

Applying the methods of system analysis to teaching assistants' evaluation

E. Beresneva <eberesneva@hse.ru>

M. Gordenko <mgordenko@hse.ru>

Department of Software Engineering,

National Research University Higher School of Economics, 20, Myasnitckaya st.,
Moscow, 101000 Russia

Abstract. This article presents the results of applying various methods of system analysis (CATWOE, Rich Picture, AHP, Fuzzy AHP) to evaluation of teaching assistants. The soft and hard methods were applied. Methods of system analysis are considered as an example at the Higher School of Economics (HSE) in program "Teaching assistant". The article shows the process of interaction of teaching assistants with students and faculty in the form of Rich Picture. Selection and analysis of criteria for the evaluation of training assistants are carried out. Three groups of criteria were defined: professional skills, communicating skills, personal qualities. Each group has some subcriteria, which were defined in brainstorm. Its own method was determined, which immediately allow drop some assistants. In addition, the application of the methods AHP and Fuzzy AHP type-2 to evaluate teaching assistants is considered. The strengths and weaknesses of each method are revealed. It is also shown that, despite the power of the methods of system analysis, it is necessary to use common sense and logic. Do not rely only on the numbers obtained by the methods of system analysis. In the process of work, the best teaching assistant is selected, and the group of the best teaching assistants is defined.

Keywords: system analysis; combination of soft and hard methods; multicriteria decision making (MCDM); AHP; type-2 fuzzy sets; Fuzzy AHP

DOI: 10.15514/ISPRAS-2018-30(3)-18

For citation: Beresneva E., Gordenko M. Applying the methods of system analysis to teaching assistants' evaluation. *Trudy ISP RAN/Proc. ISP RAS*, vol. 30, issue 3, 2018. pp. 251-270. DOI: 10.15514/ISPRAS-2018-30(3)-18

1. Introduction

At the Higher School of Economics (HSE) there is a program "Teaching assistant" which has been effective for several years. Each teacher can invite an education assistant, who will take some of the routine tasks related to teaching the course (checking homework, developing test materials, etc.).

Every student or a graduate student of the HSE, who meets the criteria established by the faculty, can be a teaching assistant. The teacher (or group of teachers) formulates

tasks for the teaching assistants and monitors the quality of their performance. The teacher is responsible for the results of the students' knowledge, the quality of materials prepared by the education assistant, methodical support of the teaching assistant's work.

At the moment, all faculties establish their own criteria for selecting teaching assistants independently. Now there is only one criterion for all disciplines: "A student must have a mark at least 8 on the course in which he/she is involved, or he/she must have a recommendation from the department, to which teaching of this discipline is fixed." However, the practice shows that it is not enough to have only this criterion. There were no special studies about it before, but annual evidence showed that an excellent mark does not fully correlate with being a good teaching assistant. Recent year revealed that 60% of assistants were not able to cope with their work according to teachers. Most problems were connected with personal qualities, professional and communicative skills. For example, somebody did all the tasks slowly and did not do everything in time, or just did not have enough knowledge in the subject area. There were even some facts of disclosure of confidential information: one teaching assistant shared answers to the tests with students. Thus, there is a strong necessity to define a group of selective factors in a clever manner.

Recently, the head of Computer Science faculty has ordered each teacher (or group of teachers) on all disciplines to choose the best teaching assistant to give him/her an incentive award. In addition, next year the number of students is reduced, and it is necessary to decrease the number of assistants. Now there is a tendency on «Discrete mathematics» course that the education assistants who come from year to year are the same. This situation prompted the idea that at the moment when assessing teaching assistants, it is worth using additional criteria that will allow the group of teachers to select the best assistant and choose the group of the most successful assistants.

Thus, two tasks are faced – to choose the best assistant on «Discrete mathematics» course and to select the group of the most successful assistants, with whom it is possible to continue working on this course.

The purpose of this work is the development of searching method, which will select the best assistant and select the group of the most successful ones according to the criteria set by the group of teachers.

The rest of the paper is organized as follows. We discuss the problem specification in Section 2 and introduce our premises for model, which we use to illustrate our main results on Section 9. Sections 3, 4 and 6 present the different methods used for solution the problem. In sections 5 and 7 the derivations for the AHP and Fuzzy AHP are discussed. Section 8 presents a sensitivity analyze.

2. The Difference between Previous Works and Our Approach

The literature review shows that there are a lot of researches that reveal a high success of applying the teaching assistant program in general. The most recent one is [3].

However, no one article is aimed neither at selection criteria for teaching assistants nor at searching methodology.

The closest study to our problem is devoted to a proposed framework for evaluating student's performance [4]. This work is based on the hard approach only. It uses the variation of the most widely used approach for multi-criteria decision-making – Analytic Hierarchy Process that combines mathematics and expert judgment. Since Analytic Hierarchy Process suffers from the problem of imprecision and subjectivity, their paper proposes to use Fuzzy AHP instead of traditional method.

However, there is an opinion about useless of applying Fuzzy AHP method. In [3] it is said that “the numerical representation of judgments in the AHP is already fuzzy” and “making fuzzy judgments more fuzzy does not lead to a better more valid outcome and it often leads to a worse one.”

Our article proves that Fuzzy AHP with type-2 modification can still be used in a decision making process. Moreover, our study combines both hard and soft approaches because this problem consists of not only main criteria but also it has a lot of additional ones. And these auxiliary factors can not be described using only formal algorithms.

3. Problem Definitions

The problem of finding the best teaching assistant and the group of teaching assistants is closely related with searching the criteria by which the teaching assistants should be selected.

To analyze the domain and determine its boundaries, the rich picture can be applied. Rich Picture is a collection of sketches, pictures, photos, symbols, signatures which represent a particular situation or a question of the real world from the point of view of the person or group of people who create it. Image components are people (stakeholders), systems, processes, inter-faces, data streams, information sources, infrastructure objects, attendant and impeding factors, emotions, points of view and attitude to them, etc.

Rich Picture can reflect the interaction and connections of the system components (or the surrounding world), their influence, cause and effect. It can also represent such subjective elements as attitude (perception), point of view, prejudice [1].

It is used to explore and aggregate the physical, conceptual and emotional aspects of the actual situation (sys-tem/problem/need).

Rich picture on subject «Teaching assistants» interactions in discipline «Discrete mathematic» is provided in Fig. 1.

To analyze the subject area and project boundaries, the CATWOE technique is a good addition to Rich Pictures.

CATWOE is defined by Peter Checkland as a part of his Soft Systems Methodology (SSM). It is a simple checklist for thinking. CATWOE is an acronym, each letter

stands for a specific word: Clients, Actors, Transformation, World view, Owner, Environmental constraints [2].

