Contemporary Issues in Business, Management and Education 2013

# Synergies of data mining and multiple attribute decision making

Mohammad Hasan Aghdaie[a,b], Sarfaraz Hashemkhani Zolfani[c,d], Edmundas Kazimieras Zavadskas[a]*

[a] *Vilnius Gediminas Technical University, Institute of Internet and Intelligent Technologies Sauletekio al. 11, LT-10223 Vilnius, Lithuania*
[b] *Shomal University, Department of Industrial Engineering, P. O. Box 731, Amol, Mazandaran, Iran*
[c] *Amirkabir University of Technology (Tehran Polytechnic), Technology Foresight Group, Department of Management, Science and Technology, P.O. Box 1585-4413, Tehran, Iran*
[d] *Amirkabir University of Technology (Tehran Polytechnic), Futures Studies Research Institute, P.O. Box 1585-4413, Tehran, Iran*

## Abstract

Data Mining (DM) and Multiple Attribute Decision Making (MADM) are two fast growing trends in Operations Research (OR) /Management Science (MS). In this article, we identify the synergies of data mining and MADM. Synergies can be attained by integration of MADM techniques into data mining and vice versa. The primary goal of the paper is to show a wide range of interactions between these two fields from a new perspective with an example of the integrated approach in supplier clustering and ranking. The integrated approach includes cluster analysis as a data mining tool and Step-wise Weight Assessment Ratio Analysis (SWARA) and VIseKriterijumskao ptimizacija i KOmpromisno Resenje (VIKOR) as the two MADM tools. More precisely, the features for clustering were selected and weighted by SWARA method and suppliers are clustered using two-stage cluster analysis. In addition, VIKOR technique is used to rank the clusters from the best to the worst one. The proposed integrated approach is presented to demonstrate the applicability of the proposed methodology.

*Keywords*: data mining; Multiple Attribute Decision Making (MADM); Clustering; SWARA;VIKOR; Supplier clustering and ranking.

## 1. Introduction

Recently, modern decision making needs much more sophisticated and accurate techniques. Organizations have very large databases of information and rival companies have to engage in cut-throat competition. Nowadays, we

_____

* Corresponding author . Tel.: tel:+37069820779
*E-mail address:* edmundas.zavadskas@vgtu.lt

can use supercomputers to simulate a real decision making environment and solve the complicated real models. A demand for the integration between different fields in order to make better decisions is unavoidable.

Management Science (MS) /operations Research (OR) is an old and famous field that deals with making decisions. Data mining (DM) can be defined as the process of extracting important and useful information from large sets of data (Abello *et al.*, 2002). According to Han and Kamber (2001), some of the most important functions of data mining include concept description (characterization and discrimination), association, classification, clustering, and prediction. In addition, DM is an interdisciplinary field that combines artificial intelligence, database management, data visualization, machine learning, mathematic algorithms, and statistics (Tsai, 2012). Recently, there has been an increasing interest in the integration of OR and DM (Meisel & Mattfeld, 2010; Corne *et al.*, 2012). For instance, data mining can be helpful in many OR application areas and can be used in a complementary way to optimization method to identify constraints and reduce the search space (Olafsson *et al.*, 2008).

Multiple criteria decision making (MCDM) or multiple criteria decision analysis (MCDA) is a sub-discipline of OR which deals with multiple criteria in decision environment. Two main categories of MCDM are multiple attributes decision making (MADM) and multi objective decision making (MODM). The MADM methods deal with the process of making decisions in finding the optimum/best alternative in the presence of multiple, usually conflicting, decision criteria. MADM techniques can be used as an analytical approach to assess, weigh or rank a set of criteria or alternatives.

In the literature there are some papers about the integration of DM-MCDM techniques. For example, Rad *et al.* (2011) used K-mean clustering and analytic hierarchy process (AHP) to cluster and rank university majors. Peng *et al.* (2011) integrated data integration, data mining, and multi-criteria decision making and designed an incident information management framework. Khalili-Damghania *et al.* (2013) applied fuzzy logic, data mining and MODM in project selection. In addition, they developed a hybrid framework based on a hybrid fuzzy rule-based multi-objective framework and data envelope analysis (DEA). A combination between DM and MCDM methods in the decision support system (DSS) was introduced by Khademolqorani and Hamadani, (2012). Aghdaie *et al.* (2013a) used data mining and MADM for market segmentation and market segment evaluation and selection. Kim *et al.* (2011) proposed an approach which was comprised of two methods: association rule mining (ARM) and the analytic network process (ANP).

