All Sciences Proceedings
All Sciences Proceedings
http://as-proceeding.com/
5th International Conference on Applied Engineering and Natural Sciences
July 10-12, 2023 : Konya, Turkey

https://www.icaens.com/

Data Scientist Selection in IT Industry: SMCDM Approach

Emine Nur Nacar*

[1]Department of Industrial Engineering, Faculty of Engineering and Natural Sciences, Ankara Yildirim Beyazit University, Turkey

*(ennacar@ybu.edu.tr) Email of the corresponding author

Abstract **–** Traditional Multi-Criteria Decision Making (MCDM) methods cannot provide solutions to problems that may be encountered in the near future. Basically, MCDM methods have a very rich algorithm literature for decision making problems. Traditional MCDM methods do not take short-term changes into account. In a new development, it is necessary to start the process from the beginning and solve the problem from the beginning. This leads to huge loss of money and time. The purpose of developing the MCDM method is to make decision-making problems more efficient and to prevent losses in the process. However, the MCDM method is a method that has been developed based on the changes in the process, which is a problem that has not been addressed until now. Since it will be tiring and difficult to solve the problem by assigning new weights from the beginning, the Stratified Multi-Criteria Decision Making (SMCDM) Method has been developed for the possibilities that may occur in the near future. In this study, an exemplary study of the SMCDM method was carried out in a IT company. In the data scientist selection problem, the best alternative was selected with the SMCDM method, taking into account the events that may occur in the near future.

Keywords – Smcdm, Candidate Selection, Logistics Industry, Stratified, Mcdm

I. Introduction

Multi-criteria decision making (MCDM) methods generally aim to select the best alternative by sorting among the alternatives, taking into account the decision variables and their weights. Solving an MCDM problem relies on an objective investigation of the impact of alternatives on attributes and a subjective assessment of the decision maker's preference system [1]. Therefore, once the decision is made, the decision maker may doubt that the correct weights are given to the criteria. Because these criteria may change in the near future. Events that may occur in the near future should not be ignored in traditional multi-criteria decision making methods. The structure of MCDM methods does not seem suitable for adding the events expected to happen in the near future to the algorithm. For this reason, a layered multi-criteria decision-making method (SMCDM), which takes into account the events expected to happen in the near future, has been developed [2].

In this method, which has just been introduced to the literature, an integrated approach is used by hybridizing the concept of stratification and the multi-criteria decision-making method. The concept of stratification is easy to understand and implement, with potential for important applications in planning, robotics, optimal control, tracking, multi-target optimization, exploration, exploration, and other fields [3].

The concept of stratification describes a system that goes through several states to reach a desired state. This concept is an effective and easy problem solving approach. Therefore, it is emerging as a major area of interest in the coming years [4].

In this method, which has just been introduced to the literature, an integrated approach is used by hybridizing the concept of stratification and the MCDM method. The concept of stratification is easy to understand and implement, with potential for important applications in planning, robotics, optimal control, tracking, multi-target optimization, exploration, exploration, and other fields [3]. The concept of stratification describes a system that goes through several states to reach a desired state. This concept is an effective and easy problem solving approach. It is assumed that this approach will be applied to different fields in the coming years [5].

II. Materials and Method

In this approach developed by Asadabadi, a new method was created by using the traditional multi-criteria decision-making method with an integrated approach by hybridizing with the concept of stratification. The operation of the method is as follows [2]:

The set of alternatives $\{a_1, a_2, \dots, a_n\}$ and the set of criteria $\{c_1, c_2, \dots, c_m\}$ are created. The matrix A created according to these alternatives and criteria is as follows:

$$A = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1m} \\ q_{21} & q_{22} & \cdots & q_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ q_{n1} & q_{n2} & \cdots & q_{nm} \end{bmatrix} \quad (1)$$

Since the weights of the criteria may change depending on whether the current situation persists, the set of criteria weights is as follows:

Wt: $\{wt_1, wt_2, \dots, wt_m\}$ (2)

Considering that the decision in the system is in the wth phase, h different phases in the system are a continuation of each other and the set is expressed as follows:

W: $\{w_1, w_2, \dots, w_h\}$ (3)

Weights may vary for each phase. In order for the method to be applied, it is necessary to ensure that which phases are more likely to occur than others, that the probability of occurrence of the phases is calculable or predictable, and that the criteria weights in each phase are predictable. The set of criterion weights in the kth phase is as follows:

Wt_k: $\{wt_{k1}, wt_{k2}, \dots, wt_{km}\}$ (4)

The transition matrix P, in which the probabilities of the phases are reflected, is given.

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1h} \\ p_{21} & p_{22} & \cdots & p_{2h} \\ \cdots & \cdots & \cdots & \cdots \\ p_{h1} & p_{h2} & \cdots & q_{hh} \end{bmatrix} \quad (5)$$

The equation expressing the whole process, which is formed by assuming that the system is in the kth phase, is as follows:

$$V_{ai} = \sum_{t=1}^{m} q_{it} \sum_{j=1}^{h} wt_{jt} p_{kj} \quad (6)$$

The case study was carried out in an IT firm. As a result of the need arising in the workplace, the company wants to hire a data scientist. There are 3 candidates deemed suitable to work in the company. Alternatives are Candidate 1, Candidate 2, and Candidate 3. There are 3 criteria determined by the company. These are data preparing, programming and coding, and predictive modeling. The study was evaluated on the basis of these criteria and alternatives. Accordingly, the criteria weights given in Table 1 were created.

According to the business, there are certain events that may happen in the near future. These are as follows.

1. Purchase of a new program.
2. The departure of the manager of the department. Since the expectations of the potential new manager will be different, it is thought that it will affect the candidate selection.
3. A new project that will make a difference in the sector is on the agenda.

Since these events that may occur in the near future will affect the selection of candidates, it has been decided that SMCDM is the most suitable method for the problem. The elements of the input vector of the system, which has been prepared considering the conditions above, with the input vector ut =(a, b, c) and the output vector vt=(x, y, z), will be determined according to the following properties.

$$a = \begin{cases} 1, & \text{purchase new program} \\ 0, & \text{otherwise} \end{cases}$$

$$b = \begin{cases} 1, & \text{hiring new manager} \\ 0, & \text{otherwise} \end{cases}$$

$$c = \begin{cases} 1, & \text{starting new project} \\ 0, & \text{otherwise} \end{cases}$$

Depending on the realization of the above events, the input vector will change. The weights of the output vector vt=(x, y, z)=(cost, accessibility to resources, environmental risks) criteria that will arise from this input vector will be different from the weights (0.30, 0.25, 0.45) given in Table 1.

All events that may happen in the near future have been analysed. Then another event that may happen after each event; By calculating new weights, it is desired to reach the target in a way that all layers required by the system follow each other. The Chart of the Concept of Stratification is given in Table 2, showing all possible inputs and the varying outputs obtained as a result of the company managers' foresight. Then, by arranging the data in Table 2, criterion weights were created for the different stages in Table 3.

Thus, weights were calculated using the criteria specified for Candidate 1, Candidate 2, Candidate 3 alternatives in w1, w2, w3, w4, w5, w6, w7, w8 stages.

Since the probability of realization of each stage is different from each other, the probabilities shown in Table 4 were obtained by using the data of the factory managers. Since the probability of occurrence for each stage is taken into account, a more objective result was desired.

The final weights were obtained by using the formula given in equation (6), and the order in Table 5 was found in accordance with these total weights.

III. Results

In this study, the Stratified Multi-Criteria Decision Making (SMCDM) method was investigated. For this, a case study was conducted for candidate selection in an enterprise operating in the IT sector. 3 alternative candidates, 3 main criteria were determined and the problem was solved. As a result of the method, Candidate 3 was selected as the best candidate.