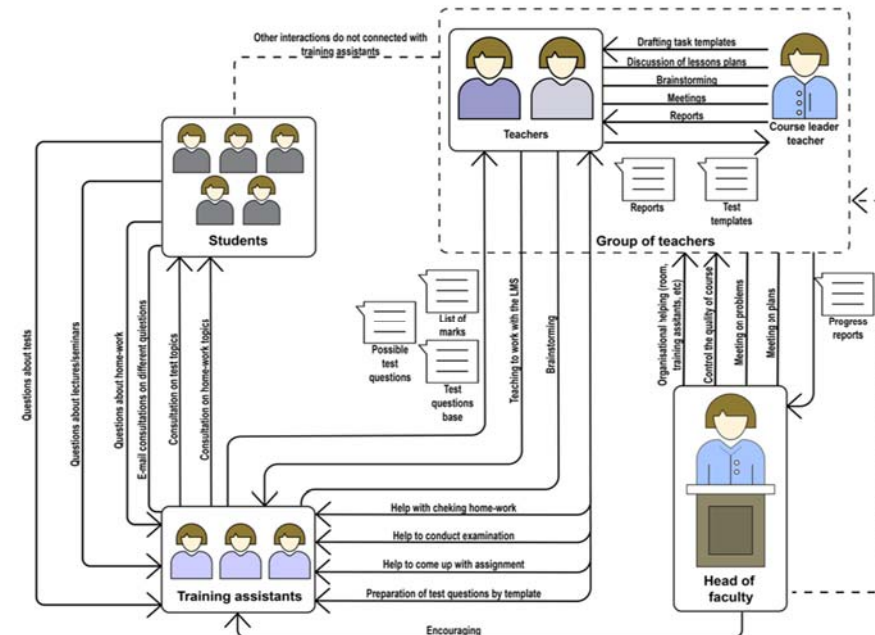


Fig. 1. Rich picture on subject “Teaching assistants’ interactions in “Discrete mathematics”

Table 1. The CATWOE analysis

Role	Description
Clients	Teachers who want to assess their teaching assistants. Students who need assistants' help.
Actors	Groups of teachers who interested in evaluating of skills of teaching assistants and choosing the group of the best teaching assistants. The head of faculty who wants to encourage the best teaching assistant.
Transformation	Teaching assistant receives points for certain evaluation criteria.
World View	It is needed to define a group of the best teaching assistants and the best teaching assistant. The definition of a group of best teaching assistants is necessary in order to reduce the risks associated with incompetent and disinterested teaching assistants with the next year group of teaching assistants.
Owner	Teachers and the head of faculty.
Environmental constraints	National educational and assessment standards.

After analyzing the processes and interactions associated with the members of the system, a clear understanding of the subject area is emerged.

There are three teachers: one lecturer (the leading teacher) and two seminarists at "Discrete mathematics" course. They compose a decision group for choosing best assistants. Fair and reliable evaluation results would be obtained by this group because its members have a strong relationship with teaching assistants during the course.

In order to evaluate the assistants, it is decided to come up with evaluation criteria. After the first brainstorm, the list of criteria is similar to a chaotic list of records. The next meeting of the teachers shows that some of the criteria identified in the first stage for assessing the assistants turned out to be duplicated or unnecessary. After long discussions and joint brainstorming, three main groups of criteria are identified: professional skills, communicating skills, personal qualities.

The professional skills include the following sub-criteria:

- active involvement in the process of forming the program of discipline;
- initiative to compile new types of tasks for tests;
- knowledge of the subject domain;
- quality of homework checking;
- speed of homework checking;
- experience of active use of the LMS;

The communicating skills include the following sub-criteria:

- pedagogical experience, the ability to correctly present information;
- openness to student issues (e.g. quick response to questions, competent answers);
- participation in counseling sessions before the tests and examinations;
- active communicating with teachers, participation in weekly meetings;
- the ability to listen carefully.

The personal qualities include the following sub-criteria:

- ethical compliance;
- punctuality;
- self-motivation, the desire for development;
- responsibility for work;
- teamwork skills;
- subordination;
- striving to achieve common results;
- resistance to conflict situations;
- the ability to generate new and innovative ideas;
- the ability to compromise;
- benevolence.

From the first group the next criteria are deleted:

- active involvement in the process of forming the program of discipline. *The*

teachers should do it, because drawing up a discipline program requires experience and entails a great responsibility;

- knowledge of the subject domain. *Taking into account that each assistant is selected among the best students of the course, this requirement should be fulfilled by default.*

And the next criteria are combined as they characterize the checking of homework and are closely interrelated:

- quality of homework checking;
- speed of homework checking.

From the second group the next criteria are deleted:

- pedagogical experience, the ability to correctly present information. *This ability can be learned. One of the goals of the Teaching Assistant program is the development of teaching skills;*
- the ability to listen. *In our opinion, this parameter is almost impossible to estimate.*

From the third group the next criteria are combined, because they are very interrelated and cannot be separated:

- self-motivation, the desire for development;
- responsibility for work;

And the next criteria are deleted:

- teamwork skills. *It is related with the responsibility of work criteria;*
- ability to be subordinate. *By default, the main person on the course is the teacher. This is necessary to understand at first;*
- striving to achieve common results. *It is related with the responsibility of work criteria;*
- resistance to conflict situations. *It is the responsibility of the teacher to resolve and prevent the emergence of conflict situations;*
- the ability to generate new and innovative ideas. *This is not a paramount task of the teaching assistant. And the teaching assistant can work great, but do not come up with ideas, it's not scary;*
- the ability to compromise. *The last word for the teacher;*
- benevolence. *It is related with the ethical compliance and punctuality of work criteria.*

The final elected criteria and subcriteria are shown in Fig. 2. All the criteria and subcriteria have their own identification numbers.

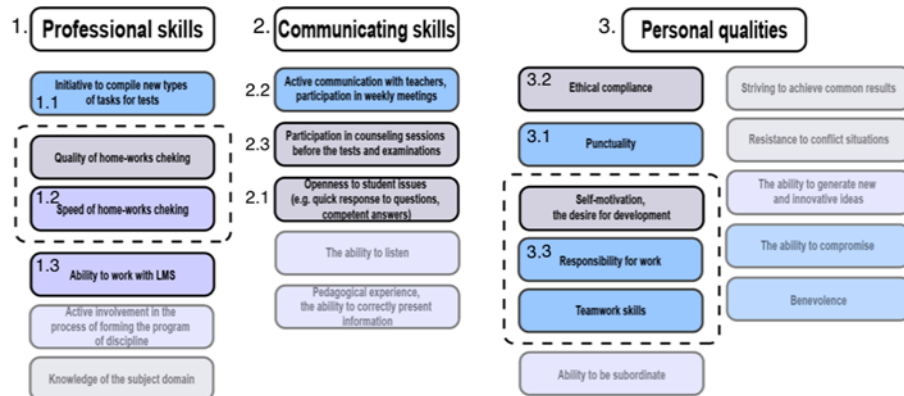


Fig. 2. The final list of criteria.