The aim of this paper is to propose a new hybrid DM-MADM approach to cluster suppliers, rank the clusters and finally grade them. A two-stage cluster analysis which is developed by Punj and Steward (1983) is applied as a DM tool. Two-stage clustering is based on a combination of K-means (MacQueen, 1967) algorithm and Ward's method (Ward, 1963). Two-stage clustering manipulates benefits of the two methods. In addition, in order to select and weigh attributes which two-stage cluster analysis is needed, SWARA is applied. Finally, two MADM methods including SWARA and VIKOR are used to rank the clusters of the suppliers.

The remainder of the paper is organized as follows. The next section outlines the proposed methodology combining, Clustering, SWARA, and VIKOR. In Section 3, a real-world data is given to prove the applicability of the proposed method on a car manufacturing industry in Iran. Also in Section 3, the results obtained. In Section 4, the article's conclusion will be presented and future directions are discussed.

## 2. The proposed model

The proposed model section describes three points. In the first part of this section, the proposed integrated DM-MADM model as a new approach is explained. In other parts, SWARA and VIKOR methods are explained.

### 2.1. The proposed integrated DM-MADM model

This section describes a three phase methodology which is used for suppliers clustering and cluster evaluation and selection (see Fig. 1). In this conceptual model two approaches, including DM and MADM have been combined.

The first phase is a data preparation phase and includes two steps. In the first step, the most important features needed for cluster analysis were selected by literature survey and experts' opinions. In DM approaches, feature/attribute selection is a very important part. A lot of output is based on the selected features. The attributes
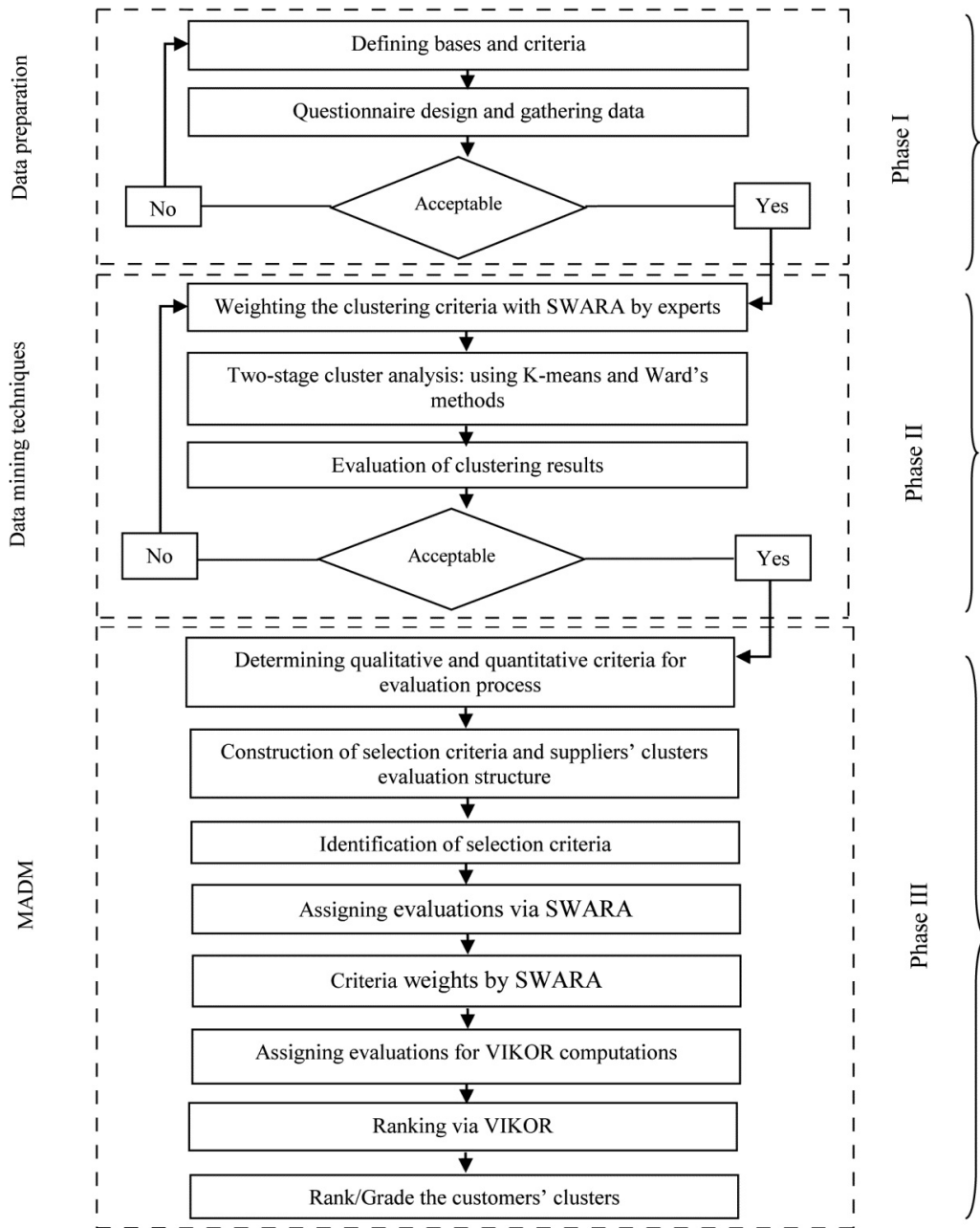
Fig. 1. Flow chart illustrating the integrated DM-MADM approach

