Selection of Most Suitable Secondary School Alternative by Multi-Criteria Fuzzy Analytic Hierarchy Process

N.S.M. Rezaur Rahman[1], Md. Abdul Ahad Chowdhury[2], Abdullah-Al Nahian Siraj[3], Rashedur M. Rahman[4], Rezaul Karim[5], K. M. Arafat Alam[6]

[1,2,3,4,5,6] Department of Electrical and Computer Engineering,
North South University,
Plot-15, Block-B, Bashundhara Residential Area, Dhaka, Bangladesh.

[1]shaon13nov@gmail.com, [2]ahad.chowdhury@northsouth.edu,
[3]nahian.siraj@northsouth.edu,
[4]rashedur.rahman@northsouth.edu,
[5]karim.rezaul@northsouth.edu, [6]arafatalam42@gmail.com

Abstract. Selecting the most suitable secondary school for a student is a multi-criteria decision problem. In this paper we propose a system that uses fuzzy analytic hierarchy process (AHP) to define and evaluate linguistic variables that represent these criteria. This system takes some of them as crisp inputs and a set of fuzzy linguistic values from the user that denote relative importance measures, as deemed by the user, for every possible pair of these factors. Then it uses Fuzzy AHP to calculate absolute weights for each factor and comparative scores for a few user-selected schools for each criterion. These values are used to assign a final score to each school, which is the measurement of its suitability for the given input.

Keywords: Fuzzy logic, Analytic Hierarchy Process, Facility Location Selection, Multi-Attribute Decision Making

1 Introduction

A student's future prospects for professional success can highly depend on the selection of educational institute for him. In Bangladesh, such decisions for millions of students are made every year by their parents and guardians. Unfortunately, the outcome of this decision-making in the conventional way often does not happen to be very suitable. In this paper, we propose fuzzy analytic hierarchy process (Fuzzy AHP) for helping the parents or the guardians to select the most suitable school for his child. Fuzzy AHP was originally introduced by Saaty [1], which combined analytic hierarchy process with fuzzy logic for solving multi-attribute decision making problems. In our method, we take fuzzy linguistic inputs for relative importance of each pair of criteria and various information of a student. We then use these values to compute suitability scores of the schools and suggest the best school(s).

2 Related Works

Application of Fuzzy AHP for multi-attribute decision making is very common for facility location selection problems. Boltürk, et al. [2] proposed a new Hesitant Fuzzy AHP where both triangular and trapezoidal fuzzy numbers were used as input through Hesitant Fuzzy AHP and comparative scores for the alternatives were calculated.

Singh [3] proposed Extent Fuzzy AHP to determine facility location, where a square matrix of triangular fuzzy numbers was used to determine relative importance of the criteria through Fuzzy AHP. Another similar approach was proposed by Chou et al. [4] that integrated Fuzzy Set Theory with Factor Rating System and Simple Additive Weighting to evaluate facility location alternatives for a supply chain management problem.

3 Methodology

We present the description of our methodology in the following five sections.

3.1 Data Collection and Pre-processing

We obtained all of our data from Bangladesh Open Data [5], the government-maintained portal for datasets on national statistics. We combined multiple datasets into a single dataset for our purpose. Table 1 depicts our dataset and its different attributes.

Table 1. Dataset and Selected Attributes

Datasets	Selected Fields
Institutional Information of Bangladesh	EIIN, Name of institution, Division, District, Thana, Mauza, Post, Institute type,
Institute-wise Teacher Student Info	EIIN, Number of total teachers, Number of female teachers, Number of total students, Number of female students
Students Information of Bangladesh	EIIN, Number of students by age in classes 6–10
Information Related Socio-Economic Background of Parents/Guardian of Secondary Level Students	EIIN, Number of students categorized by parent's or guardian's profession in classes 6–10
Grade Wise Student in School	EIIN, Number of male and female students in classes 6–10

3.2 User Inputs

We want to ensure that our system always gives the user a personalized output. For this purpose, we take student's age, sex, intended class (grade) for admission, guardian's living division, district, thana and union (smaller sub regions of a district) as input from the user. We also take linguistic input denoting relative importance for each possible pair of the attributes of a school: student-teacher ratio (TSR), socio-economic status (SES), ratio of male or female students in the school (MFR), age of school (AS), distance from the user's home (DIST) and age gap of the student from the average age of students in the specific class (ADS).