4. Exploring the alternatives

There are ten teaching assistants *A, B, C, D, E, F, G, H, I, J* on the course.

We can reduce the number of evaluating teaching assistants after assessing the involvement of teaching assistants in educational process.

We have 3 groups of criteria, consisting of 9 sub-criteria. In order to assess the involvement of assistants in the educational process, we did not use the values of the last three subcriteria (3.1-3.3). These sub-criteria refer to a group of personal qualities and cannot be regarded as involvement in the educational process. Then the involvement of the teaching assistant in the educational process for each criterion is evaluated, based on expert judgment. The results are presented in Table 2.

Table 2. The involvement in educational process

	A	B	C	D	E	F	G	H	I	J
1.1.	5	5	5	5	4	5	3	2	1	1
1.2.	4	5	5	4	4	4	4	4	1	4
1.3.	1	2	1	4	5	1	1	1	1	1
2.1.	4	3	3	4	2	5	5	1	3	1
2.2.	5	4	4	4	3	5	2	1	2	2
2.3.	4	4	3	4	1	4	5	2	1	2

Let us understand which assistants are least involved in the work process, according to experts. Calculations of threshold equals to 3, 4 and 5 are shown in Tables 3, 4 and 5.

Table 3. Treshhold is equal to 3

	A	B	C	D	E	F	G	H	I	J
1.1.	1	1	1	1	1	1	1	0	0	0
1.2.	1	1	1	1	1	1	1	1	0	1
1.3.	0	0	0	1	1	0	0	0	0	0
2.1.	1	1	1	1	0	1	1	0	1	0
2.2.	1	1	1	1	1	1	0	0	0	0
2.3.	1	1	1	1	0	1	1	0	0	0

Table 3 and 4 allow to identify teaching assistants who are least involved in the educational process.

The Table 5 with threshold equals to 5 shows that no one from H, I, J is not indispensable.

Table 4. Treshhold is equal to 4

	A	B	C	D	E	F	G	H	I	J
1.1.	1	1	1	1	1	1	0	0	0	0
1.2.	1	1	1	1	1	1	1	1	0	1
1.3.	0	0	0	1	1	0	0	0	0	0
2.1.	1	0	0	1	0	1	1	0	0	0
2.2.	1	1	1	1	0	1	0	0	0	0
2.3.	1	1	0	1	0	1	1	0	0	0

Table 5. Treshhold is equal to 5

	A	B	C	D	E	F	G	H	I	J
1.1.	1	1	1	1	0	1	0	0	0	0
1.2.	0	1	1	0	0	0	0	0	0	0
1.3.	0	0	0	0	1	0	0	0	0	0
2.1.	0	0	0	0	0	1	1	0	0	0
2.2.	1	0	0	0	0	1	0	0	0	0
2.3.	0	0	0	0	0	0	1	0	0	0

Thus, it is decided not to consider further the last three teaching assistants (H, I, J). However, little involvement in the educational process has its own explanations:

- H was ill two month;
- I was out of connection;
- J decided to switch to another faculty. Preparation for the exams took all his spare time.

Thus, seven candidates are remained. It is difficult to find the best one because each of them is successful in one or more criteria.

Stakeholders are about to choose A as a winner because this assistant took part in all teacher meetings and he suggested new types of tasks for tests so regularly (approximately once every two weeks). Assistant A communicated with teachers a lot (flushed before their eyes), that is why they prefer him.

However, this decision can be too unfair, that's why multicriteria decision making (MCDM) process is decided to be applied.

5. Analytical Hierarchy Process

Analytic Hierarchy Process (AHP) which is one of the most used MCDM approaches [3] is a structured multicriteria technique for organizing and analyzing complex decisions including many criteria. In this paper we use a classical AHP proposed by the author [4].

At the first step of AHP a model for the decision is developed. Experts break down the decision into a hierarchy of goals, criteria, sub-criteria and alternatives.

After that, decisioners derive priorities (weights) for the criteria with respect to the desired goal. It is made in the form of pairwise comparisons using individual questionnaires. Since the evaluation criteria are subjective and qualitative in nature, it is very difficult for the decision maker to express the preferences using exact numerical values. That is why a special numerical scale [4] which consists of interpretation of linguistic terms is used (see Table 6).

Table 6. Saaty's pairwise comparison scale

Numeric value	Linguistic terms
1	Equally important
3	Moderately more important
5	Strongly more important
7	Very Strongly more important
9	Extremely important
2, 4, 6, 8	The intermediate values that are used to address situations of uncertainty between the two adjacent judgments

Results of comparisons of all experts are presented in the form of matrices (see Table 7).

Table 7. Criteria pairwise comparisons obtained by experts

	Professional skills			Communicative skills			Personal qualities		
	Exp.1	Exp.2	Exp.3	Exp.1	Exp.2	Exp.3	Exp.1	Exp.2	Exp.3
Professional skills	1	1	1	5	5	4	3	2	1
Communicative skills	1/5	1/5	1/4	1	1	1	1/3	1/4	1/5
Personal qualities	1/3	1/2	1	3	4	5	1	1	1

Before calculating the weights, the consistency of the comparison matrix is checked. As a rule, only if consistency is less than 0.1, it considered as acceptable, otherwise the pair-wise comparisons should be revised. In this decision making process, all of them are less than 0.092 that shows answers are consistence.

On the basis of Table 7 the final matrix is created by finding a mean between estimates of all experts (see Table 8). This metric is used because of solid decision to make all experts' voices to be equal.

Table 8. Aggregate matrix with criteria pairwise comparisons

	Professional skills	Communicative skills	Personal qualities
Professional skills	1	4,66	2
Communicative skills	0,22	1	0,25
Personal qualities	0,6	4	1

The matrix from Table 8 is used in order to calculate criteria priority weights. The same way as it was earlier, a pairwise comparison of all the sub-criteria, with respect to each criterion, included in the decision-making model, is made. Obtained results are shown in Table 9.