that were used for cluster analysis were both qualitative and quantitative, including a number of supplier's employees, a variety of products, product quality, certifications, risk, finance position, warranties and claim policies, delivery lead time, delivery on time, transportation costs, product price, and complexity of the products. Some of these attributes were qualitative, so 1-9 scoring scale was used to score every supplier for qualitative criteria. They were denoted as F1, F2, F3, F4, F5, F6, F7, F8, F9, F10, F11, and F12, respectively. Then, a questionnaire was

designed and data were gathered. Next, the prepared data were checked to make it clear that the data is acceptable for analysis or not.

The second phase is a data mining phase. In the first step, by using one of the famous MADM tools, namely SWARA weight of each feature for cluster analysis was identified. Then, two-stage cluster analysis was done in order to classify suppliers into different groups. Finally, clustering results were evaluated and if the results were acceptable, the next phase could be started.

The last phase called MADM phase. In this phase, different clusters of suppliers were assessed and graded. In addition, for conducting the third phase, eight steps passed. In the first step of this phase, the most important criteria for evaluation of clusters were identified. Next, the qualitative and quantitative criteria were selected. The criteria list was selected based on the literature survey. Finally, the project team constructed the selection criteria and problem structure. Criteria weights were calculated applying SWARA method and based on experts' evaluations. At this stage, all the clusters were assessed by a group of experts and VIKOR method was applied to achieve the final ranking results.

### 2.2. Step-wise Weight Assessment Ratio Analysis (SWARA)

Weight assessment is an important issue in many MADM problems (Hashemkhani Zolfani *et al.*, 2013). SWARA method is one of the brand-new ones in which an expert has an important role in evaluations and calculating weights. Based on this method, the most significant criterion is given rank 1, and the least significant criterion is given rank last. The overall ranks of the group of experts are determined according to the mediocre value of ranks (Kersuliene *et al.*, 2010; Kersuliene & Turskis, 2011). Therefore, SWARA can be useful for some issues that priorities are known formerly according to situations and finally SWARA proposed for applying in certain environments of decision making.

### 2.3. VIseKriterijumska optimizacija i KOompromisno Resenje (VIKOR) VIKORmethod

The VIKOR method or multi-criteria optimization and compromise solution is a compromise MADM method, developed by Opricovic and Tzeng (Opricovic, 1998; Opricovic, Tzeng 2002). The concept of VIKOR is based on the compromise programming of MCDM by comparing the measure of ''closeness'' to the ''ideal'' alternative (Wu *et al.*, 2009).

1) Calculate the normalized value:

$$f_{ij} = \frac{x_{ij}}{\sqrt{\sum_{j=1}^{n} x_{ij}^2}} i = 1, 2, \dots, m; j = 1, 2, \dots, n \tag{1}$$

2) Determine the best and worst values:

For all the attribute functions the best value was $f_j^+$ and the worst value was $f_j^-$ that is, for attribute $J = 1 - n$, we get formulas (2) and (3)

$$f_j^+ = \max f_{ij}, i = 1, 2, \dots, m \tag{2}$$

$$f_j^- = \min f_{ij}, i = 1, 2, \dots, m \tag{3}$$

Where $f_j^+$ is the positive ideal solution for the $j$ th criteria, $f_j^-$ is the negative ideal solution for the $j$ th criteria.

