International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:5, No:11, 2011

Using the Combined Model of PROMETHEE and Fuzzy Analytic Network Process for Determining Question Weights in Scientific Exams through Data Mining Approach

Hassan Haleh, Amin Ghaffari and Parisa Farahpour

***Abstract*—**Need for an appropriate system of evaluating students' educational developments is a key problem to achieve the predefined educational goals. Intensity of the related papers in the last years; that tries to proof or disproof the necessity and adequacy of the students assessment; is the corroborator of this matter. Some of these studies tried to increase the precision of determining question weights in scientific examinations. But in all of them there has been an attempt to adjust the initial question weights while the accuracy and precision of those initial question weights are still under question. Thus In order to increase the precision of the assessment process of students' educational development, the present study tries to propose a new method for determining the initial question weights by considering the factors of questions like: difficulty, importance and complexity; and implementing a combined method of PROMETHEE and fuzzy analytic network process using a data mining approach to improve the model's inputs. The result of the implemented case study proves the development of performance and precision of the proposed model.

***Keywords*—**Assessing students, Analytic network process, Clustering, Data mining, Fuzzy sets, Multi-criteria decision making, and Preference function.

I. Introduction

EVALUATING educational development is a very important and applicable process in any educational system. Although teachers have paid much attention to this process of transforming qualitative variables to quantitative ones, there has been skepticism about its precision. These kinds of exams are not decisive due to their nature. since humanities are not exact sciences; their results are not decisive by nature. Moreover, a highly qualified evaluation system paves the ground for individual enhancement and ensures fair ranking and scoring for all the students. Thus, assessment systems have to be not only logical and transparent but also easily-implementable on computers. So, it seems that Fuzzy Analytic Network Process (FANP) is suitable for designing such a system due to its capabilities in implementation on computers and its uncertain inputs. In addition, PROMETHEE is a method which has recently attracted much attention because of its mathematical properties and easy application.

Faculty of Mechanical & Industrial Engineering, Islamic Azad University, Qazvin Branch, Qazvin, Iran (Tel: +98-281-367-0051; E-mail: hhaleh@cc.iut.ac.ir,

Data mining approach is also a new interdisciplinary and developing approach which tries to extract hidden knowledge and information from a substantial amount of data. Therefore, it can be used effectively in analyzing the data obtained from the exams with a large number of questions and test-takers. Also, the theory of fuzzy sets since its introduction in 1965 by Professor Lotfi Zadeh [12] has been widely used for problem solving in different science fields. Recently, this theory has been used for educational assessment and ranking.

Biswas presented a fuzzy method for educational assessment in 1995 [5]. In 1999, Chen & Lee also introduced a fuzzy method for solving the problem of ranking students who receive the same score using the Biswas method [6]. For transforming traditional scores to rankings, Echaus & Vachtsevanos (1995) offered a fuzzy logic system [7]. Law (1996) proposed a fuzzy-structured model which was an educational ranking system and aggregated the scores of different exams in order to produce one specific score for each student [8].

Wilson & Karr & Freeman (1998) presented an automatic ranking system based on fuzzy rules and genetic algorithm [9]. Ma & Zhou (2000) presented a fuzzy set approach for investigating the students' learning outcomes using the assessment of their peers and teachers [10]. Wang & Chen (2008) proposed a method for assessing students' answer sheets using fuzzy numbers associated with the degrees of rater reliability [4].

It can be found from the previous research that fuzzy numbers, fuzzy logic systems and fuzzy rules have been used in different educational ranking systems. Weon & Kim (2001) improved an assessment strategy which was based on fuzzy Membership Functions (MF). In this method, they emphasized three important factors of questions in their assessment including questions' difficulty, complexity and importance [11]. Bai & Chen (2008 a) presented a method to automatically construct the ranking MFs of fuzzy rules for assessing students' educational development [2]. Additionally, Bai & Chen (2008 b) applied a method for using fuzzy MFs and fuzzy rules for a similar purpose. To overcome subjectivity problem of the difficulty factor in the Weon and Kim method (2001), they considered difficulty as a function of answer accuracy and the spent time for answering each question [3]. However, their method still suffered from the subjectivity problem since the obtained results were largely dependent on the weights specific to the factors which were determined by the subjective knowledge of domain experts. Ibrahim Saleh & Seong in Kim (2009) proposed an improved alternative of the Bai & Chen method which presented a fuzzy logic system for considering the questions' difficulty, complexity and importance based on the Mamdani's fuzzy inference (Mamdani, 1974) and the Center of Gravity (COG) defuzzification [1].Their method was also defective in application since measuring the answering time for each question is a main problem in the integrated exams and the exams with a large number of questions.In these similarly conducted studies, there has been an attempt to adjust the initial question weights while the accuracy and precision of determining the weights are still under question. Thus, in this paper, a new method is presented for determining question weights by considering the factors of question difficulty, importance and complexity and implementing a combined method of PROMETHEE and fuzzy analytic network process using a data mining approach.