Table 9. Criteria and subcriteria priority weights

1. Professional skills	54,772%	1.1. New task types creation	25,232%
		1.2. HW checking	26,068%
		1.3. Experience in LMS	3,472%
2. Communicative skills	10,069%	2.1. Openness to students	2,946%
		2.2. Communication with teachers	1,288%
		2.3. Participation in consultations	5,834%
3. Personal qualities	35,159%	3.1. Punctuality	13,086%
		3.2. Ethical compliance	10,062%
		3.3. Self-motivation	12,011%

Next step consists of deriving the relative priorities (preferences) of the alternatives with respect to each criterion. Overall priority weights of assistants are calculated by summing all local priorities. Final figures are shown in Table 10. Bar chart is built on the basis of overall preferences of the alternatives (see Fig. 3).

Table 10. Local and overall priorities of alternatives

	A	B	C	D	E	F	G
1.1.	8,352%	6,711%	3,784%	2,728%	1,821%	1,214%	0,622%
1.2.	3,140%	6,060%	9,131%	3,836%	1,700%	0,966%	1,234%
1.3.	0,160%	0,249%	0,148%	1,001%	1,618%	0,148%	0,148%

2.1.	0,203%	0,154%	0,101%	0,441%	0,082%	0,803%	1,161%
2.2.	0,444%	0,089%	0,117%	0,062%	0,062%	0,444%	0,072%
2.3.	0,414%	0,387%	0,198%	0,833%	0,144%	1,628%	2,230%
3.1.	1,035%	2,677%	1,921%	3,670%	1,663%	1,041%	1,078%
3.2.	0,254%	1,928%	0,254%	3,822%	1,782%	0,993%	1,029%
3.3.	0,923%	2,040%	0,761%	4,201%	1,640%	1,108%	1,339%
Totals	14,924%	20,296%	16,416%	20,595%	10,010%	8,846%	9,114%

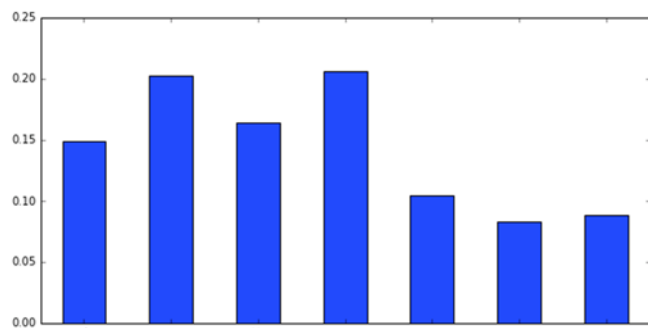


Fig. 3. Overall preferences of the alternatives

6. A Discussion on AHP Results

AHP analysis shows that the prompt decision of choosing A as the best assistant is totally unfair. Results reveal that experts did not take into account other important criteria that in general over weighted those, which were chosen at first. Another discovered problem of A is some of his/her estimates, which are the worst in comparison with others (for instance, criteria 3.1 and 3.2). This fact also decreases his/her chances to be a leader.

The main interesting point of results are the highest figures which belong to both two assistants B and D. Let's describe each of them.

Assistant B cannot be named as a brilliant employee. Nevertheless, he/she has showed good stable work without having bad results in any of the activities during the course. Despite not being the best in any of the criteria, B always was close to the leader.

In the same manner as B, assistant D has shown quite strong results in technical and communicative estimates. In addition, D was on the top in the personal qualities criteria. He/she produces an impression of too self-motivated person and D was never late on any events. Result of D exceeds B at an inconspicuous gap of 0,3. Since experts make an arrangement on having no less than 2% advantage taking by the leader, such difference is admitted being not crucial for them.

In addition, there is a problematic situation with evaluation of the five best assistants. Four employees can be determined more or less clearly (A, B, C, and D). However,

the difference between E and the closest competitor G is less than 1%, which is also insignificant.

7. Fuzzy Type-2 AHP

Since experts want to be more confident in fairness of their choice, we decide to apply another MCDM approach for purpose of aiming our goal. It is called Fuzzy AHP. In classical AHP crisp numbers are used, for pairwise comparison evaluations. However, in Fuzzy AHP, the linguistic variables are represented as fuzzy numbers instead of crisp. In this case a fuzzy logic provides a mathematical strength to capture the uncertainties associated with human cognitive process. Many researchers [5], [6] who have studied the Fuzzy AHP have provided evidence that it shows relatively more sufficient description of decision making processes compared to the traditional AHP methods.

According to [7], the membership functions of type-1 fuzzy sets have no uncertainty associated with it. Type-2 fuzzy sets generalize type-1 fuzzy sets and systems so that more uncertainty for defining membership functions can be handled. That's why type-2 fuzzy logic is used.

A type-2 fuzzy set $\tilde{\tilde{A}}$ in the universe of discourse X can be represented by a type-2 membership function $\mu_{\tilde{\tilde{A}}}$, shown as follows [8]:

$$\tilde{\tilde{A}} = \left\{ \left((x, u), \mu_{\tilde{\tilde{A}}}(x, u) \right) \mid \forall x \in X, \forall u \in J_x \subseteq [0, 1], 0 \leq \mu_{\tilde{\tilde{A}}}(x, u) \leq 1 \right\},$$

where J_x denotes an interval $[0, 1]$.

$\tilde{\tilde{A}}$ is called an interval type-2 fuzzy set if all $\mu_{\tilde{\tilde{A}}} = 1$ [8]. Interval type-2 fuzzy sets are the most commonly used type-2 fuzzy sets because of their simplicity and reduced computational effort with respect to general type-2 fuzzy sets. For this reason, we use interval type-2 fuzzy sets.

A trapezoidal interval type-2 fuzzy set is illustrated as

$$\tilde{\tilde{A}}_i = \left(\tilde{\tilde{A}}_i^U; \tilde{\tilde{A}}_i^L \right) = \left(a_{i1}^U, a_{i2}^U, a_{i3}^U, a_{i4}^U; H_1(\tilde{\tilde{A}}_i^U), H_2(\tilde{\tilde{A}}_i^U) \right), \left(a_{i1}^L, a_{i2}^L, a_{i3}^L, a_{i4}^L; H_1(\tilde{\tilde{A}}_i^L), H_2(\tilde{\tilde{A}}_i^L) \right),$$

where $\tilde{\tilde{A}}_i^U$ and $\tilde{\tilde{A}}_i^L$ are type-1 fuzzy sets, $a_{i1}^U, a_{i2}^U, \dots, a_{i3}^U, a_{i4}^U$ are the references points of the interval type-2 fuzzy set $\tilde{\tilde{A}}_i$; $H_j(\tilde{\tilde{A}}_i^U)$ denotes the membership value of the element $a_{j(j+1)}^U$ in the upper trapezoidal membership function $\tilde{\tilde{A}}_i^U$ and $H_j(\tilde{\tilde{A}}_i^L)$ denotes the membership value of the element $a_{j(j+1)}^L$ in the lower trapezoidal membership function $\tilde{\tilde{A}}_i^L$, $j = \overline{1..2}$ [7].