If one associates all $f_j^+$ one will have the optimal combination, which gets the highest scores, the same as $f_j^-$.
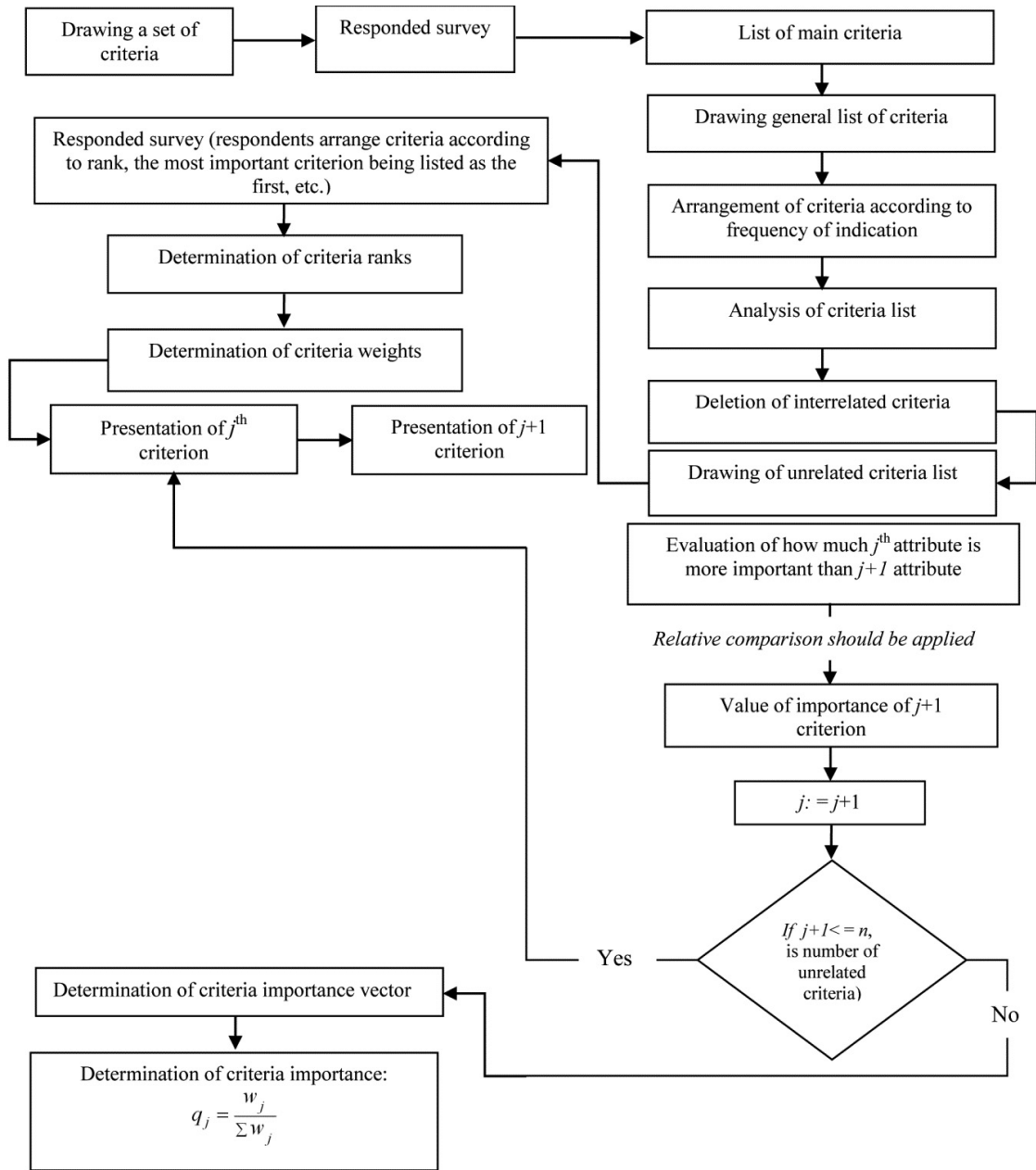
Fig. 2. Determining of the criteria weights based on (Kersuliene *et al.,* 2010)

3) Determine the weights of attributes:
The weights of attribute should be calculated to express their relative importance.
4) Compute the distance of alternatives to ideal solution:
This step is to calculate the distance from each alternative to the positive ideal solution and then get the sum to obtain the final value according to formula (4) and (5).

$$S_i = \sum_{j=1}^{n} w_j \left( f_j^+ - f_{ij} \right)\left( f_j^+ - f_j^- \right) \tag{4}$$

$$R_i = Max_j \sum_{j=1}^{n} w_j \frac{\left( f_j^+ - f_{ij} \right)}{\left( f_j^+ - f_j^- \right)} \tag{5}$$

Where $S_i$ represents the distance rate of the $i$ th alternative to the positive ideal solution (best combination), $R_i$ represents the distance rate of the $i$ th alternative to the negative ideal solution (worst combination). The excellence ranking will be based on $S_i$ values and the worst rankings will be based on $R_i$ values. In other words, $S_i$, $R_i$ indicate $L_{1i}$ and $L_{0i}$ of $L_P - metirc$ respectively.