The relative importance of these attributes, which are comparative to each other, is taken by giving the user a set of linguistic values as options to choose from. These linguistic values and their corresponding triangular fuzzy number (TFN) values are given in table 2 in $(l \quad m \quad u)$ format.

Table 2. Linguistic Values and Corresponding Triangular Fuzzy Numbers

Linguistic Values	Triangular Fuzzy Numbers
Equally Important	$(1 \quad 1 \quad 1)$
Moderately Important	$(2 \quad 3 \quad 4)$
Fairly Important	$(4 \quad 5 \quad 6)$
Very Important	$(6 \quad 7 \quad 8)$
Absolutely Important	$(9 \quad 9 \quad 9)$

3.3 Criteria Weight Calculation

First, we convert the linguistic values into TFNs according to the scale in table 2 and construct a comparison matrix with them. For example, if criterion A is "Moderately Important" compared to criterion B, then the comparative importance value of criterion A with respect to criterion B in the matrix will be $(2 \quad 3 \quad 4)$, and the value of criterion B against criterion A will be $(1/4 \quad 1/3 \quad 1/2)$ in the matrix. The relation can be shown in mathematical terms like equation (1):

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ \vdots & & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \tag{1}$$

Here, C is the comparison matrix and c_{ij} is the comparative importance value of i-th criterion with respect to j-th criterion. We take geometric mean of TFNs for each row of this matrix. According to Buckley [6], this is given by equation (2):

$$r_i = \left(\prod_{j=1}^{n} c_{ij}\right)^{\frac{1}{n}} \tag{2}$$

Here, each r_i value is a TFN, denoted by TFN $(l_i \quad m_i \quad u_i)$. Let us assume the new set of values for each row form a vector R is given in equation (3)

$$R = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} \quad (3)$$

Then we calculate a vector V by taking the vector summation of TFNs of R, taking the inverse of each member of the sum and sorting the members in ascending order. Since $l \leq m \leq u$ for any TFN $(l \quad m \quad u)$, this process can be described in mathematical terms like equation (4):

$$V = \begin{bmatrix} \frac{1}{U} & \frac{1}{M} & \frac{1}{L} \end{bmatrix} \quad (4)$$

Here $L = \sum_{i=1}^{n} l_i$, $M = \sum_{i=1}^{n} m_i$ and $U = \sum_{i=1}^{n} u_i$ for each r_i from R.

Now we multiply every member of matrix R individually with V. This gives us the fuzzy weight matrix W as given in equation (5)

$$W = \begin{bmatrix} \frac{l_1}{U} & \frac{m_1}{M} & \frac{u_1}{L} \\ & \vdots & \\ \frac{l_n}{U} & \frac{m_n}{M} & \frac{u_n}{L} \end{bmatrix} \quad (5)$$

These fuzzy weight values for each criterion is defuzzified using the Center of Area (COA) method. Since these are all triangular fuzzy numbers, this can be calculated by taking the arithmetic mean of l, m, u as given in equation (6)

$$v_i = \frac{l_i + m_i + u_i}{3} \quad (6)$$

These crisp values are normalized by their sum as given by equation (7):

$$w_i = \frac{v_i}{\sum_{p=1}^{n} v_p} \quad (7)$$

Therefore, by applying this process on matrix C, we get $w_1, w_2, \ldots w_n$ for n criteria.

After this, the user selects m alternatives from the list of schools. These schools are compared against each other based on each criterion, which produces an m-by-m comparison matrix for each criterion. For this, we calculate a "score" for each criterion.

3.4 Criteria-wise Score Calculation for Alternatives

To calculate score for some of the attributes, we apply cumulative distribution function (CDF) of standard normal distribution, in order to obtain a value between 0 and 1. The formula we use for some of these values is shown in equation (8). Here, μ is the mean and σ is the standard deviation. For calculating mean and standard deviation, we consider the data of all the schools available in our dataset.

$$cdf(x, \mu, \sigma) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right] \quad (8)$$

In this equation, $\operatorname{erf}(x)$ means the error function, denoted by equation (9):

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{9}$$

To calculate the score for “student sex ratio”, we apply the following equation:

$$score_{mfr} = \begin{cases} \frac{\#ms}{\#st} & if\ sex = 'male' \\ \frac{\#fs}{\#st} & if\ sex = 'female' \end{cases} \tag{10}$$

Here, $score_{mfr}$ means number of male or number of female in ratio depending on the user’s input sex, $\#ms$ means the number of male students, $\#fs$ means the number of female students and $\#st$ number of total students.