Pairwise comparison matrices got from experts for AHP are directly applied for our needs in Fuzzy AHP. We introduce interval trapezoidal type-2 fuzzy scales of the linguistic variables (see Table 11). They represent a modified version of scales proposed by [8] and include intermediate values between the two adjacent judgments like in AHP.

The priority weights of criteria (Table 12) and sub-criteria (Table 13) are demonstrated.

Type 2 fuzzy and defuzzified overall weights of the alternatives are shown in Tables 14 and 15. For defuzzification the Defuzzified Trapezoidal Type-2 Fuzzy Set (DTraT) approach is used proposed by [8].

Bar chart is built on the basis of overall preferences of the alternatives (see Fig. 4).

Table 11. Trapezoidal interval type-2 fuzzy scales

Numeric value from AHP	Trapezoidal interval type-2 fuzzy scales
1	(1, 1, 1, 1; 1, 1) (1, 1, 1, 1; 1, 1)
2	(1, 1, 3, 4; 1, 1) (1.2, 1.2, 2.8, 3.8; 0.8, 0.8)
3	(1, 2, 4, 5; 1, 1) (1.2, 2.2, 3.8, 4.8; 0.8, 0.8)
4	(2, 3, 5, 6; 1, 1) (2.2, 3.2, 4.8, 5.8; 0.8, 0.8)
5	(3, 4, 6, 7; 1, 1) (3.2, 4.2, 5.8, 6.8; 0.8, 0.8)
6	(4, 5, 7, 8; 1, 1) (4.2, 5.2, 6.8, 7.8; 0.8, 0.8)
7	(5, 6, 8, 9; 1, 1) (5.2, 6.2, 7.8, 8.8; 0.8, 0.8)
8	(6, 7, 8.5, 9; 1, 1) (6.2, 7.2, 8.3, 8.8; 0.8, 0.8)
9	(7, 8, 9, 9; 1, 1) (7.2, 8.2, 8.8, 9; 0.8, 0.8)

Table 12. Interval type-2 fuzzy weights of criteria

Criteria	Interval type-2 weights
1. Professional skills	(0.275, 0.377, 0.754, 1.005; 1, 1) (0.304, 0.410, 0.706, 0.935; 0.8, 0.8)
2. Communicative skills	(0.057, 0.073, 0.14, 0.211; 1, 1) (0.061, 0.078, 0.13, 0.19; 0.8; 0.8)
3. Personal qualities	(0.188, 0.257, 0.519, 0.708; 1, 1) (0.203, 0.274, 0.477, 0.639; 0.8, 0.8)

Table 13. Interval type-2 fuzzy weights of sub-criteria

Sub-criteria	Interval type-2 weights
1.1.	(0.071, 0.134, 0.447, 0.811; 1, 1) (0.085, 0.154, 0.396, 0.703; 0.8, 0.8)
1.2.	(0.074, 0.138, 0.453, 0.811; 1, 1) (0.088, 0.158, 0.402, 0.705; 0.8, 0.8)
1.3.	(0.013, 0.022, 0.069, 0.124; 1, 1) (0.015, 0.025, 0.061, 0.108; 0.8, 0.8)
2.1.	(0.008, 0.014, 0.062, 0.149; 1, 1) (0.009, 0.016, 0.052, 0.118; 0.8, 0.8)
2.2.	(0.004, 0.007, 0.031, 0.071; 1, 1) (0.004, 0.008, 0.025, 0.055; 0.8, 0.8)
2.3.	(0.014, 0.029, 0.115, 0.24; 1, 1) (0.017, 0.034, 0.099, 0.199; 0.8, 0.8)
3.1.	(0.055, 0.087, 0.212, 0.317; 1, 1) (0.062, 0.095, 0.191, 0.28; 0.8, 0.8)
3.2.	(0.046, 0.069, 0.168, 0.265; 1, 1) (0.051, 0.075, 0.151, 0.23; 0.8, 0.8)
3.3.	(0.046, 0.075, 0.196, 0.317; 1, 1) (0.052, 0.082, 0.175, 0.274; 0.8, 0.8)