A review of the combined model of promethee and fuzzy analytic network process using data mining approach is shown in fig. 1.

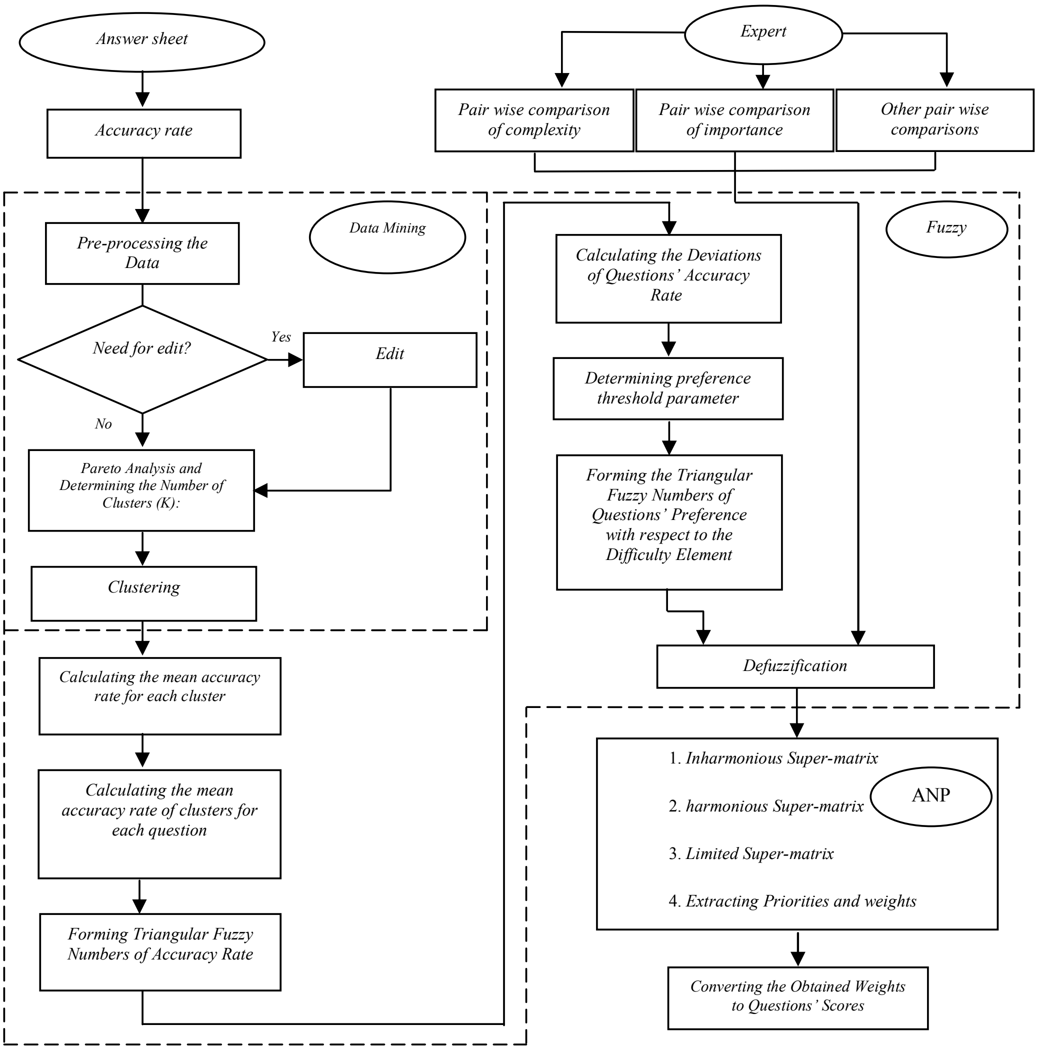

Fig. 1 General View of the proposed model

II. A Case Study Conducted At Qazvin Islamic Azad University

In this case study, the results of the Mechanic Assembling course for mechanical engineering students was investigated. This exam included 11 questions and 91 students took part in

A. Data Mining

1. Pre-processing the data:
2. Pareto analysis and determining the number of clusters (K): The results showed that the number of clusters should be k times as large as 7 in this study.