For criterion “teacher-student ratio” we apply formula described by equation (11).

$$score_{tsr} = 1 - cdf(tsr, \mu_{TSR}, \sigma_{TSR}) \tag{11}$$

Here, tsr means ratio of number of students and number of teachers in the school, $score_{tsr}$ means teacher-student ratio score, μ_{TSR} denotes the mean of all the TSR values and σ_{TSR} denotes the standard deviation in TSR values.

For criterion “average age of students” we take the absolute value of the difference between the user’s given age and the average age of students in the class, ads. We normalize it using equation (8) and obtain the score for this criterion, $score_{ads}$ by equation (12). Here, μ_{ADS} and σ_{ADS} denote the mean and the standard deviation of all the ADS values.

$$score_{ads} = 1 - cdf(ads, \mu_{ADS}, \sigma_{ADS}) \tag{12}$$

For criterion “age of school” we take the age of school, as, from the year of establishment of the school given in the dataset. We normalize it using equation (8) and calculate the score for tis criterion, $score_{as}$.

$$score_{as} = cdf(as, \mu_{AS}, \sigma_{AS}) \tag{13}$$

For the criterion “socio-economic status” we have conducted a survey on 15 participants. They have assigned 15 professions and 7 administrative levels a number between 1 and 10 (inclusive), based on their perception of social status. The data for the profession score obtained from the survey is rounded to the nearest multiple of 2.5, to represent 4 distinct social classes. The data from the survey is shown in tables 3 and 4.

Table 3. Professions and Their Status

Professions	Profession Status, p_j
Lawyer, Doctor, Engineer, Businessman	10.0
Government job, Private job, Teacher	7.5
Expatriate, Cultivation, Small business owner	5.0
Worker, Fisherman, Weaver, Potter, Blacksmith	2.5

Table 4. Administrative Levels and Their Status

Administrative Level	Area Status, A
Metropolitan	10.0
City Corporation	9.0
District Sadar Municipality	7.0
Upazilla Sadar Municipality	5.0
Other Municipality Area	4.0
Upazilla Sadar but not Municipality	2.0
Rural Area	1.0

In our dataset the number of students whose guardians are working in these 15 professions and the level of the areas in which the schools are located are given. For each selected school, we calculate the socio-economic score by the equation (14):

$$S = \left(\frac{1}{N}\sum_{j=1}^{15} p_j n_j\right) \times W_P + A \times W_A \tag{14}$$

Here N is the total number of students in the school, p_j is the socio-economic status score of the j-th profession, n_j is the number of students in the school whose parents or guardians work in the j-th profession. A is the area-based status of the school's location. $W_P = 10$ and $W_A = 5$, which are the weights that guardian's occupation and area status carry. These values were also obtained from the survey. We normalize this value according to equation (8), shown in equation (15).

$$score_{ses} = cdf(S, \mu_S, \sigma_S) \tag{15}$$

Finally, for the criterion "distance", at first we assign a crisp value $d = 10$ for each school. If the school is in the same division as the user's inputted division, we subtract 4 from d. Similarly, if the school is in the same district, thana or union/ward, we subtract 3, 2 and 1 respectively. This value is used to calculate the score for this criterion, as shown in equation (16).

$$score_{dist} = 1 - \frac{d}{10} \tag{16}$$

3.5 Criteria-wise Comparison Matrix Formation for Alternatives and Final Calculation

After gathering all these scores, we construct a comparison matrix for each criterion. At first, for comparing i-th alternative to j-th alternative on n-th criterion, we assign an integer, $cs_{n_{ij}}$, between -10 and 10 (inclusive) according to equation (17), given that $score_{n_i}$ and $score_{n_j}$ are the score values of i-th alternative and j-th alternative on n-th criterion.

$$cs_{n_{ij}} = \begin{cases} \lceil \left| score_{n_i} - score_{n_j} \right| \rceil & if\ score_{n_i} \geq score_{n_j} \\ \lfloor \left| score_{n_i} - score_{n_j} \right| \rfloor & else \end{cases} \tag{17}$$