5) Calculate the VIKOR values $Q_i$ for $i = 1, 2, ..., m$ which are defined as:

$$Q_i = v \left[ \frac{S_i - S_i^+}{S^- - S^+} \right] + (1-v)\left[ \frac{R_i - R_i^+}{R^- - R^+} \right] \tag{6}$$

Where $S^- = Max_i S_i$, $S^+ = Min_i S_i$, $R^- = Max_i R_i$, $R^+ = Min_i R_i$, and $v$ is the weight of the strategy of "the majority of criteria'' (or "the maximum group utility"). $\left[ (S - S^*)/(S^- - S^*) \right]$ represents the distance rate from the positive ideal solution of the $i$ th alternative's achievements In other words, the majority agrees to use the rate of the $i$ th. $\left[ (R - R^*)/(R^- - R^*) \right]$ represents the distance rate from the negative ideal solution of the $i$ th alternative; this means the majority disagree with the rate of the $i$ th alternative. Thus, when the $v$ is larger ($> 0.5$), the index of $Q_i$ will tend to majority agreement; when $v$ is less ($< 0.5$), the index $Q_i$ will indicate majority negative attitude; in general, $v = 0.5$, i.e. compromise attitude of evaluation experts.

6) Rank the alternatives by $Q_i$ values:

According to the $Q_i$ values calculated by step (4), we can rank the alternatives and to make-decision.

## 3. Case study

A real case study was chosen to show the performance and application of the model. The study was conducted by one of the well-known manufacturing automobile producers in Iran. This company is located near Tehran, and it is a giant company. In addition, the enormous automobile manufacturer has hundreds of suppliers. Recently, there is a policy to grade suppliers by using analytical tools. Therefore, after receiving general agreement about defining a new grading project, because of the mentioned reasons; a project team which consists of nineteen experts with different education background was constructed. The project team identified twelve features that were denoted in Table 1, for clustering. In order to improve the quality of cluster analysis, SWARA was used to weight the features.

For receiving general agreement in every step of this project, Delphi method was used, since Delphi is a very famous method for receiving general agreement in complicated decision making situations (Aghdaie *et al.*, 2013b). After weighting the features and calculating the weight of each criterion, the data which was gathered from 300 suppliers of the company, was used for cluster analysis. A two-stage cluster analysis which is a combination of Ward's and K-means methods was applied to group customers based on twelve features. In the first stage of clustering, the optimal number of clusters based on Ward's method was calculated. The optimal number of clusters was four and consequently customers were classified into four segments. Also, Schwarz's Bayesian criterion (BIC) gave the best fit for four clusters. Besides, an average silhouette measure that is slightly greater than 0.6 indicates a reasonable partitioning. In the second stage, K-means method was used to improve the results from Ward's method. In performing the K-means method, the initial seed points were taken from the derived cluster centers which derived from Ward's method.

Table 1. Final results of SWARA method in weighting criteria

| Criterion | The comparative importance of average value $S_j$ | Coefficient $k_j = s_j + 1$ | Recalculated weight $w_j = \dfrac{x_{j-1}}{k_j}$ | Weight $q_j = \dfrac{w_j}{\Sigma\, w_j}$ |
|---|---|---|---|---|
| $F_3$ |  | 1 | 1 | 0.130 |
| $F_9$ | 0.15 | 1.15 | 0.870 | 0.113 |
| $F_8$ | 0.05 | 1.05 | 0.828 | 0.108 |
| $F_{11}$ | 0.05 | 1.05 | 0.789 | 0.103 |
| $F_5$ | 0.10 | 1.10 | 0.717 | 0.093 |
| $F_{12}$ | 0.05 | 1.05 | 0.683 | 0.089 |
| $F_2$ | 0.20 | 1.20 | 0.569 | 0.074 |
| $F_6$ | 0.05 | 1.05 | 0.542 | 0.071 |
| $F_7$ | 0.05 | 1.05 | 0.516 | 0.067 |
| $F_4$ | 0.10 | 1.10 | 0.469 | 0.061 |
| $F_9$ | 0.20 | 1.20 | 0.391 | 0.051 |
| $F_1$ | 0.25 | 1.25 | 0.313 | 0.041 |

The most important interpretations of the segments can be defined as follows:

Cluster 1:This cluster consists of 115 (38%) suppliers and it is the biggest cluster. Most of the suppliers in this cluster produce a few number of different products. They do not have special certificates andthe number of employees are under 50. These suppliers manufacture cheap and simple products. The delivery lead time bias in this group is large.