IV. DISCUSSION

Since the structure of traditional Multi-Criteria Decision Making methods does not allow for any additional changes, it causes a waste of time since it is necessary to solve the problem from the beginning. For this reason, SMCDM is used for events that may occur in the near future. As a result of the method, Candidate 3 was selected as the best candidate. This difference shows that the events that may occur in the near future affect the optimal choice considerably. Since the method has just been introduced to the literature, this article is expected to shed light on future studies. It is foreseen that the studies will be carried out in other sectors.

V. CONCLUSION

The difference between traditional methods shows that the events that may happen in the near future affect the optimal choice. Since the method has just been introduced to the literature, this article is expected to shed light on future studies. It is foreseen that the studies will be carried out in other sectors. The main conclusions of the study should be summarized in a short Conclusions section.

VI. Acknowledgment

I would like to thank Assoc. Prof. Dr. Babek Erdebilli for his support in this study.

VII. References

[1] M. I. Henig, & Buchanan, J. T. (1996). Solving MCDM problems: Process concepts. *Journal of Multi-Criteria Decision Analysis*, *5*(1), 3–21, 1996.. https://doi.org/10.1002/(SICI)1099-1360(199603)5:1<3::AID-MCDA85>3.0.CO;2-6

[2] Asadabadi, M. R. (2018). The stratified multi-criteria decision-making method. *Knowledge-Based Systems*, *162*(June), 115–123. https://doi.org/10.1016/j.knosys.2018.07.002

[3] Zadeh, L. A. (2016). Stratification, target set reachability and incremental enlargement principle. *Information Sciences*, *354*, 131–139. https://doi.org/10.1016/j.ins.2016.02.047

[4] Asadabadi, M. R., Saberi, M., & Chang, E. (2017). Logistic informatics modelling using concept of stratification (CST). *IEEE International Conference on Fuzzy Systems*. https://doi.org/10.1109/FUZZ-IEEE.2017.8015510

[5] Rajabi Asadabadi, M., Saberi, M., & Chang, E. (2018). Letter: The concept of stratification and future applications. *Applied Soft Computing Journal*, *66*, 292–296. https://doi.org/10.1016/j.asoc.2018.02.035

Table 1. The weights of the criteria

Criteria	**Weights**
Data preparing	0.3
Programming and coding	0.25
Predictive modeling	0.45

Table 2. Criteria weights in different phases

Criteria/Phases	w1	w2	w3	w4	w5	w6	w7	w8
Data preparing	0.3	0.34	0.18	0.28	0.42	0.13	0.15	0.27
Programming and coding	0.25	0.23	0.41	0.54	0.28	0.23	0.24	0.38
Predictive modeling	0.45	0.43	0.41	0.18	0.3	0.64	0.61	0.35

Table 3. Weights of alternatives in different phases according to criteria

Criteria/Phases	w1	w2	w3	w4	w5	w6	w7	w8
Data preparing	0.23	0.23	0.21	0.17	0.21	0.26	0.25	0.21
Programming and coding	0.22	0.22	0.23	0.22	0.21	0.24	0.24	0.22
Predictive modeling	0.24	0.23	0.23	0.18	0.21	0.27	0.27	0.22

Table 4. Probability of the phases

p11	p12	p13	p14	p15	p16	p17	p18
0.17	0.12	0.07	0.06	0.13	0.08	0.09	0.08

Table 5. Weights of alternatives

Alternatives/Weights	w1	w2	w3	w4	w5	w6	w7	w8	Total	Ranking
Candidate 1	0.039	0.027	0.015	0.010	0.027	0.020	0.023	0.016	0.177	3
Candidate 2	0.038	0.026	0.016	0.013	0.027	0.019	0.021	0.018	0.179	2
Candidate 3	0.040	0.028	0.016	0.011	0.027	0.022	0.024	0.017	0.185	1