 \\ s_5 & s_4 & - & s_5 & s_4 \\ s_1 & s_3 & s_3 & - & s_3 \\ s_3 & s_2 & s_1 & s_3 & - \end{pmatrix} P_2 = \begin{pmatrix} - & s_2 & s_1 & s_1 & s_2 \\ s_1 & - & s_2 & s_3 & s_2 \\ s_5 & s_5 & - & s_5 & s_4 \\ s_1 & s_2 & s_3 & - & s_3 \\ s_1 & s_1 & s_2 & s_1 & - \end{pmatrix} P_3 = \begin{pmatrix} - & s_2 & s_1 & s_2 & s_3 \\ s_2 & - & s_2 & s_2 & s_1 \\ s_5 & s_5 & - & s_5 & s_5 \\ s_1 & s_3 & s_3 & - & s_1 \\ s_2 & s_2 & s_2 & s_2 & - \end{pmatrix}$$

- **Aggregating preference values into the collective preference matrix**: By applying expression (1), the

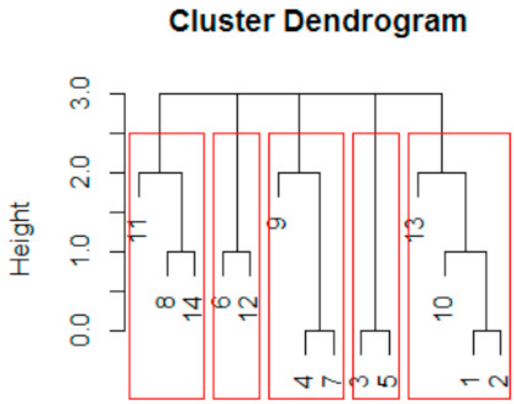

Fig. 1. Co-occurrence matrix associated dendogram.

information is aggregated. The next results are obtained:

$$P_c = \begin{pmatrix} - & 2 & 1 & 1.333 & 2.667 \\ 1.667 & - & 2 & 2.333 & 1.667 \\ 5 & 4.667 & - & 5 & 4.333 \\ 1 & 2.667 & 3 & - & 2.333 \\ 2 & 1.667 & 1.667 & 2 & - \end{pmatrix}$$

- **Calculating the groups ranking**: By using expressions (3) and (2), the ranking of alternatives groups is performed. It should be noticed that all the values are expressed in the interval [0,1]. Results are shown in Table 6.

	x_1	x_2	x_3	x_4	x_5
GNDD	0.8	0.833	1	0.8833	0.8166
GDD	0.15	0.2166	0.933	0.366	0.2166
Ranking value	0.475	0.525	0.966	0.625	0.5166
Ranking	x_3	x_4	x_2	x_5	x_1

Table 6. Ranking of categories.

- **Calculating consensus**: The three presented levels of consensus are calculated. Results are shown below:

$$C_c = \begin{pmatrix} - & 1 & 1 & 0.555 & 0.555 \\ 0.555 & - & 1 & 0.555 & 0.555 \\ 1 & 0.555 & - & 1 & 0.555 \\ 1 & 0.555 & 1 & - & 0.111 \\ 0.333 & 0.555 & 0.555 & 0.333 & - \end{pmatrix}$$

Consensus reached on each alternative group is shown in Table 7. Global consensus value is 0.6533. Since the obtained value is greater than 0.60, it is considered that there is enough consensus. Therefore, there is no need of more GDM rounds.

x_1	x_2	x_3	x_4	x_5
0.82	0.64	0.73	0.64	0.42

Table 7. Consensus results obtained for the categories related GDM process.

The most preferred alternatives group is the one that have the following alternatives: $\{x_7, x_9, x_{12}\}$. Therefore, a new GDM process must be carried out over that alternatives. Consensus is again set to 0.6. Also, the same linguistic label set is is used for providing preferences to the system. The GDM process is repeated in order to rank the alternatives of the selected alternatives group:

- **Preference providing step**: Experts provide the following preferences:

$$P_1 = \begin{pmatrix} - & s_4 & s_5 \\ s_1 & - & s_2 \\ s_1 & s_1 & - \end{pmatrix} P_2 = \begin{pmatrix} - & s_4 & s_4 \\ s_1 & - & s_2 \\ s_1 & s_2 & - \end{pmatrix}$$

$$P_3 = \begin{pmatrix} - & s_5 & s_5 \\ s_1 & - & s_2 \\ s_1 & s_2 & - \end{pmatrix}$$

- **Calculating the collective preference matrix**: The provided preferences are aggregated. The following result is obtained:

$$P_c = \begin{pmatrix} - & 4.333 & 4.666 \\ 1 & - & 2 \\ 1 & 1.666 & - \end{pmatrix}$$

- **Temporary alternatives ranking calculation**: Next, the ranking of alternatives is obtained by using selection operators. Results are shown in Table 8.

	x_7	x_9	x_{12}
GDD	0.888	0.983	0.111
GNDD	1	0.722	0.666
Ranking value	0.944	0.402	0.388
Ranking	x_7	x_9	x_{12}

Table 8. Ranking of alternatives.

- **Calculating consensus measures**: In this step, the level of agreement among the experts is obtained by using consensus measures. For each pair of alternatives, the following results are obtained:

$$C_c = \begin{pmatrix} - & 0.555 & 0.555 \\ 1 & - & 1 \\ 1 & 0.555 & - \end{pmatrix}$$

  Consensus results for each alternative can be seen in Table 9. The obtained level of consensus in this step is 0.707, which is above the established consensus threshold, that is, 0.6. Therefore, the calculated ranking is considered as the final one and x_7 is obtained as the most voted alternative.

x_7	x_9	x_{12}
0.55	1	0.55

Table 9. Obtained consensus value in the alternative selection process.