Cluster 2:This cluster consists of 84 (28%) and they are responsible for producing a host of different simple products with average prices. They have general certificates and the companies are medium size. They are far from the company.

Cluster 3: This cluster consists of 45 (15%)and the size of this cluster is small. However, most of the companies in it are very large with more than 300 employees. They produce a variety of products that are often complicated. The price of their goods is high since the delivery time and lead time of them are in high standard.

Cluster 4:This cluster consists of 56 (19%) and the quality of their products is very high. The price of their goods is average. Although the transportation costsare very high because of the distance. They can produce an average number of goods with good guarantee.

In phase III, the group of experts determined the qualitative and quantitative criteria for the evaluation process and grade the clusters. Then, they created a structure for suppliers' clusters evaluation and identified the final criteria (see Fig. 3). The grading project is very important for the company as it helps the company to efficiently deal with its suppliers.

As mentioned before, Fig. 3 shows the hierarchical structure of the problem, which were designed by the group of experts. Every cluster which consists of a group of suppliers, considered as an alternative. Thus, alternatives $A_1$, $A_2$, $A_3$, and $A_4$, are equal to Cluster 1, Cluster 2, Cluster 3, and Clsuter 4, respectively. Table 2 shows the definition of all criteria for the evaluation process. In the evaluation process, an average score of the cluster for quantitative criterion was used. Then, all the qualitative and quantitative criteria assessed qualitatively based on 1 to 9 scoring scale.

Among all the criteria one criterion C1 is cost criterion (the minimum amount of this criterion is desirable) and others are benefit criteria. This kind of classification is important for VIKOR calculations.
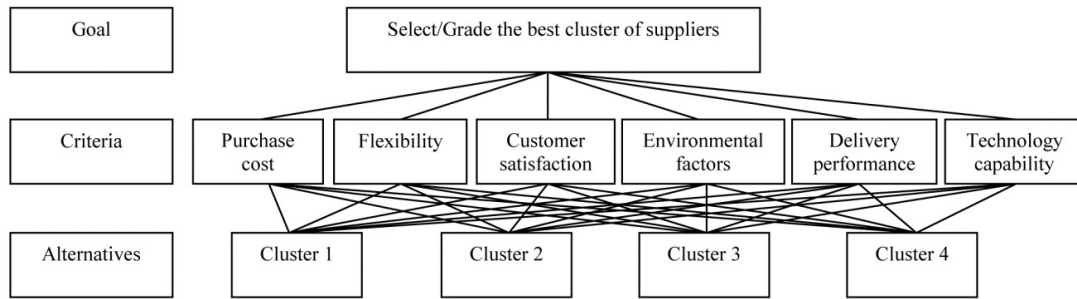
Fig. 3. The hierarchical structure of the problem, selection criteria and alternatives

Table 2. The supplier cluster evaluation criteria

| Criteria | Definition |
|---|---|
| C1: Purchase cost | The purchase cost that determines the product price includes material cost, logistics cost, quantity discount, maintenance cost, and warranty cost, etc. |
| C2: Flexibility | The flexibility includes items such as product flexibility, manufacture flexibility, product volume flexibility, multi-skilled and flexible people. |
| C3: Customer satisfaction | The degree of customers' needs satisfaction |
| C4: Environmental factors | How much the topics including, green production, ergonomy, pollution, energy are considered? |
| C5: Delivery performance | The delivery lead time and delivery on time criteria are considered as delivery performance |
| C6: Technology capability | Ability to meet current and future demands of the firm. |

The first column of the Table 3 shows criteria weights based on SWARA evaluations and the final weight of each criterion. The last part of the table indicates decision making evaluations by the experts' group.

Four clusters (alternatives) were assessed by the group of experts based on six criteria in order to grade each cluster. Nineteen experts used the Delphi method to reach consensus. Their scores are shown in the decision making part in Table 3. Then, Eqs (1) to (6) were applied to rank clusters from the best to the worst one. The final ranking of the alternatives is given in Table 4. The sixth column of this table shows the final ranking according to the VIKOR methodology.