Finally, for each $cs_{n_{ij}}$ value we take a TFN from table 5 and assign this value in the i-th row and j-th column in the comparison matrix for n-th criterion. Thus, for n criteria we obtain n matrices C_1, C_2, … C_n that are m-by-m in dimension, in which a member a_{ijk} represents the comparison TFN of j-th school with the k-th school based on the i-th criterion, shown in equation (18).

$$C_n = \begin{bmatrix} a_{n11} & a_{n12} & \cdots & a_{n1m} \\ \vdots & & \ddots & \vdots \\ a_{nm1} & a_{nm2} & \cdots & a_{nmm} \end{bmatrix} \tag{18}$$

Table 5. Possible Crisp Scores ($cs_{n_{ij}}$) and Corresponding TFNs

Negative Crisp Scores	Corresponding TFNs	Positive Crisp Scores	Corresponding TFNs
-10	(1/9 1/9 1/9)	0	(1 1 1)
-9	(1/9 1/9 1/9)	1	(1 1 1)
-8	(1/9 1/8 1/7)	2	(1 2 3)
-7	(1/8 1/7 1/6)	3	(2 3 4)
-6	(1/7 1/6 1/5)	4	(3 4 5)
-5	(1/6 1/5 1/4)	5	(4 5 6)
-4	(1/5 1/4 1/3)	6	(5 6 7)
-3	(1/4 1/3 1/2)	7	(6 7 8)
-2	(1/3 1/2 1)	8	(7 8 9)
-1	(1 1 1)	9	(9 9 9)
0	(1 1 1)	10	(9 9 9)

For each of these comparison matrices, Fuzzy AHP is applied using equations (2) through (7). Applying this process of matrices C_1, C_2, … C_n yields sets of alternative-wise scores for all the criteria, $\{v_{11}, v_{21}, \dots v_{m1}\}$, $\{v_{12}, v_{22}, \dots v_{m2}\}$, … $\{v_{1n}, v_{2n}, \dots v_{mn}\}$. Now we calculate the final scores for each of these m schools, based on the n criteria. We obtain this final score Z_i for the i-th school by

$$Z_i = \sum_{p=1}^{n} w_p v_{ip} \tag{19}$$

Here, w_p is the absolute weight of p-th criterion that had been calculated earlier.

4 Demonstration and Result Analysis

4.1 Demonstration of Calculation

Let us go through an example to demonstrate how this process works. We assume that a user has given the crisp and fuzzy inputs shown in tables 6 and 7. The fuzzy inputs are mapped with TFNs from table 2 to generate a comparison matrix, shown in table 8.

Table 6. Crisp Inputs

Fields	Inputs	Fields	Inputs
Class	7	Division	RAJSHAHI
Student's age	15	District	NAWABGANJ
Student's sex	Female	Thana	CHAPAI NAWABGANJ SADAR
		Union/Ward	WARD NO-03

Table 7. Fuzzy Inputs

Ab. Imp	Ve. Imp	Fa. Imp	Mo. Imp	Crit.	Eq. Imp	Crit.	Mo. Imp	Fa. Imp	Ve. Imp	Ab. Imp
				TSR		SES	√			
		√		TSR		MFR				
			√	TSR		AS				
				TSR		DIST			√	
			√	TSR		ADS				
	√			SES		MFR				
			√	SES		AS				
				SES	√	DIST				
	√			SES		ADS				
				MFR	√	AS				
				MFR		DIST		√		
				MFR	√	ADS				
				AS		DIST		√		
				AS	√	ADS				
	√			DIST		ADS				

We apply AHP on this comparison matrix, which is described in equations (2) to (7). This gives us crisp weights of each criterion.

Now we assume that the user selects 3 schools. The details of these schools are shown in table 8. For each school, its values are used to obtain scores for each criterion using equations (10) to (17). These scores are shown in table 9, which are used to construct comparison matrices for each criterion. Each matrix is inputted through Fuzzy AHP, as described in equations (2) to (7). They generate crisp values for each criterion and for each alternative. These values, along with crisp weights of each criterion, are used as described in equation (19) to calculate final score for each alternative. This has been shown in table 11.