5. Conclusions

A novel GDM method that deals with environments that employ a large set of alternatives has been presented. Clustering methods were used in order to create groups of alternatives that the experts can use to discuss about. The initial GDM process with a high number of alternatives was converted into two processes where both have a reduced set of alternatives that the experts have to debate about. Thanks to this, the amount of information that the experts have to deal with at the same time is drastically reduced. First, experts create groups of alternatives. Next, they discuss which group is the most adequate. Finally, the ranking of alternatives is calculated by employing a GDM procedure.

Acknowledgements

The authors would like to thank the FEDER financial support for the Project TIN2016-75850-P by the Spanish Ministry of Science, Innovation and Universities. Also, they want to acknowledge the Juan de la Cierva program of the Ministry of Science, Innovation and Universities.

References

[1] J. A. Morente-Molinera, R. Al-Hmouz, A. Morfeq, A. S. Balamash, E. Herrera-Viedma, A decision support system for decision making in changeable and multi-granular fuzzy linguistic contexts., Journal of Multiple-Valued Logic & Soft Computing 26 (2016) 485–514.
[2] F. Herrera, S. Alonso, F. Chiclana, E. Herrera-Viedma, Computing with words in decision making: foundations, trends and prospects, Fuzzy Optimization and Decision Making 8 (4) (2009) 337–364.
[3] Y. Liu, Y. Dong, H. Liang, F. Chiclana, E. Herrera-Viedma, Multiple attribute strategic weight manipulation with minimum cost in a group decision making context with interval attribute weights information, IEEE Transactions on Systems, Man, and Cybernetics: Systems.
[4] J. A. Morente-Molinera, G. Kou, Y. Peng, C. Torres-Albero, E. Herrera-Viedma, Analysing discussions in social networks using group decision making methods and sentiment analysis, Information Sciences 447 (2018) 157–168.
[5] P. Arabie, L. J. Hubert, G. De Soete, Clustering and classification, World Scientific, 1996.
[6] L. de la Fuente-Tomas, B. Arranz, G. Safont, P. Sierra, M. Sanchez-Autet, A. Garcia-Blanco, M. P. Garcia-Portilla, Classification of patients with bipolar disorder using k-means clustering, PloS one 14 (1) (2019) e0210314.
[7] C.-C. Li, Y. Dong, F. Herrera, A consensus model for large-scale linguistic group decision making with a feedback recommendation based on clustered personalized individual semantics and opposing consensus groups, IEEE Transactions on Fuzzy Systems 27 (2) (2019) 221–233.
[8] X. Liu, Y. Xu, R. Montes, R.-X. Ding, F. Herrera, Alternative ranking-based clustering and reliability index-based consensus reaching process for hesitant fuzzy large scale group decision making, IEEE Transactions on Fuzzy Systems 27 (1) (2019) 159–171.
[9] J. A. Morente-Molinera, G. Kou, K. Samuylov, R. Ureña, E. Herrera-Viedma, Carrying out consensual group decision making processes under social networks using sentiment analysis over comparative expressions, Knowledge-Based Systems 165 (2019) 335–345.
[10] H. Zhang, Y. Dong, F. Chiclana, S. Yu, Consensus efficiency in group decision making: A comprehensive comparative study and its optimal design, European Journal of Operational Research 275 (2) (2019) 580–598.
[11] J. Morente-Molinera, I. Pérez, M. Ureña, E. Herrera-Viedma, On multi-granular fuzzy linguistic modeling in group decision making problems: A systematic review and future trends, Knowledge-Based Systems 74 (2015) 49–60.
[12] J. A. Morente-Molinera, G. Kou, C. Pang, F. J. Cabrerizo, E. Herrera-Viedma, An automatic procedure to create fuzzy ontologies from users' opinions using sentiment analysis procedures and multi-granular fuzzy linguistic modelling methods, Information Sciences 476 (2019) 222–238.
[13] M. J. del Moral, F. Chiclana, J. M. Tapia, E. Herrera-Viedma, A comparative study on consensus measures in group decision making, International Journal of Intelligent Systems 33 (8) (2018) 1624–1638.
[14] I. J. Pérez, F. J. Cabrerizo, S. Alonso, Y. Dong, F. Chiclana, E. Herrera-Viedma, On dynamic consensus processes in group decision making problems, Information Sciences 459 (2018) 20–35.
[15] N. H. Kamis, F. Chiclana, J. Levesley, Geo-uninorm consistency control module for preference similarity network hierarchical clustering based consensus model, Knowledge-Based Systems 162 (2018) 103–114.
[16] V. Siless, K. Chang, B. Fischl, A. Yendiki, Anatomicuts: Hierarchical clustering of tractography streamlines based on anatomical similarity, Neuroimage 166 (2018) 32–45.
[17] F. J. Cabrerizo, E. Herrera-Viedma, W. Pedrycz, A method based on pso and granular computing of linguistic information to solve group decision making problems defined in heterogeneous contexts, European Journal of Operational Research 230 (2013) 624–633.
[18] I. J. Pérez, F. J. Cabrerizo, E. Herrera-Viedma, A mobile decision support system for dynamic group decision-making problems, Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on 40 (6) (2010) 1244–1256.
[19] F. J. Cabrerizo, J. M. Moreno, I. J. Pérez, E. Herrera-Viedma, Analyzing consensus approaches in fuzzy group decision making: advantages and drawbacks, Soft Computing 14 (5) (2010) 451–463.