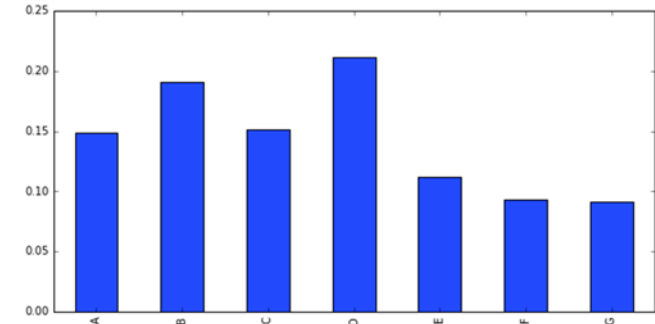


Fig. 4. Overall preferences of the alternatives

Table 14. Local and overall priorities of alternatives A, B, C, D

	A	B	C	D
1.1.	(0.115, 0.212, 0.505, 0.806; 1, 1) (0.133, 0.234, 0.462, 0.7; 0.8, 0.8)	(0.098, 0.174, 0.414, 0.692; 1, 1) (0.113, 0.191, 0.38, 0.612; 0.8, 0.8)	(0.048, 0.091, 0.252, 0.484; 1, 1) (0.06, 0.1, 0.23, 0.4; 0.8, 0.8)	(0.037, 0.063, 0.173, 0.342; 1, 1) (0.042, 0.07, 0.15, 0.3; 0.8, 0.8)
1.2.	(0.04, 0.067, 0.158, 0.272; 1, 1) (0.045, 0.073, 0.14, 0.24; 0.8, 0.8)	(0.083, 0.144, 0.314, 0.471; 1, 1) (0.095, 0.157, 0.3, 0.429; 0.8, 0.8)	(0.186, 0.265, 0.477, 0.657; 1, 1) (0.202, 0.283, 0.45, 0.6; 0.8, 0.8)	(0.067, 0.108, 0.26, 0.406; 1, 1) (0.077, 0.12, 0.24, 0.37; 0.8, 0.8)
1.3.	(0.038, 0.042, 0.055, 0.067; 1, 1) (0.039, 0.043, 0.05, 0.064; 0.8, 0.8)	(0.038, 0.051, 0.082, 0.106; 1, 1) (0.041, 0.054, 0.08, 0.1; 0.8, 0.8)	(0.032, 0.037, 0.051, 0.067; 1, 1) (0.033, 0.04, 0.05, 0.063; 0.8, 0.8)	(0.177, 0.229, 0.382, 0.536; 1, 1) (0.187, 0.24, 0.36, 0.5; 0.8, 0.8)
2.1.	(0.029, 0.044, 0.086, 0.133; 1, 1) (0.032, 0.047, 0.08, 0.12; 0.8, 0.8)	(0.027, 0.038, 0.072, 0.11; 1, 1) (0.029, 0.04, 0.07, 0.1; 0.8, 0.8)	(0.021, 0.026, 0.046, 0.071; 1, 1) (0.022, 0.03, 0.043, 0.06; 0.8, 0.8)	(0.064, 0.104, 0.237, 0.409; 1, 1) (0.072, 0.113, 0.22, 0.36; 0.8, 0.8)
2.2.	(0.185, 0.262, 0.448, 0.587; 1, 1) (0.2, 0.278, 0.426, 0.554; 0.8, 0.8)	(0.041, 0.05, 0.077, 0.106; 1, 1) (0.04, 0.05, 0.07, 0.098; 0.8, 0.8)	(0.044, 0.055, 0.084, 0.111; 1, 1) (0.046, 0.057, 0.08, 0.1; 0.8, 0.8)	(0.038, 0.043, 0.061, 0.076; 1, 1) (0.04, 0.05, 0.06, 0.07; 0.8, 0.8)
2.3.	(0.033, 0.046, 0.091, 0.137; 1, 1) (0.035, 0.05, 0.08, 0.124; 0.8, 0.8)	(0.034, 0.046, 0.083, 0.122; 1, 1) (0.036, 0.049, 0.08, 0.1; 0.8, 0.8)	(0.019, 0.025, 0.046, 0.07; 1, 1) (0.02, 0.026, 0.04, 0.06; 0.8, 0.8)	(0.065, 0.1, 0.22, 0.366; 1, 1) (0.072, 0.108, 0.2, 0.323; 0.8, 0.8)
3.1.	(0.029, 0.042, 0.086, 0.13; 1, 1) (0.03, 0.045, 0.08, 0.116; 0.8, 0.8)	(0.102, 0.161, 0.324, 0.462; 1, 1) (0.113, 0.174, 0.3, 0.425; 0.8, 0.8)	(0.038, 0.052, 0.095, 0.138; 1, 1) (0.041, 0.055, 0.09, 0.126; 0.8, 0.8)	(0.134, 0.215, 0.483, 0.71; 1, 1) (0.153, 0.24, 0.45, 0.649; 0.8, 0.8)
3.2.	(0.015, 0.02, 0.035, 0.053; 1, 1) (0.016, 0.021, 0.03, 0.048; 0.8, 0.8)	(0.068, 0.102, 0.198, 0.295; 1, 1) (0.075, 0.109, 0.2, 0.27; 0.8, 0.8)	(0.015, 0.02, 0.035, 0.053; 1, 1) (0.02, 0.02, 0.03, 0.05; 0.8, 0.8)	(0.163, 0.269, 0.543, 0.78; 1, 1) (0.184, 0.3, 0.51, 0.72; 0.8, 0.8)
3.3.	(0.05, 0.059, 0.09, 0.115; 1, 1) (0.052, 0.061, 0.09, 0.108; 0.8, 0.8)	(0.061, 0.089, 0.23, 0.325; 1, 1) (0.07, 0.099, 0.2, 0.294; 0.8, 0.8)	(0.038, 0.047, 0.081, 0.115; 1, 1) (0.04, 0.05, 0.076, 0.105; 0.8, 0.8)	(0.184, 0.262, 0.492, 0.632; 1, 1) (0.2, 0.28, 0.46, 0.59; 0.8, 0.8)
Total fuzzy weight	(0.018, 0.052, 0.372, 1.069; 1, 1) (0.023, 0.064, 0.3, 0.826; 0.8, 0.8)	(0.026, 0.074, 0.496, 1.336; 1, 1) (0.034, 0.091, 0.4, 1.044; 0.8, 0.8)	(0.023, 0.061, 0.384, 1.062; 1, 1) (0.03, 0.074, 0.32, 0.83; 0.8, 0.8)	(0.035, 0.09, 0.552, 1.46; 1, 1) (0.045, 0.11, 0.454, 1.14; 0.8, 0.8)
Total defuz. weight	0.331	0.426	0.337	0.471
Total norm. defuz. weight	14.868%	19.119%	15.124%	21.128%