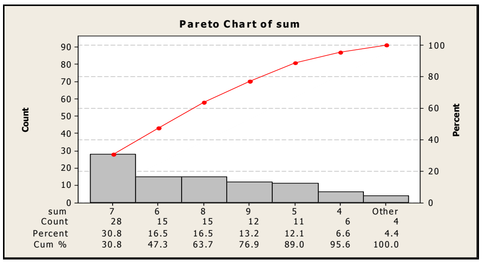

Fig. 2 Pareto analysis for determining the number of clusters

3. Clustering: The K-Means algorithm was used in this paper for clustering.
4. Calculating the Averages and Forming Triangular Fuzzy Numbers of Accuracy rate: The results for this step are mentioned in Table 1. In this table, the component 1-1 indicates the average of accuracy rate for question 1 in cluster 1. The last three rows show triangular fuzzy numbers which are in proportion with each question's accuracy rate (x: A, B, C).

TABLE I
CALCULATING THE AVERAGES AND FORMING TRIANGULAR FUZZY NUMBERS IN PROPORTION WITH EACH QUESTION'S ACCURACY RATE(%)

11	10	9	8	7	6	5	4	3	2	1	Q.N / Cluster
65	57	65	69	78	93	80	26	64	39	58	1
42	65	23	11	11	36	71	80	46	38	26	2
67	61	22	20	49	75	60	77	62	30	65	3
70	68	89	66	86	91	73	88	81	59	61	4
66	57	93	47	65	79	8	58	56	12	28	5
65	43	53	41	54	71	62	73	92	34	53	6
43	40	73	7	0	83	65	83	60	84	60	7
42	40	22	7	0	36	8	26	46	12	26	A
59	56	60	37	49	75	60	69	66	42	50	B
70	68	93	69	86	93	80	88	92	84	65	C

B. Forming Pair wise Comparison Matrices

1. Calculating the Degree of Questions' Preference with respect to Difficulty Element using PROMETHEE.

a. Calculating the Deviations of Questions' Accuracy Rate: In this step, all the deviations of the possible pairs of triangular fuzzy numbers, which were obtained from Section 3.1.4., were calculated using Relation 1.

$$D_i - D_j = (L_i - U_j, M_i - M_j, U_i - L_j) \quad (1)$$

b. Forming the Triangular Fuzzy Numbers of Questions' Preference with respect to the Difficulty: In order to determine the rate of question preference with respect to the difficulty element, the exponential preference function presented in Relation 2 was used taking into account that the accuracy rate of each question is in a reverse relationship with that question's difficulty.

$$P_k(a,b) = \begin{cases} 1 - e\text{^}(-(-d+t)\text{^}2)/(2*s\text{^}2))*9 & d < 0 \\ \\ 1/P_k(b,a) & d < 0 \end{cases} \quad (2)$$

TABLE II
DIFFERENT VALUES OF PARAMETER "T "FOR EACH CORRESPONDING VALUE OF PARAMETER "S"

t	0.024	0.048	0.073	0.097	0.121	0.146	0.170	0.194	0.218
s	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45

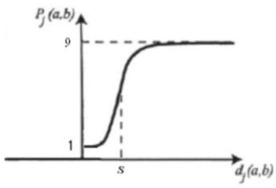

Fig. 3 Exponential Preference Function

c. Defuzzification and Forming Pair wise Comparison Matrix of Question Preference with respect to the Difficulty Element: The Center of Gravity (COG) defuzzification method was used in this paper. Also, to reduce the required calculation time and its errors, the pseudo-codes below were used in MATLAB Software. The final results are presented in Table III.

```
clc
clear all
fuzzy=xlsread('book1.xlsx')
x=0:0.1:10
for i=1:size(fuzzy,1)
    mf=trimf(x,[fuzzy(i,1) fuzzy(i,2) fuzzy(i,3)])
    fuzzy(i,4)=defuzz(x,mf,'centroid')
end
xlswrite('centroid.xlsx',fuzzy
```

"The pseudo-code of the MATLAB Software for defuzzification"