Table 3. Final results of SWARA method in weighting criteria

| Criterion | Criteria weights based on SWARA | | | | Decision matrix | | | |
|---|---|---|---|---|---|---|---|---|
| | Comparative importance of average value $S_j$ | Coefficient $k_j = s_j + 1$ | Recalculated weight $w_j = \dfrac{x_{j-1}}{k_j}$ | Weight $q_j = \dfrac{w_j}{\Sigma w_j}$ | $A_1$ | $A_2$ | $A_3$ | $A_4$ |
| $C_3$ | | 1 | 1 | 0.241 | 2 | 4 | 6 | 7 |
| $C_5$ | 0.15 | 1.15 | 0.870 | 0.210 | 1 | 3 | 8 | 3 |
| $C_1$ | 0.20 | 1.20 | 0.725 | 0.175 | 2 | 4 | 7 | 7 |
| $C_2$ | 0.30 | 1.30 | 0.558 | 0.135 | 6 | 3 | 4 | 5 |
| $C_6$ | 0.10 | 1.10 | 0.507 | 0.122 | 2 | 4 | 7 | 3 |
| $C_4$ | 0.05 | 1.05 | 0.483 | 0.117 | 4 | 7 | 5 | 6 |

Table 4. Ultimate results and ranking of the alternatives

| Alternatives | $S_i$ | $R_i$ | $V_i$ | $Q_i$ | Ranking | Grade |
|---|---|---|---|---|---|---|
| $A_1$ | 0.75 | 0.241 | 0.5 | 1.0000 | 4 | D |
| $A_2$ | 0.5728 | 0.15 | 0.5 | 0.2221 | 3 | C |
| $A_3$ | 0.4312 | 0.175 | 0.5 | 0.1374 | 1 | A |
| $A_4$ | 0.5266 | 0.175 | 0.5 | 0.1851 | 2 | B |

According to Table 4 which shows ultimate results of VIKOR methodology, the third alternative (Cluster 3) is the best cluster of suppliers. Based on this table, cluster 3hasgrade A, while others can be categorized B, C, and D. Also, the proposed hybrid model provides a systemically analytic model for clustering and ranking in a real situation with a combination of DM and MADM tools.

## 4. Conclusions and future research directions

Nowadays, the capability to collect and generate data has been vastly expanded and the huge amount of available data provides better resources for companies to make their decisions. OR/MS is the science of finding the best option for a decision making problem and data mining is the science of extracting information from a vast amount of data.

Data is the backbone of data mining and decision making. Every decision needs information, so the DM can be used as a tool to provide useful information for OR/MS tools. In addition, OR methods can be used in DM techniques for optimization. Recently, there has been a growing interest in integrating DM and OR in many fields.

One of the sub-disciplines of OR is MADM and this paper introduces new synergies between data mining and MADM. Also, in this paper, a novel hybrid approach for supplier clustering and cluster evaluation and selection with integrating DM mining and MADM methods was proposed. Two-stage cluster analysis was applied as a DM tool to group suppliers. The SWARA method was used to weigh the criteria for cluster analysis. In addition, SWARA method was applied as a decision making tool for extracting weights of criteria which is needed in VIKOR. Therefore, the results of SWARA were used as weighted inputs for VIKOR. The proposed SWARA-VIKOR integrated approach can be viewed as another meaningful contribution of this study.

Interest in integrating DM techniques with OR methods continue to grow and sub-disciplines of OR will absorb the new wave of interest soon. This interest is being addressed in the field of MADM at the beginning of the paper. However the field is very young, we hope that this paper provides a new perspective with motivation for researchers to this field. Furthermore, a new synergy can be integrated DM with MODM methods.