Table 8. Comparison Matrix

	TSR	SES	MFR	AS	DIST	ADS
TSR	(1 1 1)	(1/2 1/3 1/4)	(4 5 6)	(2 3 4)	(1/6 1/7 1/8)	(2 3 4)
SES	(4 3 2)	(1 1 1)	(6 7 8)	(2 3 4)	(1 1 1)	(6 7 8)
MFR	(1/6 1/5 1/4)	(1/8 1/7 1/6)	(1 1 1)	(1 1 1)	(1/4 1/5 1/6)	(1 1 1)
AS	(1/4 1/3 1/2)	(1/4 1/3 1/2)	(1 1 1)	(1 1 1)	(1/4 1/5 1/6)	(1 1 1)
DIST	(8 7 6)	(1 1 1)	(6 5 4)	(6 5 4)	(1 1 1)	(6 7 8)
ADS	(1/4 1/3 1/2)	(1/8 1/7 1/6)	(1 1 1)	(1 1 1)	(1/8 1/7 1/6)	(1 1 1)

Table 9. Details of User-selected School Alternatives

	TSR	SES	MFR	AS	DIST	ADS
A1	128.58	82.99	50.94	91.00	1.00	2.53
A2	38.83	133.23	35.61	171.00	6.00	3.20
A3	29.75	119.87	22.69	27.00	6.00	2.97

Table 10. Scores of User-selected School Alternatives

	TSR	SES	MFR	AS	DIST	ADS
A1	0.00	0.98	0.51	0.46	0.90	0.73
A2	0.46	1.00	0.36	0.46	0.40	0.39
A3	0.62	1.00	0.23	0.45	0.40	0.51

Table 11. Alternative-based Scores for Each Criterion and Final Score of Each Alternative

Criteria	Weights	A1	A2	A3
TSR	0.13	0.09	0.42	0.48
SES	0.32	0.33	0.33	0.33
MFR	0.05	0.40	0.32	0.27
AS	0.06	0.33	0.33	0.33
DIST	0.38	0.71	0.14	0.14
ADS	0.05	0.54	0.21	0.25
TOTAL		0.46	0.27	0.27

Therefore, the first alternative (A1) is the best choice for the student.

4.2 Analysis

From table 7 we see that the user has thought “distance” is one of the two highest-priority criteria along with “socio-economic status”. For this, in table 10, we see that criterion “distance” has the highest weight, and the alternative that has the highest score in “distance” criterion ultimately comes to be the best alternative. As “socio-economic status” is also among the top priority of the user, so it is chosen as the more important one in every comparison, except when comparing with “distance”. This is why it comes a close-second in terms of criterion-weight. Then the user ranks the criteria “teacher-student ratio”. So, this criterion comes third in terms of criterion-weight. But here we see the effect of linguistic terms and corresponding fuzzy numbers. The user has thought “teacher-student ratio” is important, but only “moderately” or “fairly” important, not “very” or “absolutely” important. For this reason, this criterion comes in third, but in a distant third. This affects the result, as we see that even though A1 has the lowest score in this criterion, it still has gained the highest final score. The remaining three criteria have very poor relative importance and therefore have less impact on the final score.

5 Conclusion

In context of Bangladesh, selection of secondary school is an important decision to make for any guardian. In this paper we propose to use Fuzzy AHP to assist the decision-maker. In our system we take into account most of the common criteria usually considered for choosing a school. Our target is to provide the user with various information about the schools, but at the same time give enough options to provide his or her personal preferences. This method can be used to solve many other problems of facility selection and multi-attribute decision-making as well.

References

1. T. L. Saaty, "The Analytic Hierarchy Process," McGraw-Hill, New York, 1980.
2. E. Boltürk, S. C. Onar, B. ÖztayşI, C. Kaharman, and K. Göztepe, “Multi-Attribute Warehouse Location Selection in Humanitarian Logistics Using Hesitant Fuzzy AHP,” International Advanced Research Journal in Science, Engineering and Technology. Vol. 3, Issue 2, 2016.
3. R. K. Singh, "Facility Location Selection Using Extent Fuzzy AHP," International Advanced Research Journal in Science, Engineering and Technology. Vol. 3, Issue 2, 2016.
4. S. Chou, Y. Chang, and C. Shen, “A Fuzzy Simple Additive Weighting System Under Group Decision-Making for Facility Location Selection with Objective/Subjective Attributes,” European Journal of Operational Research 189, 2008, pp. 132–145
5. Bangladesh Open Data, (n. d.), retrieved from http://data.gov.bd/dataset.
6. J. J. Buckley, “Fuzzy Hierarchical Analysis,” Fuzzy Sets Systems, Vol.17 (1), 1985, pp. 233–247.