Table 15. Local and overall priorities of alternatives E, F, G

	E	F	G
1.1.	(0.025, 0.042, 0.123, 0.257; 1, 1) (0.029, 0.046, 0.109, 0.212; 0.8, 0.8)	(0.019, 0.03, 0.082, 0.161; 1, 1) (0.021, 0.033, 0.073, 0.134; 0.8, 0.8)	(0.013, 0.018, 0.04, 0.077; 1, 1) (0.014, 0.019, 0.036, 0.065; 0.8, 0.8)
1.2.	(0.029, 0.046, 0.106, 0.192; 1, 1) (0.033, 0.05, 0.097, 0.165; 0.8, 0.8)	(0.021, 0.028, 0.055, 0.091; 1, 1) (0.022, 0.03, 0.05, 0.08; 0.8, 0.8)	(0.024, 0.033, 0.076, 0.131; 1, 1) (0.026, 0.036, 0.068, 0.111; 0.8, 0.8)
1.3.	(0.283, 0.378, 0.56, 0.674; 1, 1) (0.302, 0.4, 0.532, 0.648; 0.8, 0.8)	(0.032, 0.037, 0.051, 0.067; 1, 1) (0.033, 0.038, 0.049, 0.063; 0.8, 0.8)	(0.032, 0.037, 0.051, 0.067; 1, 1) (0.033, 0.038, 0.049, 0.063; 0.8, 0.8)
2.1.	(0.015, 0.02, 0.037, 0.06; 1, 1) (0.016, 0.021, 0.034, 0.053; 0.8, 0.8)	(0.128, 0.193, 0.392, 0.613; 1, 1) (0.14, 0.208, 0.363, 0.553; 0.8, 0.8)	(0.177, 0.282, 0.546, 0.772; 1, 1) (0.198, 0.304, 0.509, 0.715; 0.8, 0.8)
2.2.	(0.038, 0.043, 0.061, 0.076; 1, 1) (0.039, 0.045, 0.058, 0.072; 0.8, 0.8)	(0.226, 0.292, 0.461, 0.589; 1, 1) (0.239, 0.306, 0.44, 0.559; 0.8, 0.8)	(0.042, 0.049, 0.068, 0.086; 1, 1) (0.044, 0.05, 0.066, 0.081; 0.8, 0.8)
2.3.	(0.014, 0.017, 0.031, 0.049; 1, 1) (0.014, 0.018, 0.029, 0.044; 0.8, 0.8)	(0.135, 0.2, 0.409, 0.6; 1, 1) (0.148, 0.216, 0.377, 0.543; 0.8, 0.8)	(0.179, 0.272, 0.537, 0.747; 1, 1) (0.2, 0.295, 0.5, 0.693; 0.8, 0.8)
3.1.	(0.064, 0.094, 0.194, 0.316; 1, 1) (0.07, 0.1, 0.179, 0.28; 0.8, 0.8)	(0.046, 0.064, 0.128, 0.209; 1, 1) (0.049, 0.068, 0.118, 0.184; 0.8, 0.8)	(0.05, 0.067, 0.13, 0.197; 1, 1) (0.053, 0.071, 0.12, 0.175; 0.8, 0.8)
3.2.	(0.088, 0.146, 0.327, 0.538; 1, 1) (0.099, 0.159, 0.3, 0.48; 0.8, 0.8)	(0.047, 0.067, 0.145, 0.263; 1, 1) (0.051, 0.072, 0.132, 0.227; 0.8, 0.8)	(0.056, 0.077, 0.145, 0.225; 1, 1) (0.06, 0.082, 0.135, 0.203; 0.8, 0.8)
3.3.	(0.111, 0.129, 0.195, 0.233; 1, 1) (0.117, 0.135, 0.186, 0.222; 0.8, 0.8)	(0.059, 0.074, 0.131, 0.175; 1, 1) (0.063, 0.078, 0.122, 0.159; 0.8, 0.8)	(0.064, 0.078, 0.137, 0.167; 1, 1) (0.067, 0.082, 0.127, 0.154; 0.8, 0.8)
Total weight	(0.021, 0.049, 0.283, 0.789; 1, 1) (0.026, 0.059, 0.233, 0.604; 0.8, 0.8)	(0.015, 0.035, 0.227, 0.681; 1, 1) (0.018, 0.042, 0.183, 0.509; 0.8, 0.8)	(0.016, 0.037, 0.232, 0.652; 1, 1) (0.019, 0.044, 0.188, 0.496; 0.8, 0.8)
Total defuzzy weight	0.251	0.208	0.205
Total norm. defuzzy weight	11.347%	9.334%	9.079%

8. Discussion on Fuzzy-Type-2 AHP Results

Now, we see that assistant D has higher priority weight than B and difference between them (2%) is suitable for experts. In addition, it can be noticed that E should be in the top five group, for sure (difference is also about 2%). Thus, Fuzzy AHP does not change ranks of alternatives but makes it clearer. It means that more reliable results are maintained since interval type-2 fuzzy sets can better represent uncertainties.

It is important to note that, contrary to the common belief, the system does not determine the decision we should make, rather, the results should be interpreted as a blueprint of preference and alternatives based on the level of importance obtained for the different criteria taking into consideration our comparative judgments. In other words, the AHP methodology allows us to determine which alternative is the most consistent with our criteria and the level of importance that we give them.

Taking this point into account, Sensitivity Analysis is used. It performs a “what-if” analysis to see how the final results would have changed if the weights of the criteria would have been different [9].

Let’s start with a goal of finding the best teaching assistant. The first criterion has the highest weight in our results ($\approx 50\%$). If we decrease its weight and proportionally

increase other weights then D will still be a leader. In this case D will have even more clear-cut victory. Otherwise, if we increase weight of this criterion up to 60% and more, then B will become a new leader. However, stakeholders come to one opinion that no one criterion should cost more than a half and they have highlighted that the first criterion (professional skills) should stay as the most important one.

It means that weight of the first criterion should be in the next approximate range [33%; 50%].

Let’s now tune proportions of the second and the third criteria. Calculations show that D can stop be a winner only and only if the third criterion will cost more than the second. Thus, this point was brought to expert discussion and they have unanimously decided that personal qualities (third criterion) should be appreciated higher than communicative ones.

Another important note is change of proportions of subcriteria inside their criteria. There are no strong disputes about subcriteria weights, experts’ opinions differ no more than 10%. In this case change of subcriteria preferences in that range does not influence on the leader.

It means that there is no opportunity to have another leader than D by introducing small changes in current proportions of criteria weights.

At the same time, there is a complex situation with choosing top five assistants group. Analysis shows that four assistants are determined clearly. They are A, B, C, and D. The fifth assistant can be either E or G.

Calculations reveal that position of assistant G is directly connected with the second criteria and if its weight is equal or more than 15% than G will be in top five group instead of E. However, now second criterion has only nearly 10%.

Finally, after Sensitivity Analysis is done, next recommendations for the experts are given:

- to choose assistant D as a winner;
- to prolongate contracts with A, B, C and D;
- to prolongate contract with E if experts think that personal qualities should be at least twice more important than communicating skills (finally, communicating skills should have a weight less than 15);
- to prolongate contract with G, in other case.

9. Final Result and Conclusion

Taking into consideration recommendations mentioned above, group of teachers has decided to follow first two instructions. They have selected D as the best teaching assistant on the course of «Discrete mathematics». Also, they have prolonged contracts with D and A, B, and C assistants.

The main important step now is to choose the fifth assistant. Before making a choice, experts decide to use a retrospective and to look through all methods that were applied earlier. Lecture of the course noticed that since A, B, C, and D assistants are already

confirmed it means that nobody will be responsible for communication with students (answering questions, having consultations) because assistant F did it before. However, now there is a choice between either E or G. And in this case G demonstrates a clear superiority compared with others as he/she is one of the top in this kind of work. Finally, G is chosen.

At the very beginning teachers wanted to choose assistant A as the best teaching assistant. However, the soft methods of analysis helped us to choose another assistant. Also, neither AHP nor Fuzzy AHP chose G teaching assistant as the 5th best assistant in the group. Only a sound logic helped us to do this.

The application of methods of system analysis can help to make a decision but it does not make a choice for us. We should look carefully at the results of system analysis methods, but make a balanced and considered decision.