## References

Abello, J., Pardalos, P. M. & Resende, M. (2002). *Handbook of massive data sets*. Kluwer, New York.
Aghdaie, M.H., HashemkhaniZolfani, S. & Zavadskas, E.K. (2013a). 'A hybrid approach for market segmentation and market segment evaluation and selection: an integration of data mining and MADM', *Transformations in Business & Economics, 12 (2B (29B)),* 42–59.
Aghdaie, M. H., HashemkhaniZolfani, S. & Zavadskas, E. K. (2013b). Decision making in machine tool selection: An integrated approach with SWARA and COPRAS-G methods. *Inzinerine Ekonomika – Engineering Economics*, *24(1),* 5–17.
Corne, D., Dhaenens, C. & Jourdan., L. (2012). Synergies between operations research and data mining: The emerging use of multi-objective approaches,*European Journal of Operational Research, 221(3),* 469–479. http://dx.doi.org/10.1016/j.ejor.2012.03.039
Han, J. & Kamber, M. (2001). *Data mining concepts and techniques*. Morgan Kaufmann Publisher.
Hashemkhani Zolfani, S., Aghdaie, M. H., Derakhti, A., Zavadskas, E. K. & Varzandeh, M. H. M. (2013). Decision making on business issues with foresight perspective; an application of new hybrid MCDM model in shopping mall locating, *Expert Systems with Applications*, *40(17),*7111–7121. http://dx.doi.org/10.1016/j.eswa.2013.06.040
Kersuliene, V. & Turskis, Z. (2011). Integrated fuzzy multiple criteria decision making model for architect selection, *Technological and Economic Development of Economy, 17(4),* 645–666. http://dx.doi.org/10.3846/20294913.2011.635718
Keršulienė, V., Zavadskas, E. K. & Turskis, Z. (2010). Selection of rational dispute resolution method by applying new step-wise weight assessment ratio analysis (SWARA). *Journal of Business Economics and Management, 11(2),* 243–258. http://dx.doi.org/10.3846/jbem.2010.12
Khademolqorani, S. & Zeinal Hamadani, A. (2012). An Adjusted Decision Support System through Data Mining and Multiple Criteria Decision Making, *Proceedings of the 2nd International Conference on Integrated Information (IC-ININFO 2012)*, Budapest, Hungary, September *3,* 388–395.

Khalili-Damghania, K., Sadi-Nezhadb, S., Lotfic, F.H. & Tavanad, M. (2013). A hybrid fuzzy rule-based multi-criteria framework for sustainable project portfolio selection, *Information Sciences, 220,* 442–462. http://dx.doi.org/10.1016/j.ins.2012.07.024

Kim, C., Lee, H., Seol, H. & Lee, C. (2011). Identifying core technologies based on technological cross-impacts: An association rule mining (ARM) and analytic network process (ANP) approach, *Expert Systems with Applications, 38(10),* 12559–12564. http://dx.doi.org/10.1016/j.eswa.2011.04.042

MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 281– 297.

Meisel, S. & Mattfeld, D. (2010). Synergies of Operations Research and Data Mining, *European Journal of Operational Research, 206 (1),* 1–10. http://dx.doi.org/10.1016/j.ejor.2009.10.017

Olafsson, S., Li, X. & Wu, S. (2008). Operations research and data mining, *European Journal of Operational Research*, *187(3),* 1429–1448. http://dx.doi.org/10.1016/j.ejor.2006.09.023

Opricovic, S. (1998). Multi criteria optimization of civil engineering systems, "Faculty of Civil Engineering", Belgrade (In Serbian).

Opricovic, S. & Tzeng, G. H. (2002). Multi-criteria planning of post earthquake sustainable reconstruction, *Computer-Aided Civil and Infrastructure Engineering, 17,* 211–220. http://dx.doi.org/10.1111/1467-8667.00269

Peng, Y., Zhang, Y., Tang, Y. & Li. S. (2011). An incident information management framework based on data integration, data mining, and multi-criteria decision making, *Decision Support Systems, 51(2),* 316–327. http://dx.doi.org/10.1016/j.dss.2010.11.025

Punj, G., & Stewart, D.W. (1983). Cluster analysis in Marketing Research: Review and Suggestions for Applicationd, *Journal of Marketing Research, 20,* 134–148. http://dx.doi.org/10.2307/3151680

Rad, A., Naderi, B. & Soltani, M. (2011).Clustering and ranking university majors using data mining and AHP algorithms: A case study in Iran, *Expert Systems with Applications, 38(1),* 755–763. http://dx.doi.org/10.1016/j.eswa.2010.07.029

Tsai, H. (2012). Global data mining: An empirical study of current trends, future forecasts and technology diffusions, *Expert Systems with Applications, 39(9),* 8172–8181. http://dx.doi.org/10.1016/j.eswa.2012.01.150

Ward, J. (1963). Hierarchical grouping to optimize an objective unction. *Journal of the American Statistical Association*, *58,* 236–244. http://dx.doi.org/10.1080/01621459.1963.10500845

Wu, C., Barnes, D., Rosenberg, D. & Luo, X. X. (2009). An analytic network process- mixed integer multi-objective programming model for partner selection in agile supply chains, *Production Planning & Control*, *20(3),* 254–275. http://dx.doi.org/10.1080/09537280902856047