TABLE III
PAIRWISE COMPARISON MATRIX OF THE ELEMENTS OF THE CLUSTER OF QUESTIONS WITH RESPECT TO THE DIFFICULTY ELEMENT

Difficulty	2	3	4	5	6	7	8	9	10	11
1	0.71	1.69	1.74	1.50	1.78	0.96	0.62	1.50	1.30	1.45
2		1.78	1.78	1.73	1.78	1.35	0.80	1.73	1.64	1.71
3			1.14	0.77	1.45	0.59	0.56	0.77	0.67	0.74
4				0.69	1.30	0.57	0.56	0.69	0.62	0.67
5					1.67	0.65	0.56	1	0.84	0.96
6						0.56	0.56	0.60	0.57	0.59
7							0.63	1.54	1.35	1.50
8								1.77	1.74	1.77
9									0.84	0.96
10										1.14

2. Forming the Pair wise Comparison Matrix of Questions' Preference with respect to the Elements of Complexity and Importance: In this section, fuzzy terms are used for the pair wise comparison of the elements. Thus, the expert views are turned into triangular fuzzy numbers considering the fuzzy terms' definition.

TABLE IV
CRISP VALUES CORRESPONDING TO FUZZY TERMS

Crisp values	fuzzy terms
9	Extremely
7	Very Strongly
5	Strongly
3	Moderately
1	Equally

TABLE V
PAIR WISE COMPARISON MATRICES OF THE ELEMENTS OF THE CLUSTER OF QUESTIONS WITH RESPECT TO THE ELEMENTS OF COMPLEXITY AND IMPORTANCE

Complexity	2	3	4	5	6	7	8	9	10	11
1	7	5	5	5	5	7	3	5	1	1
2		1/3	1	1	1/3	1/3	1/5	1/3	1/7	1/9
3			3	3	1	1	1/3	1	1/5	1/7
4				1	1/3	1/3	1/5	1/5	1/7	1/9
5					1/3	1/3	1/3	1/3	1/7	1/7
6						1	1	1	1/5	1/5
7							1	1	1/7	1/7
8								1	1/5	1/7
9									1/5	1/5
10										1/3

Importance	2	3	4	5	6	7	8	9	10	11
1	7	7	7	7	7	7	7	7	3	1/3
2		1/3	1	1	1/3	1/3	1/3	1/3	1/5	1/7
3			3	3	1	1	1	1	1/3	1/5
4				1	1/3	1/3	1/3	1/3	1/5	1/7
5					1/3	1/3	1/3	1/3	1/5	1/7
6						1	1	1	1/3	1/5
7							1	1	1/5	1/5
8								1	1/5	1/7
9									1/3	1/5
10										1/3

3. Forming Other Pair wise Comparison Matrices: To form these matrices, we follow the way mentioned in previous section.

C. Analytic Network Process (ANP)

All the calculations of this section were conducted using Super Decisions software which was confirmed by Mr. Thomas L. Saati. A

1. Inharmonious Super-matrix
2. Cluster-matrix
3. Harmonious Super-matrix
4. Limited Super-matrix
5. Extracting Priorities from a Super-matrix

D. Converting the Obtained Weights to Questions' Scores

In this step, final calculations for determining the questions' score weights were conducted. Relation 3 was used for this purpose in which the amount of parameter S was 20. The final results can be observed in Table 6.

$$g_i = w_i \times (S / \sum_{i=1}^{m} w_i) \quad (3)$$

TABLE VI
CONVERTING WEIGHTS TO THE FINAL SCORES OF QUESTIONS

Question Number(i)	Normalized Weight (Wi)	Final Score(Gi)
1	*0.14926*	*2.9852*
2	*0.07107*	*1.4214*
3	*0.06212*	*1.2424*
4	*0.06853*	*1.3706*
5	*0.06028*	*1.2056*
6	*0.07278*	*1.4556*
7	*0.07525*	*1.505*
8	*0.10121*	*2.0242*
9	*0.07088*	*1.4176*
10	*0.11732*	*2.3464*
11	*0.1513*	*3.026*