References

- [1] J. Lopa. Using Undergraduate Students as Teaching Assistants. *The Professional & Organizational Development Network in Higher Education*, vol. 21, 2009, pp. 50-62.
- [2] M. Asad, S. Kermani and H. Hora. A Proposed Framework for Evaluating Student's Performance and Selecting the Top Students in E-Learning System, Using Fuzzy AHP Method. In *Proceedings of the International Conference on Management, Economics and Humanities*, Istanbul-Turkey, 2015.
- [3] T. Saaty. There is no mathematical validity for using fuzzy number crunching in the analytic hierarchy process. *Journal of Systems Science and Systems Engineering*, vol. 15, no. 4, 2006, pp. 457-464.
- [4] Monk and S. Howard. The Rich Picture: A Tool for Reasoning about Work Context. [Online]. Available: <http://www-users.york.ac.uk/~am1/RichPicture.pdf>, accessed 12.05.2018.
- [5] Changing Minds. CATWOE. [Online]. Available: <http://creatingminds.org/tools/catwoe>, accessed 12.05.2018.
- [6] M. Velasquez and P. Hester. An Analysis of Multi-Criteria Decision Making Methods. *International Journal of Operations Research*, vol. 10, no. 2, 2013, pp. 56-66.
- [7] T. Saaty. *The Analytic Hierarchy Process*, New York: McGraw Hill, 1980.
- [8] C. Boender and J. De Graan. Multicriteria Decision Analysis with Fuzzy Pairwise Comparisons. *Fuzzy Sets and Systems*, vol. 29, 1989, pp. 133-143.
- [9] D. Chang. Applications of the Extent Analysis Method on Fuzzy-AHP. *European Journal of Operational Research*, vol. 95, 1996, pp. 649-655.
- [10] J. Mendel and R. John. Type-2 fuzzy sets made simple. *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 2, 2002, pp. 117-127.
- [11] C. Kahraman and B. Öztayşi. Fuzzy analytic hierarchy process with interval type-2 fuzzy sets. *Knowledge-Based Systems*, vol. 59, 2014, pp. 48-57.
- [12] E. Mu and M. Pereyra-Rojas. *An Introduction to the Analytic Hierarchy Process (AHP). Using Super Decisions*, vol. 2, New York: Springer, 2016.

Применение методов системного анализа к оцениванию работы учебных ассистентов

Е.Н. Береснева <eberesneva@hse.ru>

М.К. Горденко <mgordenko@hse.ru>

Департамент программной инженерии,

Национальный исследовательский университет "Высшая школа экономики",
101000, Россия, г. Москва, ул. Мясницкая, д. 20

Аннотация. В этой статье представлены результаты применения различных методов системного анализа (CATWOE, Rich Picture, AHP, Fuzzy AHP) для оценки учебных ассистентов для преподавателей. В статье рассмотрено применение soft- и hard-методов системного анализа. Методы системного анализа рассматриваются на примере реализации программы «Учебный ассистент» в Национальном Исследовательском Университете «Высшая школа экономики» (НИУ ВШЭ) на дисциплине «Дискретная математика». В статье показан процесс взаимодействия преподавателей с учениками и преподавателями в форме Rich Picture. Определены связующие мероприятия, встречи и даже отчеты, которые предоставляют ассистенты преподавателю. Затем описано каким образом были выбраны критерии для оценки ассистентов и оценена важность каждого критерия. Были определены три группы критериев: профессиональные навыки, навыки общения, личные качества. Каждая группа имеет некоторые подкритерии, которые были определены посредством уточняющих встреч и мозгового штурма. Также в работе был определен собственный метод оценки, который явился пререквизитом для AHP и позволивший сразу же отбросить наиболее неперспективных ассистентов. Кроме того, рассматривается применение методов AHP и Fuzzy AHP типа 2 для оценки учебных ассистентов. Выявлены сильные и слабые стороны каждого метода. Также показано, что, несмотря на мощь методов системного анализа, необходимо использовать здравый смысл и логику. Нельзя полагаться только на числа, полученные методами системного анализа, необходимо затем производить анализ результатов. В процессе работы выбирается лучший учебный ассистент, и определяется группа лучших учебных ассистентов.

Ключевые слова: системный анализ; комбинация soft и hard методов; многокритериальное принятие решений; AHP; Fuzzy Type-2 AHP; нечеткие множества.

DOI: 10.15514/ISPRAS-2018-30(3)-18

Для цитирования: Береснева Е.Н., Горденко М.К. Применение методов системного анализа к оцениванию работы учебных ассистентов. *Труды ИСП РАН*, том 30, вып. 3, 2018 г., стр. 251-270 (на английском языке). DOI: 10.15514/ISPRAS-2018-30(3)-18

Список литературы

- [1] J. Lopa. Using Undergraduate Students as Teaching Assistants. *The Professional & Organizational Development Network in Higher Education*, vol. 21, 2009, pp. 50-62.
- [2] M. Asad, S. Kermani and H. Hora. A Proposed Framework for Evaluating Student's Performance and Selecting the Top Students in E-Learning System, Using Fuzzy AHP

- Method. In Proceedings of the International Conference on Management, Economics and Humanities, Istanbul-Turkey, 2015.
- [3] T. Saaty. There is no mathematical validity for using fuzzy number crunching in the analytic hierarchy process. *Journal of Systems Science and Systems Engineering*, vol. 15, no. 4, 2006, pp. 457-464.
- [4] Monk and S. Howard. The Rich Picture: A Tool for Reasoning about Work Context. [Online]. Available: <http://www-users.york.ac.uk/~aml/RichPicture.pdf>, accessed 12.05.2018.
- [5] Changing Minds. CATWOE. [Online]. Available: <http://creatingminds.org/tools/catwoe>, accessed 12.05.2018.
- [6] M. Velasquez and P. Hester. An Analysis of Multi-Criteria Decision Making Methods. *International Journal of Operations Research*, vol. 10, no. 2, 2013, pp. 56-66.
- [7] T. Saaty. *The Analytic Hierarchy Process*, New York: McGraw Hill, 1980.
- [8] C. Boender and J. De Graan. Multicriteria Decision Analysis with Fuzzy Pairwise Comparisons. *Fuzzy Sets and Systems*, vol. 29, 1989, pp. 133-143.
- [9] D. Chang. Applications of the Extent Analysis Method on Fuzzy-AHP. *European Journal of Operational Research*, vol. 95, 1996, pp. 649-655.
- [10] J. Mendel and R. John. Type-2 fuzzy sets made simple. *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 2, 2002. pp. 117-127.
- [11] C. Kahraman and B. Öztayşi. Fuzzy analytic hierarchy process with interval type-2 fuzzy sets. *Knowledge-Based Systems*, vol. 59, 2014, pp. 48-57.
- [12] E. Mu and M. Pereyra-Rojas. *An Introduction to the Analytic Hierarchy Process (AHP). Using Super Decisions*, vol. 2, New York: Springer, 2016.