III. CONCLUSION

Evaluating educational development is a highly important and applicable process in any educational system. In this process, the process of changing qualitative variables to quantitative ones has always attracted teachers' attention; however, its precision has been skeptical since many years ago. In this paper, a combined method is presented for determining the weight of scientific exam questions. In this model, a part of PROMETHEE method, called preference function, formed the inputs of one of the pair wise comparison matrices, which was used in the fuzzy analytic network process. In data preparation step, data mining approach was used before PROMETHEE. Compared with the previously conducted researches, this method has the following advantages: first, the proposed method is the method of constructing the initial questions weights while previous methods only tried to adjust initial questions weights presented by teachers using similar criteria. Considering that the goal of the conducted studies was to remove human evaluation errors, using the initial score weights as the main basis does not make sense. Second, the analytic network process gives acceptable results in spite of all the current dependency in the network. In addition, preference function provides an opportunity to use the real information extracted from the answer sheets; these data are real; so they are more reliable. Moreover, data mining approach leads to the reduction in the influence of noisy and incomplete data in the conducted studies and, as a result, improves the performance and precision of the method in the exams with a large number of test-takers and questions. By investigating the previous works in the field of evaluating educational development and also the combined methods of multi-criteria decision making show that the proposed method is unique due to integrating four different fields of data mining, analytic network process, fuzzy sets and preference function.

REFERENCES

[1] Ibrahim saleh, Seong – in Kim (2008). A fuzzy system for evaluating students' learning achievement. Expert Systems with Applications, 36 (2009) 6236 – 6243
[2] Bai, S.-M., & Chen, S.-M. (2008a). automatically constructing grade membership functions of fuzzy rules for students' evaluation. Expert Systems with Applications, 35(3), 1408–1414

[3] Bai, S.-M., & Chen, S.-M. (2008b). Evaluating students' learning achievement using fuzzy membership functions and fuzzy rules. Expert Systems with Applications, 34, 399–410.
[4] Wang, H. Y., & Chen, S. M. (2008). Evaluating students' answer scripts using fuzzy numbers associated with degrees of confidence. IEEE Transactions on Fuzzy Systems, 16(2), 403–415.
[5] Biswas, R. (1995). An application of fuzzy sets in students' evaluation. Fuzzy Sets and Systems, 74(2), 187–194.
[6] Chen, S. M., & Lee, C. H. (1999). New methods for students' evaluating using fuzzy sets. Fuzzy Sets and Systems, 104(2), 209–218.
[7] Echauz, J. R., & Vachtsevanos, G. J. (1995). Fuzzy grading system. IEEE Transactions on Education, 38(2), 158–165.
[8] Law, C. K. (1996). Using fuzzy numbers in education grading system. Fuzzy Sets and Systems, 83(3), 311–323.
[9] Wilson, E., Karr, C. L., & Freeman, L. M. (1998). Flexible, adaptive, automatic fuzzy-based grade assigning system. In Proceedings of the North American fuzzy information processing society conference (pp. 334–338).
[10] Ma, J., & Zhou, D. (2000). Fuzzy set approach to the assessment of student-centered learning. IEEE Transactions on Education, 43(2), 237–241.
[11] Weon, S., & Kim, J. (2001). Learning achievement evaluation strategy using fuzzy membership function. In Proceedings of the 31st ASEE/IEEE frontiers in education conference, Reno, NV (Vol. 1, pp. 19–24).
[12] Weon, S., & Kim, J. (2001). Learning achievement evaluation strategy using fuzzy membership function. In Proceedings of the 31st ASEE/IEEE frontiers in education conference, Reno, NV (Vol. 1, pp. 19–24).
[13] Zadeh, L. A. (1965). Fuzzy sets. Information and Control, 8, 338–353.

H. Haleh he has born in IRAN, TEHRAN, 1968. He has B.Sc. degree in Industrial Engineering from Sharif University of Technology, TEHRAN, IRAN, 1991. Also he has M.Sc. degree in same area and same University, 1994. He has received his Ph.D. degree in Automation and Production Engineering from Gunma University, KIRYU, GUNMA, JAPAN, 2004.He is now dean of Mechanical and Industrial Engineering faculty in I.Azad University, QAZVIN branch. He is also an assistant professor in System and Industrial Engineering faculty, Isfahan University of Technology, Isfahan, IRAN. His major field is Automation, Robotics and Decision Support Systems.Dr. HALEH has published more than 10 journal papers and 10 conference papers in last 2 years.

A. Ghaffar he has born in IRAN, TEHRAN, 1982. He has B.Sc. degree in Industrial Engineering from Mazandaran University of Science and Technology (MUST), IRAN, BABOL, 2007. Also he has M.Sc. degree in Industrial Engineering from I.Azad University, QAZVIN branch, IRAN, QAZVIN, 2011. His major field is Multi Criteria Decision Making (MCDM), Fuzzy theory and he has experiences in SCM and Quality systems.